[00166] Figure 17 shows an alternative method of producing a polynucleotide having defined 3' and 5' sequences from a RNA target and a chimeric oligonucleotide primer. A chimeric oligonucleotide extension product is bound to a solid support through sequence B' and is used to generate a bound polynucleotide which comprises from its 5' end, a defined sequence (B), a sequence representative of a target polynucleotide, a sequence (A') and a sequence (C').

[00167] Figure 18 illustrates an isothermal amplification using a composite primer and the bound polynucleotide product produced in Figure 17.

[00168] Figure 19 shows an alternative method of producing a polynucleotide having defined 3' and 5' sequences from a DNA target and a chimeric oligonucleotide primer. A chimeric oligonucleotide extension product is bound to a solid support through sequence B' and is used to generate a bound polynucleotide which comprises from its 5' end, a defined sequence (B), a sequence representative of a target polynucleotide, a sequence (A') and a sequence (C').

DETAILED DESCRIPTION OF THE INVENTION

[00169] The invention provides methods, compositions, and kits useful in the analysis of nucleic acids. Some aspects of the invention relate to the preparation of polynucleotides comprising ligands and binding such ligands to solid surfaces for analysis and for archiving. The polynucleotides bound to the surface are generally produced in a manner that allows for them to be readily manipulated, for example, by incorporating defined 3' and/or 5' regions into the polynucleotides. The bound polynucleotides can then treated by the methods described herein to produce amplified product. The methods of the invention provide for producing amplified product with defined 3' and/or 5' ends. The amplified products having defined 3' and 5' ends can be used for further manipulation and analysis. They can be used, for example in molecular inversion probe (MIP) analysis.

[00170] The polynucleotides and amplified product of the present invention generally comprise sequences that are related to (i.e. either equivalent to or complementary to) the target RNA or target DNA from which they are derived. In some embodiments, the methods of the invention can be used for the global isolation and amplification of target RNA or DNA, e.g. messenger RNA (mRNA) or genomic DNA. Thus, the methods can be used to produce a plurality of polynucleotides and/or amplification products having sequence related to the target RNA or DNA, and also having defined, e.g. universal, sequences at their 3' and/or 5' ends. This plurality of polynucleotides and or amplification products can be representative of the target nucleic acid or subset of the target nucleic acid such as to comprise a library of representative sequences.

[00171] One aspect of the invention relates to the clonal amplification of nucleic acid sequences of interest. The polynucleotides of the present invention attached to solid substrates can be attached in a manner in which polynucleotides of different specific sequences representative of the target nucleic acid are isolated from one another, for example, each attached to a different bead, or each attached to an isolated area on a surface. In one embodiment, a plurality of polynucleotides, each comprising a specific sequence is bound to a plurality of beads or isolated areas such that only one copy of each polynucleotide is bound to each bead or each isolated surface area. These isolated polynucleotides can then be clonally amplified such that, for example each bead or surface area is retained within an isolated volume. This can be accomplished with beads, for example, by using a

IDT Exhibit 1020.0806

microemulsion in which, on average, each droplet comprises one or fewer beads per droplet. The methods of the present invention allow for the archiving of the polynucleotides bound to the beads and the subsequent analysis of the polynucleotides, for example, using clonal amplification.

[00172] In one aspect of the invention, the polynucleotide bound to the surface is produced in a manner that is amenable to the sequencing of the polynucleotide, thus revealing information about the target nucleic acid from which it is derived.

Method for generating polynucleotide bound to a solid surface

[00173] One aspect of the invention is a method for attaching a polynucleotide sequence that is representative of a sequence within a nucleic acid target molecule to a solid surface. The polynucleotide sequence that is produced is representative of the sequence within a nucleic acid target molecule if it is either the same as, or complementary to the sequence within the target nucleic acid. Where the target nucleic acid is double stranded, the method can produce sequences that are representative of both of the strands. The polynucleotide can be, for example either DNA or RNA.

[00174] The first step of the method comprises: (a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P), is complementary to a target nucleic acid, and a 5' portion of the of the primer, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid. The first primer may be extended by a DNA polymerase such as an RNA-dependent DNA polymerase for extending the first primer along an RNA target nucleic acid, or a DNA-dependent DNA polymerase for extending the first primer along a DNA target nucleic acid. In some embodiments, the 3' portion of the primer that is complementary to the target nucleic acid is a specific sequence. For example, where a specific region of interest of a target nucleic acid that is known or suspected to be upstream of a specific sequence on the target nucleic acid, sequence (P) of the composite primer can be designed to hybridize to this specific sequence on the target nucleic acid such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target nucleic acid comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence (P) can comprise a sequence that can hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence (P) comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence such as sequence (P) at the 3' end of the primer can be useful for performing a global amplification of a nucleic acid target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target nucleic acid. In some embodiments, the relative amounts of the various sequences can be

IDT Exhibit 1020.0807

used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample, or to determine gene copy number in a DNA sample.

[00175] The first primer extension product comprises a 5' portion comprising sequence (A). Sequence A comprises RNA. In some embodiments, sequence (A) is RNA, and sequence (P) is DNA. In other embodiments, sequence (A) may comprise some (DNA). In some embodiments, sequence (P) may comprise some RNA. In some embodiments, sequence (A) and sequence (P) are adjacent.

[00176] The method further comprises step (b) separating or removing the first primer extension product from the target nucleic acid. The first primer extension product can be separated from the target nucleic acid by a variety of methods. In some cases the separation can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. Other methods of separating the first primer extension product from the target nucleic acid involve selectively cleaving or degrading the target nucleic acid. Where the target nucleic acid is RNA, the cleaving or degrading can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or chemically, for example with the addition of alkali. In some embodiments, the target nucleic acid is completely cleaved or degraded. In other embodiments, the target nucleic acid is only partly cleaved or degraded. The amount of cleavage or degradation required is that amount which allows the extension of the second primer. In some embodiments, the cleavage or degradation is carried out partially, and the fragments of the target nucleic acid that remain can constitute the second primer for step (c).

[00177] The method further comprises step (c) extending a second primer to produce a double stranded product comprising a second primer extension product hybridized to the first primer extension product, wherein the second primer comprises a 3' segment complementary to a portion of the first primer extension product and a ligand, whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. The second primer can be a tailed primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (B), which is not complementary to the first primer extension product. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA or target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target nucleic acid. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second

IDT Exhibit 1020.0808

primer may ensure amplification of the entire selected target or family of selected targets. In some cases, random primer by the second primer in combination with extension by a polymerase having substantial strand-displacement activity may ensure amplification of the entire selected target, an entire transcriptome, an entire genome, an entire family of selected targets, or a substantial portion thereof. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target nucleic acid (sense copies), for example, the second primer extension products or the amplicons therefrom may comprise sequences that are 95%, 99%, 99.9%, 99.999%, 99.9999%, 99.999999% or more identical to the sequences in the target nucleic acid or their complement. The second primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody.

[00178] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded.

[00179] The method further comprises step (d) binding the ligand to a receptor bound to a solid surface whereby the second primer extension product is attached to the solid surface. The solid surface may comprise a bead or a set of beads, a magnetic bead or a set of beads, a substantially planar array, a well in a plate, a series of wells in a plate, an isolated surface, or a set of isolated surfaces. The ligand-receptor pair may comprise any pair of binding agents that are capable of specifically associating. For example, the ligand receptor pair may be avidin and biotin or biotin and avidin respectively. Similarly, the ligand-receptor pair may be any pair of proteins, peptides or small molecules that bind to each other. Additionally, the ligand-receptor pair may comprise nucleic acid sequences such a RNA, DNA, peptide nucleic acids or their analogs that bind or hybridize specifically. Other examples of receptors include aptamers, antibodies, affibodies, enzymes, or any protein, peptide, macromolecule, or small molecule specifically capable of binding to a ligand. Other examples of ligands include any protein, peptide, macromolecule, or small molecule specifically capable of binding to a receptor.

[00180] The receptor is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension product to the solid surface. In some embodiments, the second primer extension product is still hybridized to the first primer extension product when it is attached to the solid surface. In some embodiments, the first primer extension product is removed from the second primer extension product such that a single stranded polynucleotide is attached to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B) at or near its 5' end.

IDT Exhibit 1020.0809

[00181] One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00182] Step (d) of binding the polynucleotides to the solid surface through the ligand can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00183] A schematic exemplary of an embodiment of the invention relating to method for generating polynucleotide bound to a solid surface is shown in Figure 1. Step I shows the extension of a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P), is complementary to a target nucleic acid and a 5' portion of the of the primer, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid. The target nucleic acid can be, for example, DNA or RNA. The target nucleic acid as shown is single stranded. The target nucleic acid may be single stranded in the sample, or can be double stranded, and rendered single stranded, for example, but denaturation with heat. The sequence (P) can represent a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step II illustrates the separating or removing of the first primer extension product from the target nucleic acid. The separation or removing can be accomplished for instance by denaturation of the complex, or by selective degradation of the target nucleic acid. For example, where the target nucleic acid is RNA, the separation/removing/degradation can be accomplished by alkali, or by enzymatic cleavage, for example by RNase H. Step III illustrates extending a second primer to produce a double stranded product comprising a second primer extension product hybridized to the first primer extension product, wherein the second primer comprises a 3' segment complementary to a portion of the first primer extension product and a ligand, whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. In the embodiment shown the second primer comprises a sequence (B) that is 5' to the segment that is complementary to the first primer extension product, and the ligand is attached 5' to the sequence (B). This step requires both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity to product the

IDT Exhibit 1020.0810

second primer extension product. Step V illustrates binding of the ligand to a receptor bound to a solid surface whereby the second primer extension product is attached to the solid surface. In the embodiment shown, the second primer extension product is bound to the solid substrate in double stranded form while still hybridized to the first primer extension product. In some embodiments, the second primer extension product is rendered single stranded before binding of the ligand to the solid surface. The bound second primer extension product comprises defined sequence (B) at its 5' end and defined sequence (A') at its 3' end. The bound second primer extension product can be stored or archived, and later used for analysis and amplification, for example clonal amplification as described herein.

Method for generating a polynucleotide comprising a ligand for binding to a solid surface from an RNA target

[00184] The invention provides methods, compositions and kits for copying, storing, and amplifying polynucleotides having sequences related to target ribonucleic acid (RNA) sequences. The methods provide for amplification of a single RNA species or pool of RNA species. The methods are suitable for, for example, generation of libraries, including cDNA libraries. The methods can generate single stranded RNA or DNA products, which are readily suitable for multiplex analysis by microarray technologies, as well as electrophoresis-based technologies such as differential display, and for sequencing.

[00185] The methods of the invention can copy, store, and amplify of one or more species of RNA, such as a pool of RNA sequences, and is most particularly suitable for the amplification of all RNA (such as whole transcriptome or total RNA) sequences in a biological sample. Thus, one of the major advantages of the methods of the invention is the ability to copy, store, and amplify an entire pool of sequences, which is essential for the ability to analyze the gene expression profile in cells, such as the cells in a biological sample of interest. The methods of the invention have the potential of amplifying a multiplicity, a large multiplicity, and in some embodiments all RNA (such as whole transcriptome or total RNA in a sample) sequences in a sample.

[00186] Insofar as many mRNAs have a unique polyA 3'-end, the amplification initiated from the 3'-end sequence of mRNAs is most common for preparation of cDNA libraries and subsequent sequence analysis for determination of gene expression profiling or other applications. The methods of the invention are similarly suited for preparation of libraries of amplified 3'-portions of mRNAs. The sequence of the first primer used in the methods of invention can be designed to be complementary to a multiplicity, or all, of the mRNA species in the sample by using random sequences, according to methods known in the art. The methods are also useful for whole transcriptome amplification. The methods of the invention can be used for the total RNA in samples such as viral RNA.

[00187] The method for generating a polynucleotide comprising a ligand for binding to a solid surface from an RNA target comprise the step of (a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. In some embodiments, the 3' portion of the primer that is complementary to the target RNA is a

IDT Exhibit 1020.0811

specific sequence. For example, where a specific region of interest of a target RNA that is known or suspected to be upstream of a specific sequence on the target RNA, the sequence that is complementary to the target RNA of the first primer can be designed to hybridize to this specific sequence on the target RNA such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target RNA comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence that is complementary to the target RNA can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence that is complementary to the target RNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a target RNA, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target RNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample. In some embodiments more than one type of sequence that is complementary to the target RNA can be used, for instance both a primer with a random sequence and a primer, or combination of primers with a specific sequence complementary to RNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00188] The method further comprises the step of: (b) cleaving the target RNA from the RNA/DNA hybrid. In some embodiments, the cleaving of the target RNA from the RNA/DNA hybrid involves selectively cleaving or degrading the target RNA. In some cases the cleaving can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. The cleaving can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or a combination of RNase enzymes, or chemically. In some embodiments, the target RNA is completely cleaved. In other embodiments, the target RNA is only partly cleaved or degraded. The amount of cleaving required is that amount which will allow the extension of the second primer. In some embodiments, the cleaving is carried out partially, and the fragments of the target RNA that remain can constitute the second primer for step (c).

[00189] The method further comprises: step (c) extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and

IDT Exhibit 1020.0812

whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. The second primer can be a tailed primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (B), which is not complementary to the first primer extension product. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target RNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target RNA(sense copies). The second primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody.

[00190] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded. The method produces a nucleic acid that comprises a ligand allowing it to be bound to a solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid attached to the ligand also comprises sequence (B) at or near its 5' end.

[00191] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total

IDT Exhibit 1020.0813

nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in copending applications.

[00192] In some embodiments, the method further comprises step (d) binding the ligand to a receptor bound to a solid surface whereby the second primer extension product is attached to the solid surface. The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension product to the solid surface. In some embodiments, the second primer extension product is still hybridized to the first primer extension product when it is attached to the solid surface. In some embodiments, the first primer extension product is removed from the second primer extension product such that a single stranded polynucleotide is attached to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification.

[00193] In some embodiments, the amplification of the second primer extension product comprises the steps of: (e) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded; (f) annealing an amplification primer to the single stranded portion of the second primer extension product complementary to sequence (A), wherein the amplification primer has a DNA portion and a 5' RNA portion; (g) extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product hybridized to the second primer extension product; (h) cleaving the RNA from the amplified product hybridized to the second primer extension product; and (i) repeating steps (f) to (h) to produce multiple copies of amplified product. Where this method of amplification is used, and a sequence (B) is incorporated into the second primer extension product as described above, amplification product is generated which comprises sequence (B') complementary to sequence B at or near its 3' end.

[00194] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target RNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target RNA.

[00195] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some

IDT Exhibit 1020.0814

embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00196] The step of binding the polynucleotides to the solid surface through the ligand, step (d), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00197] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary the second primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (A') on the second primer extension product for the hybridization of primer. In some embodiments, the amplification produces single stranded amplified product, In some embodiments, the amplification provides double stranded product. In some embodiments, the second primer comprises a specific sequence (B), which becomes incorporated into the second primer extension product. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the second primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (A') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (A') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00198] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00199] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer second primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the second primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in

IDT Exhibit 1020.0815

microreactors can be performed on a sample of target RNA, such as whole transcriptome or total RNA, wherein the plurality of second primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00200] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single second primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00201] A schematic exemplary of an embodiment of the invention relating to method for generating a polynucleotide comprising a ligand for binding to a solid surface from an RNA target is shown in Figure 2. Figure 2 shows a target RNA and a chimeric RNA/DNA first primer. The primer is first annealed to the target RNA. Step I illustrates extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. The sequence complementary to a target RNA can be a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step II represents cleaving the target RNA from the RNA/DNA hybrid. Cleaving can be accomplished thermally, chemically, or enzymatically, e.g. with RNase H. The second primer comprising a ligand is then annealed to the first primer extension product. Step III(a) and step III(b) illustrate extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. Step III(a) shows step III for a second primer that does not comprise sequence (B). Step III(b) shows step III for a second primer that comprises a sequence (B) that is 5' of the segment complementary to a portion of the first primer extension product. In each case the second primer is shown with a ligand attached to the 5' end. The second primer extension product comprising a ligand and a defined sequence (A) at its 3' end is useful for storage, archiving and analysis as it has a ligand capable of binding to a solid surface. Such second primer extension product also comprises a sequence that is representative of (identical to or substantially identical to) a sequence in the target RNA, so analysis of this product provides information about the target RNA. Step IV shows the binding the ligand to a solid surface, whereby the second primer extension product becomes bound to the solid surface.

IDT Exhibit 1020.0816

Method for generating a polynucleotide comprising a ligand for binding to a solid surface from a DNA target

[00202] The methods of the present invention can be used to analyze the DNA (e.g. genomic DNA) samples that are important for many studies. The methods can be used for high-throughput genomic analysis, and can be used for forensic and paleoarcheology work which can be severely limited by nucleic acid sample size. The methods can be used, for example, for the genotyping of multiple loci in the study of complex diseases. The methods can also be used for the determination of genomic instability in various pathological conditions such as cancer, which is most precisely carried out in well defined cell populations, such as that obtained by laser capture micro-dissection or cell sorting. The DNA amplification technologies described herein provide global amplification of very small polynucleotide samples, for example, from one or a very few cells.

[00203] One aspect of the invention is a method for generating from a DNA target a polynucleotide comprising a ligand for binding to a solid surface.

[00204] The method comprises the steps of: (a) denaturing a double-stranded target DNA. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00205] The method further comprises step (b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end. The enzyme that carries out step (b) is generally a DNA polymerase. In some cases a mixture of DNA polymerases can be used. The sequence that is complementary to the target DNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer as sequence can be useful for performing a global amplification of a DNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target DNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the number of gene copies in a DNA sample, or obtaining sequence information. In some embodiments, the extension of one first primer, will result in the release of a downstream first primer extension product. This can occur throughout the target DNA resulting in the release of multiple first primer extension products from the target DNA. This process can occur simultaneously on both of the strands of the double-stranded DNA target, thus creating first primer extension products complementary to sequences in both strands.

[00206] In some embodiments, the first primer extension step is carried out with a DNA polymerase capable of extension at elevated temperature that is not compatible with subsequent hybridization of the random sequence to the displaced primer-extension product. For example, Bst DNA polymerase can be used which is active at elevated temperature. The reaction can be carried out stepwise, first with an incubation at a lower temperature such as about 25 °C, followed by an incubation at higher temperature such as about 50 °C. In some embodiments, the first incubation is carried out below about 30 °C, and the second incubation is carried out above about 40 °C. In some embodiments, a

IDT Exhibit 1020.0817

DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer. Mixtures of DNA polymerases can also be useful.

[00207] The method further comprises step (c) extending a second primer comprising a ligand and a 3' DNA region that comprises a random sequence, wherein the primer is optionally a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product whereby the second primer extension product comprises a ligand. This step may be carried out with or without prior denaturation. If carried out without denaturation, generally, only the single stranded displaced first primer extension product will hybridize to the second primer. Generally the second primer does not comprise RNA. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can be a tailed primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (B), which is not complementary to the first primer extension product. The second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target DNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target DNA. The second primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody.

[00208] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first composite primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded. The method produces a nucleic acid that comprises a ligand allowing it to be bound to a solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid attached to the ligand also comprises sequence (B) at or near its 5' end.

[00209] In some embodiments, The method further comprises step (d) binding the ligand to a receptor bound to a solid surface whereby the second primer extension product is attached to the solid surface. The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension product to the solid

IDT Exhibit 1020.0818

surface. In some embodiments, the second primer extension product is still hybridized to the first primer extension product when it is attached to the solid surface. In some embodiments, the first primer extension product is removed from the second primer extension product such that a single stranded polynucleotide is attached to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification.

[00210] In some embodiments, the amplification of the second primer extension product comprises the steps of: (e) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded; (f) annealing an amplification primer to the single stranded portion of the second primer extension product complementary to sequence (A), wherein the amplification primer has a DNA portion and a 5' RNA portion; (g) extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product hybridized to the second primer extension product; (h) cleaving the RNA from the amplified product hybridized to the second primer extension product; and (i) repeating steps (f) to (h) to produce multiple copies of amplified product. Where this method of amplification is used, and a sequence (B) is incorporated into the second primer extension product as described above, amplification product is generated which comprises sequence (B') complementary to sequence B at or near its 3' end.

[00211] In some embodiments, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target DNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target DNA.

[00212] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00213] The step of binding the polynucleotides to the solid surface through the ligand, step (d), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

IDT Exhibit 1020.0819

[00214] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary the second primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (A') on the second primer extension product for the hybridization of primer. In some embodiments, the amplification produces single stranded amplified product, In some embodiments, the amplification provides double stranded product. In some embodiments, the second primer comprises a specific sequence (B), which becomes incorporated into the second primer extension product. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the second primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (A') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (A') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00215] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (A') or sequence (B)). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00216] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer second primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the second primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target DNA, such as genomic DNA, wherein the plurality of second primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target DNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00217] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single second primer extension product can comprise a library. These libraries can be

IDT Exhibit 1020.0820

stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00218] A schematic exemplary of an embodiment of the invention relating to method for generating a polynucleotide comprising a ligand for binding to a solid surface from a DNA target is shown in Figure 3. Step I represents denaturing a double-stranded target DNA, for example by raising the temperature. Steps II and III illustrate annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end. The enzyme that carries out step (b) is generally a DNA polymerase. In some cases a mixture of DNA polymerases can be used. In some embodiments, a DNA polymerase with strand displacement activity is used such that a growing first primer extension product can displace a downstream first primer extension product, producing a plurality of first primer extension products, representing different regions of the sequence of the target DNA are produced. Step IV illustrates extending a second primer comprising a ligand and a 3' DNA region that comprises a random sequence, wherein the primer is optionally a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product whereby the second primer extension product comprises a ligand. Step IV(b) shows step IV in which the second primer comprises a segment (B) that is 5' of the 3' DNA region that comprises a random sequence. Step IV(a) shows step IV where the second primer does not comprise a (B) sequence. The second primer extension products comprising a ligand and a defined sequence (A) at its 3' end are useful for storage, archiving and analysis as they have a ligand capable of binding to a solid surface. Such second primer extension products also comprises sequences that are representative of (identical to or substantially identical to) a sequence in the target DNA, so analysis of these products provides information about the target DNA. Step V shows the binding the ligand to a solid surface, whereby the plurality of second primer extension products become bound to the solid surface.

Method for archiving and clonal expansion

[00219] An aspect of the invention is a method for archiving and clonal expansion of a nucleotide sequence.

[00220] The method comprises the steps of: (a) obtaining a plurality of partially double-stranded DNA products comprising a first polynucleotide and a second polynucleotide, wherein the second polynucleotide comprises a sequence (A') at its 3' end and a ligand, and the sequence (A') portion of the second polynucleotide is single-stranded, wherein optionally the second polynucleotide comprises a sequence (B) at or near its 5' end. The plurality of partially double-stranded DNA products comprising a first polynucleotide and a second polynucleotide, wherein the second

IDT Exhibit 1020.0821

polynucleotide comprises a sequence (A') at its 3' end and a ligand, and the sequence (A') portion of the second polynucleotide is single-stranded, wherein optionally the second polynucleotide comprises a sequence (B) at or near its 5' end may be obtained, for example, by the methods described above.

[00221] The method further comprises step (b) attaching the partially double stranded DNA products to a plurality of beads or a plurality of isolated areas on a surface by binding the ligands to the bead or isolated area. The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second polynucleotide to the solid surface. In some embodiments, the second polynucleotide is still hybridized to the first polynucleotide when it is attached to the solid surface. In some embodiments, the first polynucleotide is removed from the second polynucleotide such that a single stranded polynucleotide is attached to the solid surface. The method produces a polynucleotide that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the polynucleotide bound to the bead. As described above, in some embodiments, the polynucleotide bound to the bead also comprises sequence (B) at or near its 5' end. One aspect of the invention comprises amplification of the polynucleotide bound to the bead. In some embodiments, a plurality of different polynucleotides bound to a solid surface is created in which each of the polynucleotides has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different polynucleotides have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound polynucleotides thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00222] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00223] The step of binding the polynucleotides to the solid surface through the ligand, step (b), can be carried out such that only one polynucleotide is bound to an isolated area of a surface or only one polynucleotide is bound to a single bead. This isolated binding of polynucleotides can be used for clonal amplification of the specific bound polynucleotide in that area or on that bead. Such bound, isolated polynucleotides can also be stored and archived for later analysis, for example by sequencing. The bound, isolated polynucleotides can be amplified, stored, and analyzed multiple times.

[00224] The method further comprises the steps of: (c) annealing an amplification primer to the single stranded portion of the second polynucleotide complementary to sequence (A'), wherein the amplification primer has a DNA portion and a 5' RNA portion; (d) extending the amplification primer with an enzyme having strand displacement activity to produce a plurality of amplified products hybridized to the second polynucleotide products; (e) cleaving the RNA from the amplified product hybridized to the second polynucleotides using RNase H; and (f) repeating steps (c) to (e) to produce multiple copies of amplified products.

IDT Exhibit 1020.0822

[00225] In some embodiments, the DNA products are attached to beads or isolated areas of the solid surface, and on average, one DNA product is attached to one or fewer beads or isolated areas. The plurality of beads or isolated areas can be stored for later analysis, and in some cases can later be amplified with steps (c) through (f). In some embodiments, the amplification is a clonal amplification carried out in multiple isolated volumes wherein, on average, one isolated volume has one or fewer beads or isolated areas.

[00226] A schematic exemplary of an embodiment of the invention relating to amplification of polynucleotides bound to a solid surface is shown in Figure 4. The polynucleotides bound to the solid surface can be generated, for example from RNA or from DNA by the methods described above. A polynucleotides bound to the solid surface can be a second primer extension product comprising a sequence (A') at its 3' end. In some embodiments, as in those shown in Figure 4, the polynucleotides bound to the surface are hybridized to a first primer extension product comprising and RNA sequence (A). For clarity, the polynucleotides are referred to as second primer extension products. It will be understood, however, by those of ordinary skill in the art that polynucleotides of the same structure as shown here can be used in the present amplification methods. Where the polynucleotides are single stranded, it would be understood by those skilled in the art that the amplification can be carried out without the first RNase H step (Step I). In Figure 4, the steps labeled (b), e.g. Step I(b), represents the method for the case in which the second primer comprises a sequence (B) 5' to the 3' sequence that is complementary to the first primer extension product, while steps labeled (a), e.g. Step I(a) represent the method in which the second primer, and therefore the second primer extension product, does not include sequence (B).

[00227] In Figure 4, Step I illustrates cleaving the RNA from the first primer extension products such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. As shown, the cleavage is performed using RNase H. Chemical and thermal means can alternatively be employed. Step II shows annealing an amplification primer to the single stranded portion of the second primer extension products complementary to sequence (A), wherein the amplification primer has a DNA portion and a 5' RNA portion. Step III illustrates extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the second primer extension product. Step IV illustrates the step of cleaving the RNA from the amplified product hybridized to the second primer extension product. The product of step IV can be the starting point for another round of amplification in that Step II, the hybridization of and amplification primer, can occur, followed by the subsequent steps. Thus, steps II – IV can be repeated to produce multiple copies of amplified product. Note that for case (b) the amplified product that is produced comprises the sequence (B'), complementary to sequence (B) at its 3' end.

[00228] In some embodiments, the amplification illustrated in Figure 4 can be carried out on beads. In some embodiments, the beads have only one bound second primer extension product. In some embodiments such beads are individually in isolated volumes or microreactors allowing for clonal amplification.

IDT Exhibit 1020.0823

Method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target

[00229] One aspect of the invention is a method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target. The method utilizes a composite RNA/DNA oligonucleotide in order to extend the second primer extension product such that the second primer extension product comprises a sequence (C') at its 3' end than can be used as a site for isothermal amplification in a manner such that the sequence (A) is present at or near the 5' end of the amplified product produced in this amplification.

[00230] The method comprises the steps: (a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. . In some embodiments, the 3' portion of the primer that is complementary to the target RNA is a specific sequence. For example, where a specific region of interest of a target RNA that is known or suspected to be upstream of a specific sequence on the target RNA, the sequence that is complementary to the target RNA of the first primer can be designed to hybridize to this specific sequence on the target RNA such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target RNA comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence that is complementary to the target RNA can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence that is complementary to the target RNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a RNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target RNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample. In some embodiments more than one type of sequence that is complementary to the target RNA can be used, for instance both a primer with a random sequence and a primer, or combination of primers with a specific sequence complementary to RNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00231] The method further comprises step (b) cleaving the target RNA from the RNA/DNA hybrid. In some embodiments, the cleaving of the target RNA from the RNA/DNA hybrid involves selectively cleaving or degrading the target RNA. In some cases the cleaving can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent,

IDT Exhibit 1020.0824

or using a combination of heating the sample and adding denaturing agents. The cleaving can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or a combination of RNase enzymes, or chemically. In some embodiments, the target RNA is completely cleaved. In other embodiments, the target RNA is only partly cleaved or degraded. The amount of cleaving required is that amount which will allow the extension of the second primer. In some embodiments, the cleaving is carried out partially, and the fragments of the target RNA that remain can constitute the second primer for step (c).

[00232] The method further comprises step (c) extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and wherein a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. The second primer can be a tailed primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (B), which is not complementary to the first primer extension product. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target RNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target RNA (sense copies). The second primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody.

[00233] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded. The product further comprises a DNA-RNA heteroduplex region. The method produces a nucleic acid that comprises a ligand allowing it to be bound to a solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and

IDT Exhibit 1020.0825

further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid attached to the ligand also comprises sequence (B) at or near its 5' end.

[00234] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in copending application.

[00235] The method further comprises step (d) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. The cleaving of RNA can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (c).

[00236] The method further comprises step (e) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C). The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 5' DNA segment is complementary to all of sequence (A'), in other embodiments, the 5' DNA segment is complementary to portion of sequence (A'). In some embodiments, 5' RNA segment comprising sequence (C) is partly complementary to sequence (A').

[00237] The method optionally comprises step: (f) extending the oligonucleotide to form an oligonucleotide extension product hybridized to the second primer extension product. In some embodiments, the oligonucleotide is optionally extended from its 3' end to produce an oligonucleotide extension product hybridized to the second primer extension product and displacing the DNA portion of the first primer extension product. In some embodiments, the second primer comprises a sequence (B), such that the oligonucleotide extension product will comprise a sequence (B') at or near its 3' end that is complementary to sequence (B).

[00238] The method further comprises step (g) extending the second primer extension product to create a heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C). The DNA sequence (C') is created by a DNA polymerase that has RNA dependent DNA polymerase activity. This step creates an RNA/DNA heteroduplex region that can be used for further manipulation of the second primer extension product.

IDT Exhibit 1020.0826

[00239] The method further comprises step (h) cleaving the RNA from the heteroduplex created in step (g) to produce a single-stranded portion of the second primer extension product corresponding to sequence (C'). The RNA can be cleaved from the heteroduplex, for example, by RNase H.

[00240] In some embodiments, The method further comprises step (i) binding the ligand on the second primer extension product to a solid surface. The binding can be performed at various stages or steps during the procedure depending, for example, on whether it is advantageous to carry out the subsequent steps on a solid surface. The binding step is generally performed after step (c). In some embodiments, the binding the ligand to the solid surface is performed before step (h). In some embodiments, the binding the ligand to the solid surface is performed after step (h). The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension product to the solid surface. In some embodiments, the second primer extension product is still hybridized to the first primer extension product when it is attached to the solid surface. In some embodiments, the first primer extension product is removed from the second primer extension product such that a single stranded polynucleotide is attached to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') and (C') at its 3' end. The specific sequence (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (C') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. When the sequence (C') acts as a site to which a composite amplification primer binds, the amplified product that is produced has the sequence (A) (and a portion of sequence (C) at its 5' end. Where the second primer comprises the sequence (B), the amplified product also has the sequence (B'), complementary to (B) at or near its 3' end. Thus the method produced amplified product with defined sequences at or near both its 3' and 5' ends.

[00241] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') and (C') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target RNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target RNA.

[00242] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

IDT Exhibit 1020.0827

[00243] The step of binding the polynucleotides to the solid surface through the ligand, step (i), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00244] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary the second primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (C') on the second primer extension product for the hybridization of primer. In some embodiments, the amplification produces single stranded amplified product, In some embodiments, the amplification provides double stranded product. In some embodiments, the second primer comprises a specific sequence (B), which becomes incorporated into the second primer extension product. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the second primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (C') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (C') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00245] In some embodiments, the amplification is performed by a method comprising the following steps: (j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension product complementary to sequence (C'); (k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product; (l) cleaving the RNA from the amplified product; and (m) repeating steps (j) to (l) to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence complementary to sequence (A'). Where the second primer further comprises a segment (B) that is not complementary to the first primer extension product sequence, this amplification method, utilizing sequence (C'), allows for the production of an amplified product comprises a sequence (B') at or near its 3' end that is substantially complementary to sequence (B), and a sequence (A) near its 5' end that is complementary to sequence (A'), thus producing an amplified polynucleotide product with defined 3' and 5' ends.

[00246] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other

IDT Exhibit 1020.0828

portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00247] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer second primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the second primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target RNA, such as whole transcriptome or total RNA, wherein the plurality of second primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00248] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single second primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00249] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (B') at their 5' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer second primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of second primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (B') on the amplified product to the sequence (B) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it. Where these sequences are representative of the target RNA, the plurality of beads can constitute a library representative of such RNA.

[00250] After the amplified products are bound to the beads by hybridization, the (B) sequences on the beads can be extended along the amplified product by a DNA polymerase or mixture of polymerases to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (A') near their 5' ends. Where the (B) sequences are covalently attached to the beads, this method provides

IDT Exhibit 1020.0829

for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00251] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (A') and a portion of sequence (C'). This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (A') that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein.

[00252] The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

[00253] In some embodiments, the sequencing methods comprise sequencing by ligation. Sequencing by ligation involves using a DNA ligase. Although commonly represented as joining two pairs of ends at once, as in the ligation of restriction enzyme fragments, ligase can also join the ends on only one of the two strands (for example, when the other strand either already continuous or lacks a terminal phosphate necessary for ligation). DNA ligase is sensitive to the structure of DNA and has very low efficiency when there are mismatches between the bases of the two strands. Sequencing by ligation relies upon the sensitivity of DNA ligase for base-pairing mismatches.

[00254] The target molecule to be sequenced is a single strand of unknown DNA sequence, flanked on at least one end by a known sequence. A short "anchor" strand is brought in to bind the known sequence. A mixed pool of probe oligonucleotides is then brought in (eight or nine bases long), labeled (typically with fluorescent dyes or other detection means) according to the position that will be sequenced. These molecules hybridize to the target DNA sequence, next to the anchor sequence, and DNA ligase preferentially joins the molecule to the anchor when its bases match the unknown DNA sequence. Based on the fluorescence or other signal produced by the molecule, one can infer the identity of the nucleotide at this position in the unknown sequence. The oligonucleotide probes may also be constructed with cleavable linkages which can be cleaved after identifying the label. This will both remove the label and regenerate a 5' phosphate on the end of the ligated probe, preparing the system for another round of ligation.

[00255] This cycle can be repeated several times to read longer sequences. This sequences every Nth base, where N is the length of the probe left behind after cleavage. To sequence the skipped positions, the anchor and ligated oligonucleotides may be stripped off the target DNA sequence, and another round of sequencing by ligation started with an anchor one or more bases shorter. A simpler, albeit more limited, technique is to do repeated rounds of a single ligation where the label corresponds to different position in the probe, followed by stripping the anchor and ligated probe.

IDT Exhibit 1020.0830

Sequencing by ligation can proceed in either direction (either 5'-3' or 3'-5') depending on which end of the probe oligonucleotides are blocked by the label. The 3'-5' direction is more efficient for doing multiple cycles of ligation. Note that this is the opposite direction to polymerase based sequencing methods.

[00256] One feature unique to sequencing by ligation is the possibility of labeling the probe oligonucleotides according to various combinations of bases at more than one position. This has error detection capabilities not available to polymerase-based sequencing methods. The Applied Biosystems SOLiD sequencing system uses 2-base encoding to improve its error rates.

[00257] In some embodiments the methods of the invention provide for performing bridge PCR comprising making amplified product as described above with defined 3' and 5' ends, and further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the defined 3' and 5' sequences, for example, A and B' sequences, on the amplified product in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

[00258] In some embodiments the methods of the invention provide for making amplified product as described above with defined 3' and 5' ends and further performing rolling circle amplification comprising performing the steps of: (n) hybridizing the amplified product to an oligonucleotide comprising regions complementary to A and B' sequences in close proximity; (o) optionally extending the gap with a polymerase enzyme; (p) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid. In some embodiments, the rolling circle amplification uses primers complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product. In some cases, such a primer can be an oligonucleotide attached to a solid surface, thus resulting in amplified product bound to the surface

[00259] In some embodiments the methods of the invention provide for performing PCR comprising making amplified product as described above with defined 3' and 5' ends, further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

[00260] In some embodiments the methods of the invention provide for performing strand displacement amplification (SDA) comprising making amplified product as described above with defined 3' and 5' ends, wherein the defined 3' and 5' ends, for example, sequences (A) and (B'), in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

Method for generating a polynucleotide having a defined 3' and 5' sequences from a DNA target

[00261] One aspect of the invention is a method for generating a polynucleotide having a defined 3' and 5' sequences from a DNA target. The method utilizes a composite RNA/DNA oligonucleotide in order to extend the second primer such that the second primer extension products comprise a sequence (C') at its 3' end than can be used as a site for isothermal amplification in a manner such that the sequence (A) is present at or near the 5' end of the amplified product produced in this

IDT Exhibit 1020.0831

amplification, and where a second primer comprising sequence (B) is used, amplified products with defined sequences at both the 3' and 5' ends can be produced.

[00262] The method comprises the step: (a) denaturing a double-stranded target DNA. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00263] The method further comprises step (b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end. The sequence that is complementary to the target DNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a target DNA, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target DNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the number of gene copies in a DNA sample, or obtaining sequence information. In some embodiments, the extension of one first primer, will result in the release of a downstream first primer extension product. This can occur throughout the target DNA resulting in the release of multiple first primer extension products from the target DNA. This process can occur simultaneously on both of the strands of the double-stranded DNA target, thus creating first primer extension products complementary to sequences in both strands.

[00264] In some embodiments, the first primer extension step is carried out with a DNA polymerase capable of extension at elevated temperature that is not compatible with subsequent hybridization of the random sequence to the displaced primer-extension product. For example, Bst DNA polymerase can be used which is active at elevated temperature. The reaction can be carried out stepwise, first with an incubation at a lower temperature such as about 25 °C, followed by an incubation at higher temperature such as about 50 °C. In some embodiments, the first incubation is carried out below about 30 °C, and the second incubation is carried out above about 40 °C. In some embodiments, a DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer. Mixtures of DNA polymerases can also be useful.

[00265] The method further comprises step (c), extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and wherein a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. In this step, a second primer comprising a ligand and a 3' DNA region that comprises a random sequence is used. The primer is optionally a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product whereby the second primer extension products comprise a ligand.

IDT Exhibit 1020.0832

This step may be carried out with or without prior denaturation. If carried out without denaturation, generally, only the single stranded displaced first primer extension products will hybridize to the second primer. Generally the second primer does not comprise RNA. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can be a tailed primer having a 3' portion which is complementary to the first primer extension products, and a 5' portion, sequence (B), which is not complementary to the first primer extension products. The second primer comprises a random primer sequence that randomly binds to the first primer extension products. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target DNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target DNA. The second primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody.

[00266] The second primer extension products are extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first composite primer. Since sequence (A) on the first primer extension products comprise RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a products that are at least partially double stranded. The method produces a nucleic acid that comprises a ligand allowing it to be bound to a solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid attached to the ligand also comprises sequence (B) at or near its 5' end.

[00267] The method further comprises step (d) cleaving the RNA from the first primer extension products such that a portion of the second primer extension products that are complementary to sequence (A) is single stranded. The cleaving of RNA can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (c).

[00268] The method further comprises step (e) annealing to the second primer extension products an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C). The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 5' DNA segment is complementary to all of sequence (A'), in other embodiments, the 5' DNA segment is complementary to portion of sequence

IDT Exhibit 1020.0833

(A'). In some embodiments, 5' RNA segment comprising sequence (C) is partly complementary to sequence (A').

[00269] The method optionally comprises step: (f) extending the oligonucleotide to form a plurality of oligonucleotide extension products hybridized to the second primer extension products. In some embodiments, the oligonucleotide is optionally extended from its 3' end to produce a plurality of oligonucleotide extension products hybridized to the second primer extension product and displacing the DNA portion of the first primer extension product. In some embodiments, the second primer comprises a sequence (B), such that the oligonucleotide extension products will comprise a sequence (B') at or near their 3' ends that are complementary to sequence (B).

[00270] The method further comprises step (g) extending the second primer extension products to create a heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C). The DNA sequence (C') is created by a DNA polymerase that has RNA dependent DNA polymerase activity. This step creates an RNA/DNA heteroduplex region that can be used for further manipulation of the second primer extension products.

[00271] The method further comprises step (h) cleaving the RNA from the heteroduplex created in step (g) to produce a single-stranded portion of the second primer extension products corresponding to sequence (C'). The RNA can be cleaved from the heteroduplex, for example, by RNase H.

[00272] In some embodiments, The method further comprises step (i) binding the ligand on the second primer extension products to a solid surface. The binding step can be performed at various stages or steps during the procedure depending, for example, on whether it is advantageous to carry out the subsequent steps on a solid surface. The binding step is generally performed after step (c). In some embodiments, the binding the ligand to the solid surface is performed before step (h). In some embodiments, the binding the ligand to the solid surface is performed after step (h). The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension products to the solid surface. In some embodiments, the second primer extension products are still hybridized to the first primer extension products when they are attached to the solid surface. In some embodiments, the first primer extension products are removed from the second primer extension products such that a single stranded polynucleotide is attached to the solid surface. The method produces polynucleotides that are bound to a solid surface that have a specific sequence (A') and (C') at their 3' ends. The specific sequence (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (C') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. When the sequence (C') acts as a site to which a composite amplification primer binds, the amplified products that are produced have the sequence (A) (and a portion of sequence (C) at their 5' ends)). Where the second primer comprises the sequence (B), the amplified products also have the sequence

IDT Exhibit 1020.0834

(B'), complementary to (B) at or near their 3' ends. Thus the method produced amplified products with defined sequences at or near both its 3' and 5' ends.

[00273] Due to the use of random priming, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') and (C') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target DNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target DNA.

[00274] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00275] The step of binding the polynucleotides to the solid surface through the ligand, step (i), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times. This allows, for example, for the analysis of an individuals genes, or a portion of the individual genes at one time, then allows for archiving the genetic material at a later date by the same or by different tests.

[00276] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of amplification products that are substantially complementary to the second primer extension products. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (C') on the second primer extension products as for primer hybridization. In some embodiments, the amplification produces single stranded amplified products. In some embodiments, the amplification provides double stranded products. In some embodiments, the second primer comprises a specific sequence (B), which becomes incorporated into the second primer extension products. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the second primer extension products comprises both as sequence (B) at or near its 5' end and a sequence (C') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (C') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified products.

[00277] In some embodiments, the amplification is performed by a method comprising the following steps: (j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension products complementary to sequence (C'); (k) extending the amplification primer with an enzyme having

IDT Exhibit 1020.0835

strand displacement activity to produce an amplified products; (l) cleaving the RNA from the amplified products; and (m) repeating steps (j) to (l) to produce multiple copies of amplified products wherein the 5' portion of the amplified products have a sequence complementary to sequence (A'). Where the second primer further comprises a segment (B) that is not complementary to the first primer extension products sequence, this amplification method, utilizing sequence (C'), allows for the production of amplified products comprises a sequence (B') at or near their 3' ends that is substantially complementary to sequence (B), and a sequence (A) near their 5' end that are complementary to sequence (A'), thus producing an amplified polynucleotide products with defined 3' and 5' ends.

[00278] In some embodiments the amplification is carried out such that the amplified products are not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified products remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, amplified products are generated that may float freely in solution, but which comprise a sequence, for example sequence (A) or sequence (B'), that allows them to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00279] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer second primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the second primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target DNA, such as genomic DNA, wherein the plurality of second primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target DNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified products.

[00280] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single second primer extension product can comprise a library. The library can represent, for example, the genomic DNA of an individual, or the genomic DNA from a group of cells or from a single cell. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

IDT Exhibit 1020.0836

[00281] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (B') at their 5' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer second primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of second primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (B') on the amplified product to the sequence (B) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it. Where these sequences are representative of the target DNA, the plurality of beads can constitute a library representative of such DNA.

[00282] After the amplified products are bound to the beads by hybridization, the (B) sequences on the beads can be extended to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (A') near their 5' ends. Where the (B) sequences are covalently attached to the beads, this method provides for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00283] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (A') and a portion of sequence (C'). This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (A) that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein. The single stranded covalently bound polynucleotides derived from DNA can be sequenced as described above for the single stranded covalently bound polynucleotides derived from DNA, such as pyrosequencing, cycle sequencing, isothermal sequencing and other methods such as those described below.

[00284] The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

[00285] In some embodiments the methods of the invention provide for performing bridge PCR comprising making amplified product as described above with defined 3' and 5' ends, and further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the defined 3' and 5' sequences, for example, A and B' sequences, on the amplified product in the presence of components necessary for

IDT Exhibit 1020.0837

polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification. The bridge PCR can be carried out such that isolated lawns of amplified product are obtained wherein each lawn comprises polynucleotides having sequences representative of a portion of the sequence of the target DNA.

[00286] In some embodiments the methods of the invention provide for making amplified product as described above with defined 3' and 5' ends and further performing rolling circle amplification comprising performing the steps of: (n) hybridizing the amplified product to a target DNA comprising regions complementary to A and B' sequences in close proximity; (o) optionally extending the gap with a polymerase enzyme; (p) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid. In some embodiments, the rolling circle amplification uses primers complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product. In some cases, such a primer can be an oligonucleotide attached to a solid surface, thus resulting in amplified product bound to the surface

[00287] In some embodiments the methods of the invention provide for performing PCR comprising making amplified product as described above with defined 3' and 5' ends, further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

[00288] In some embodiments the methods of the invention provide for performing strand displacement amplification (SDA) comprising making amplified product as described above with defined 3' and 5' ends, wherein the defined 3' and 5' ends, for example, sequences (A) and (B'), in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

[00289] A schematic exemplary of an embodiment of the invention relating to generating a polynucleotide having a defined 3' and 5' sequences is shown in Figures 5A and 5B. The starting material in 5A can be generated as described in the methods herein. The starting material can be generated, for example, from the methods above from a target RNA or a target DNA. While the starting material is described as, for example first primer extension product, it will be understood that a complex of the same structure as shown and described can be used in the method described here.

[00290] In Figure 5A, step I shows the step of: cleaving the RNA from the first primer extension product in the DNA-RNA heteroduplex such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. As shown, the cleavage is performed using RNase H. Chemical and thermal means can alternatively be employed. Step II illustrates annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C). Step III illustrates extending the second primer extension product to create a DNA-RNA heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C). Step IV illustrates the optional step of extending the oligonucleotide to form an oligonucleotide extension product hybridized to the second primer extension product. Step V illustrates binding the ligand to a solid surface resulting in binding of the second primer extension product to the solid surface. While

IDT Exhibit 1020.0838

the step of binding the ligand to the solid surface is shown at step V is to be understood that the binding step can be performed at many steps in the process. For example, in this case, the binding step could alternately have been carried out before step I , or could be carried out after amplification.

[00291] In Figure 5B, step VI illustrates cleaving the RNA from the DNA-RNA heteroduplex created in step III to produce a single-stranded portion of the second primer extension product corresponding to sequence (C'). Step VII illustrates annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension product complementary to sequence (C'). Step VIII illustrates extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product. Step IX illustrates the step of cleaving the RNA from the extended hybridized amplification primer in the DNA-RNA heteroduplex. The product of step IX is now capable of hybridizing another amplification primer, allowing for steps VII to IX to occur again, resulting in the generation of another amplified product. These steps can thus be repeated to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence complementary to sequence (A'). In the embodiment illustrated in Figures 5A and 5B the second primer comprised a sequence (B), and as such, the amplified product shown in Figure 5B comprises a defined sequences on both the 5' and 3' ends.

[00292] A schematic exemplary of an embodiment of the invention relating to the capture of amplification product on a solid surface is shown in Figure 6. In some embodiments, the amplified product can be captured onto a surface by hybridization to the defined sequences on the 5' and/or 3' ends of the amplified product. This approach can be useful for example, where a clonal amplification of a given sequence or plurality of sequences that are representative of a target RNA or target DNA produces multiple copies of a specific sequence having defined 5' and 3' ends. The clonal amplification can be carried out in an isolated volume, or on in an isolated area on a surface. In some embodiments, the clonal amplification is carried out with beads in droplets of an inverse emulsion, wherein each droplet has on average one or fewer beads. In some embodiments, it is useful to capture the amplified product on a solid surface, for example a bead. Figure 6 illustrates how a solid surface, for example the surface of a bead can have multiple oligonucleotides attached thereto. In the embodiment illustrated, the amplified product has a defined sequence (B') at its 3' end and a defined sequence (A) near its 5' end. The bead has multiple oligonucleotides attached to its surface comprising the sequence (B), which is complementary to the sequence (B') at the 3' end of an amplified product. The bead can capture the amplified product as shown by hybridization under the right conditions. The capture of amplified product is illustrated in step I. This type of bead can be used as a library, and can be used for sequencing, for example, by extending from sequence (B). Figure 6 also illustrates in step II that the (B) sequences on the bead can be extended along the amplified product to produce a polynucleotide attached to the bead which has sequence (A') at or near its 3' end and sequence (B) at its 5' end. Such a polynucleotide can be made single stranded, for example by heat or chemical denaturation, and washing as illustrated in step III. The bead with single stranded polynucleotide which has sequence (A') at or near its 3' end and sequence (B) at its 5' can be more stable than a bead to which a polynucleotide is held by hybridization alone. The polynucleotide attached to the bead can be covalently bound. The beads made in this fashion

IDT Exhibit 1020.0839

can be used to create libraries with long storage life. The single stranded product of step III can be used for archiving, storage, and a variety of manipulations made possible by the defined sequences at the 5' and 3' ends. For example, as shown in Figure 6, a sequencing primer (A) can be hybridized to sequence (A') to allow the sequencing of the polynucleotide attached to the bead by methods such as those described herein.

[00293] Figure 7 illustrates that the amplified product with a defined sequence (B') at its 3' end and a defined sequence (A) near its 5' end can also be captured onto surfaces such as beads by hybridization to an oligonucleotide attached to the solid surface that comprises the sequence (A'). In some cases, the capture is by specific nucleic acid hybridization of sequence (A) of the amplified product to the complementary sequence (A'). In other cases, the oligonucleotide attached to the surface of the solid surface may not be a perfect complement of at least a portion of the 5' end of the amplified product. For example, the oligonucleotide may differ by about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, or more nucleotides from the perfect complement of sequence (A). The oligonucleotide may be attached to the solid surface by any method known in the art. For example, the oligonucleotide may be attached via non-covalent interaction such as by ionic, hydrogen, or hydrophobic interaction or a combination thereof. In some cases, the oligonucleotide may be attached by one or more covalent bonds between the oligonucleotide and the solid surface. In some cases, the oligonucleotide may be attached by interaction between a receptor ligand pair such as for example an avidin molecule on the surface and a biotin molecule on the oligonucleotide, or any receptor ligand pair known in the art.

[00294] Figure 8 illustrates that the amplified products with defined 3' and 5' sequences can be used for gap ligation to create a circular polynucleotide which can then be amplified and characterized using rolling circle amplification. In Figure 8, the amplified product with a defined sequence (B') at its 3' end and a defined sequence (A) near its 5' end is hybridized to a sequence D with sequences complementary to sequence (A) and sequence (B') (with B' hybridizes upstream from (A), wherein the hybridization of sequences (A) and (B') leave a gap. The gap is closed with a ligase and alternatively with a polymerase for larger gaps. The circular polynucleotide, e.g. DNA, can then be amplified and characterized using rolling circle amplification. Primers for rolling circle amplification can hybridize (3) to sequence (A), (2) to sequence (B'), or (1) to the sequence between (A) and (B') on the amplified product. These methods can be used for the accurate determination of mutations such as single nucleotide polymorphisms (SNPs).

Alternative method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target

[00295] One aspect of the present invention is an alternative method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target. Unlike the methods described in detail above, this method generally does not use a composite RNA/DNA primer as the first primer. Here, a tailed primer, usually not comprising RNA, composed, for example, of DNA is used to create the first primer extension product. In this method, the second primer comprises a composite RNA/DNA primer, and the third primer comprises a ligand for binding the third primer extension product to a solid surface.

IDT Exhibit 1020.0840

[00296] The method comprises step: (a) extending a first primer comprising a 3' portion complementary to a target RNA and a 5' portion, sequence (D), not complementary to the target RNA, to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. The first primer generally does not comprise RNA, and may be all DNA. The first primer is a tailed primer comprising a 5' portion, sequence (D) which is generally not complementary to the target RNA, and does not hybridize to the target RNA. In some embodiments, the 3' portion of the primer that is complementary to the target RNA is a specific sequence. For example, where a specific region of interest of a target RNA that is known or suspected to be upstream of a specific sequence on the target RNA, the sequence that is complementary to the target RNA of the first primer can be designed to hybridize to this specific sequence on the target RNA such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target RNA comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence that is complementary to the target RNA can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence that is complementary to the target RNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a RNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target RNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample. In some embodiments more than one type of sequence that is complementary to the target RNA can be used, for instance both a primer with a random sequence and a primer, or combination of primers with a specific sequence complementary to RNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00297] The method further comprises step (b) cleaving the target RNA from the RNA/DNA hybrid. In some cases the cleaving can be accomplished by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. The cleaving can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or a combination of RNase enzymes, or chemically. In some embodiments, the target RNA is completely cleaved. In other embodiments, the target RNA is only partly cleaved or degraded. The amount of cleaving required is that amount which will allow the extension of the second primer.

[00298] The method further comprises step (c) extending a second primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to the first

IDT Exhibit 1020.0841

primer extension product and a 5' portion, sequence (E), of the of the second primer is not complementary to the first primer extension product, to produce a double-stranded DNA product comprising the first primer extension product hybridized to a second primer extension product, whereby the second primer extension product has a sequence (D') that is complementary to sequence (D) at its 3' end. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer is a composite RNA/DNA primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (E), which is not complementary to the first primer extension product. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target RNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target RNA (sense copies).

[00299] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (D') which is complementary to sequence (D) of the first primer. The primer extension results in a product that is at least partially double stranded.

[00300] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in copending application.

[00301] The method further comprises step (d) denaturing the double-stranded DNA product. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00302] The method further comprises step (e) annealing to the second primer extension product and extending a third primer comprising, from its 5' end, a ligand, optionally a sequence (F), and a sequence (D), wherein sequence (D) is complementary to sequence (D') on the second primer

IDT Exhibit 1020.0842

extension product to produce a double-stranded DNA product comprising the second primer extension product hybridized to a third primer extension product, whereby the third primer extension product comprises a sequence (E') at its 3' end complementary to sequence (E) in a DNA-RNA heteroduplex. The third primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody. The method produces a nucleic acid that comprises a ligand allowing it to be bound to a solid surface and that has a specific sequence (E') at its 3' end. The specific sequence (E') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, the nucleic acid attached to the ligand also comprises sequence (D') at or near its 5' end, and in some embodiments comprises the sequence (F').

[00303] In some embodiments, the method further comprises the step of binding the ligand to a solid surface, whereby the third primer extension product is bound to the solid surface.

[00304] In some embodiments, the method further comprises the steps of: (f) cleaving the RNA portion of the second primer extension product in the DNA-RNA heteroduplex, whereby sequence (E') of the third primer extension product is single stranded, (g) annealing an oligonucleotide comprising a 3' DNA segment (E) that is complementary to sequence (E') and a 5' RNA segment comprising sequence (G), (h) extending the third primer extension product to produce a sequence (G') at its 3' end complementary to sequence (G), and (i) cleaving the RNA from the heteroduplex created in step (h) to produce a single-stranded portion of the third primer extension product corresponding to sequence (G'). The cleaving of RNA in steps (f) and (i) can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (e). The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 3' DNA segment is complementary to all of sequence (E'), in other embodiments, the 3' DNA segment is complementary to portion of sequence (E'). In some embodiments, 5' RNA segment comprising sequence (G) is partly complementary to sequence (E'). In some embodiments, the oligonucleotide is optionally extended from its 3' end to produce an oligonucleotide extension product hybridized to the third primer extension product and displacing the DNA portion of the second primer extension product. In some embodiments, the third primer comprises a sequence (F), such that the oligonucleotide extension product will comprise a sequence (F') at or near its 3' end that is complementary to sequence (F). The DNA sequence (G') is created by a DNA polymerase that has RNA dependent DNA polymerase activity. This step creates an RNA/DNA heteroduplex region that can be used for further manipulation of the second primer extension product.

[00305] In some embodiments the method further comprises binding the ligand to a solid surface, whereby the third primer extension product comprising sequence (G') is bound to the solid surface. The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic. The receptor bound to the solid surface is a member of the ligand-receptor pair such that

IDT Exhibit 1020.0843

binding of the ligand results in attaching the second primer extension product to the solid surface. In some embodiments, the second primer extension product is still hybridized to the first primer extension product when it is attached to the solid surface. In some embodiments, the second primer extension product is removed from the third primer extension product such that a single stranded polynucleotide is attached to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (E') and (G') at its 3' end. The specific sequence (G') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (F) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (G') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. When the sequence (G') acts as a site to which a composite amplification primer binds, the amplified product that is produced has the sequence (E) (and a portion of sequence (G) at its 5' end. The amplified product has a sequence (D') at or near its 3' end. Where the third primer comprises the sequence (F), the amplified product also has the sequence (F'), complementary to (F) at or near its 3' end. Thus the method produced amplified product with defined sequences at or near both its 3' and 5' ends.

[00306] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (E') and (G') at its 3' end, the sequence (D) at or near its 5' end (and in some embodiments also a specific sequence (F) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target RNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target RNA.

[00307] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary the second primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (G') on the third primer extension product for the hybridization of primer. In some embodiments, the amplification produces single stranded amplified product, In some embodiments, the amplification provides double stranded product. The third primer extension product comprises the specific sequence (D). In some embodiments, the third primer comprises a specific sequence (F), which thus becomes incorporated into the third primer extension product. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the third primer extension product comprises both as sequence (F) at or near its 5' end and a sequence (G') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of

IDT Exhibit 1020.0844

the sequence (G') and the other designed to hybridize to sequence (F'), complementary to sequence (F), and/or to sequence (D') to exponentially produce double stranded amplified product.

[00308] One aspect of the invention is a method of amplifying a sequence representative of a sequence within a target RNA comprising the above steps and further comprising: (j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the third primer extension product complementary to sequence (G'); (k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product; (l) cleaving the RNA from the amplified product; and (m) repeating steps (j) to (l) to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence (E) complementary to sequence (E') and the 3' end of the amplified product has sequence (D') complementary to sequence (D) and optionally sequence (F') complementary to sequence (F).

[00309] The step of binding the polynucleotides to the solid surface through the ligand can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00310] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (E) or sequence (D'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (E') or sequence (D). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00311] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer third primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the third primer extension products bound to the beads is then carried out such that the clonal amplification of a plurality of third primer extension products is achieved. This clonal amplification in microreactors can be performed on a sample of target RNA, such as whole transcriptome or total RNA, wherein the plurality of third primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (E') or sequence (D)), which are complementary to sequence (E) or sequence (D') on the amplified product.

IDT Exhibit 1020.0845

[00312] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single third primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a third primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00313] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (D) (or F) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (D') (or F') at their 5' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer third primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of third primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (D') (and/or F') on the amplified product to the sequence (D) (and/or F) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it. Where these sequences are representative of the target RNA, the plurality of beads can constitute a library representative of such RNA.

[00314] After the amplified products are bound to the beads by hybridization, the (D) sequences on the beads can be extended to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (E') near their 5' ends. Where the (D) (and or F) sequences are covalently attached to the beads, this method provides for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00315] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (E') and a portion of sequence (G'). This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (E) that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein.

[00316] The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

IDT Exhibit 1020.0846

[00317] The amplified products with defined 3' and 5' ends can be used in the methods described herein for the amplified products produced by the other methods. They can be used, for example, for bridge PCR, rolling circle amplification, and strand displacement amplification.

[00318] A schematic exemplary of an embodiment of the invention relating to an alternative method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target is shown in Figure 9. Step I illustrates the step of extending a DNA first primer comprising a 3' portion complementary to a target RNA and a 5' portion, sequence (D), not complementary to the target RNA, to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. Step II illustrates cleaving the target RNA from the RNA/DNA hybrid, and hybridizing the second primer. Step III illustrates extending a second primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to the first primer extension product and a 5' portion, sequence (E), of the of the second primer is not complementary to the first primer extension product, to produce a double-stranded DNA product comprising the first primer extension product hybridized to a second primer extension product, whereby the second primer extension product has a sequence (D') that is complementary to sequence (D) at its 3' end. Step IV illustrates denaturing the double-stranded DNA product. Step V illustrates annealing to the second primer extension product and extending a third primer comprising, from its 5' end, a ligand, optionally a sequence (F), and a sequence (D), wherein sequence (D) is complementary to sequence (D') on the second primer extension product to produce a double-stranded DNA product comprising the second primer extension product hybridized to a third primer extension product, whereby the third primer extension product comprises a sequence (E') at its 3' end complementary to sequence (E). The products of steps V and VI may be useful for SPIA amplification of a target RNA sequence or its complement.

[00319] Step VI illustrates the step of cleaving the RNA portion of the second primer extension product in the DNA-RNA heteroduplex, whereby sequence (E') of the third primer extension product is single stranded.

[00320] Figure 10 illustrates the isothermal amplification starting from the partially double-stranded polynucleotide produced either from RNA described above, or from DNA as described below. Step VII comprises annealing an amplification primer comprising 5' RNA and 3' DNA segments. Step VIII shows the extension of the amplification primer by a DNA polymerase with strand displacement activity. Step IX shows the cleavage of the RNA portion of the amplification primer. Step X illustrates the continued extension of the amplification primer by a DNA polymerase with strand displacement activity to produce an amplified product comprising sequence (D') and sequence (F') at its 3' end. The product of step X can hybridize to another amplification primer, and steps VII through X can be repeated to produce multiple copies of amplified product. While the ligand is unbound in this embodiment, it is to be understood that the ligand can be bound to the surface at various steps in the process. For example, it can be advantageous to have the species bound during the amplification step. Figure 11 illustrates the process shown in Figure 10 where the ligand and third primer extension product are bound to the support during amplification.

Alternative method for generating a polynucleotide having a defined 3' and 5' sequences from a DNA target

IDT Exhibit 1020.0847

[00321] One aspect of the invention is an alternative method for generating a polynucleotide having a defined 3' and 5' sequences from a DNA target. This method generally does not use a composite RNA/DNA primer as the first primer. Here, a tailed primer, usually not comprising RNA, composed, for example, of DNA is used to create a plurality of first primer extension products. In this method, the second primer comprises a composite RNA/DNA primer, and the third primer comprises a ligand for binding the third primer extension products to a solid surface.

[00322] The method comprises step: (a) denaturing a double-stranded target DNA. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00323] The method further comprise step: (b) annealing to the target DNA and extending a first primer comprising a 3' portion comprising a random sequence and a 5' portion, sequence (D), which is not complementary to the target DNA, to form a plurality of first primer extension products, each comprising sequence (D) at its 3' end. The first primer generally does not comprise RNA, and may be all DNA. The first primer is a tailed primer comprising a 5' portion, sequence (D) which is generally not complementary to the target DNA, and does not hybridize to the target DNA. The sequence that is complementary to the target DNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a DNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target DNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the number of gene copies in a target DNA sample. In some embodiments more than one type of sequence that is complementary to the target DNA can be used, for instance both a primer with a random sequence and a primer with a sequence complementary to DNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00324] The method further comprises step (c): extending a second primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion comprises a random sequence, and a 5' portion, sequence (E), of the of the second primer is not complementary to the first primer extension products, to produce a plurality of double-stranded DNA products comprising a first primer extension product hybridized to a second primer extension product, whereby the second primer extension products have a sequence (D') that is complementary to sequence (D) at their 3' ends. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer is a composite RNA/DNA primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (E), which is not complementary to the first primer extension product. The second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target DNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common

IDT Exhibit 1020.0848

to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target DNA (sense copies).

[00325] The second primer extension products are extended such that the 3' portion of the second primer extension products comprises a sequence (D') which is complementary to sequence (D) of the first primer. The primer extension results in products that are at least partially double stranded.

[00326] The method further comprises step (d) denaturing the double-stranded DNA product. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00327] The method further comprises step (e) annealing to the second primer extension product and extending a third primer comprising, from its 5' end, a ligand, optionally a sequence (F), and a sequence (D), wherein sequence (D) is complementary to sequence (D') on the second primer extension products to produce double-stranded DNA products comprising second primer extension products hybridized to a third primer extension products, whereby the third primer extension products comprise a sequence (E') at its 3' end complementary to sequence (E). The third primer comprises a ligand that is a member of a ligand-receptor pair. In some embodiments, the ligand is attached to the primer at the 5' end of the primer. In some embodiments, the ligand is a small molecule, such as biotin or digoxigenin. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule recognized by the antibody. The method produces nucleic acids that comprise a ligand allowing them to be bound to a solid surface and that has a specific sequence (E') at its 3' end. The specific sequence (E') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. As described above, the nucleic acid attached to the ligand also comprises sequence (D') at or near its 5' end, and in some embodiments comprises the sequence (F').

[00328] In some embodiments, the method further comprises the step of binding the ligand to a solid surface, whereby the third primer extension product is bound to the solid surface.

[00329] In some embodiments, the method further comprises the steps of: (f) cleaving the RNA portion of the second primer extension products in the DNA-RNA heteroduplex, whereby sequence (E') of the third primer extension products is single stranded, (g) annealing an oligonucleotide comprising a 3' DNA segment (E) that is complementary to sequence (E') and a 5' RNA segment comprising sequence (G), (h) extending the third primer extension products to produce a sequence (G') at their 3' end complementary to sequence (G), and (i) cleaving the RNA from the heteroduplex created in step (h) to produce a single-stranded portion of the third primer extension products corresponding to sequence (G'). The cleaving of RNA in steps (f) and (i) can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (e). The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 3' DNA segment is complementary to all of sequence (E'), in other embodiments, the 3' DNA segment is complementary to portion of sequence (E'). In some embodiments, 5' RNA segment comprising sequence (G) is partly complementary to sequence (E'). In some embodiments, the oligonucleotide is optionally extended

IDT Exhibit 1020.0849

from its 3' end to produce a plurality of oligonucleotide extension products hybridized to the third primer extension products and displacing the DNA portion of the second primer extension products. In some embodiments, the third primer comprises a sequence (F), such that the oligonucleotide extension products will comprise a sequence (F') at or near their 3' end that is complementary to sequence (F). The DNA sequence (G') is created by a DNA polymerase that has RNA dependent DNA polymerase activity. This step creates an RNA/DNA heteroduplex region that can be used for further manipulation of the second primer extension products.

[00330] In some embodiments the method further comprises binding the ligand to a solid surface, whereby the third primer extension products comprising sequence (G') are bound to the solid surface. The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic. The receptor bound to the solid surface is a member of the ligand-receptor pair such that binding of the ligand results in attaching the second primer extension products to the solid surface. In some embodiments, the second primer extension products are still hybridized to the first primer extension product when they are attached to the solid surface. In some embodiments, the second primer extension products are removed from the third primer extension products such that single stranded polynucleotides are attached to the solid surface. The method produces nucleic acids that are bound to a solid surface having a defined sequence (E') and (G') at their 3' ends. The specific sequence (G') can be a site for primer hybridization and further analysis or amplification of the nucleic acids bound to the beads. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (F) at or near its 5' end. One aspect of the invention comprises amplification of the nucleic acids bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (G') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. When the sequence (G') acts as a site to which a composite amplification primer binds, the amplified products that are produced have the sequence (E) (and a portion of sequence (G) at their 5' ends. The amplified products have a sequence (D') at or near their 3' ends. Where the third primer comprises the sequence (F), the amplified products also have the sequence (F'), complementary to (F) at or near their 3' ends. Thus the method produces amplified product with defined sequences at or near both its 3' and 5' ends.

[00331] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (E') and (G') at its 3' end, the sequence (D) at or near its 5' end (and in some embodiments also a specific sequence (F) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target DNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target DNA.

IDT Exhibit 1020.0850

[00332] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of amplification products that are substantially complementary the second primer extension products. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (G') on the third primer extension product for the hybridization of primer. In some embodiments, the amplification produces single stranded amplified products, In some embodiments, the amplification provides double stranded products. The third primer extension products comprise the specific sequence (D'). In some embodiments, the third primer comprises a specific sequence (F), which thus becomes incorporated into the third primer extension products. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the third primer extension products comprise both as sequence (F) at or near its 5' end and a sequence (G') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (G') and the other designed to hybridize to sequence (F'), the complement of sequence (B), or sequence (D), can be used to carry out a PCR reaction to exponentially produce double stranded amplified products.

[00333] One aspect of the invention is a method of amplifying a sequence representative of a sequence within a target DNA comprising the above steps and further comprising: (j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the third primer extension products complementary to sequence (G'); (k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified products; (l) cleaving the RNA from the amplified products; and (m) repeating steps (j) to (l) to produce multiple copies of amplified products wherein the 5' portion of the amplified products have a sequence (E) complementary to sequence (E') and the 3' end of the amplified products have sequence (D') complementary to sequence (D) and optionally sequence (F') complementary to sequence (F).

[00334] The step of binding the polynucleotides to the solid surface through the ligand can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00335] In some embodiments the amplification is carried out such that the amplified products are not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified products remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, amplified products are generated that may float freely in solution, but which comprise a sequence, for example sequence (E) or sequence (D'), that allows them to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface (e.g. sequence (E') or sequence (D). In some embodiments, the amplified product is a single-

IDT Exhibit 1020.0851

stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00336] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer third primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the third primer extension products bound to the beads is then carried out such that the clonal amplification of a plurality of third primer extension products is achieved. This clonal amplification in microreactors can be performed on a sample of target DNA such as genomic, wherein the plurality of third primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target DNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (E') or sequence (D)), which are complementary to sequence (E) or sequence (D') on the amplified product.

[00337] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single third primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a third primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00338] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (D) (or F) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (D') (or F') at their 5' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer third primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of third primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (D') (and/or F') on the amplified product to the sequence (D) (and/or F) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it. Where these sequences are representative of the target DNA, the plurality of beads can constitute a library representative of such DNA.

[00339] After the amplified products are bound to the beads by hybridization, the (D) sequences on the beads can be extended to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (E') near their 5' ends. Where the (D) (and or F) sequences are covalently attached to the beads, this method provides for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the

IDT Exhibit 1020.0852

beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00340] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (E') and a portion of sequence (G'). This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (E) that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein.

[00341] The sequencing can be used to reveal information about the target DNA, for example the genomic DNA. The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

[00342] The amplified products with defined 3' and 5' ends can be used in the methods described herein for the amplified products produced by the other methods. They can be used, for example, for bridge PCR, rolling circle amplification, and strand displacement amplification.

[00343] A schematic exemplary of an embodiment of the invention relating to an alternative method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target is shown in Figure 12. The steps shown in Figure 12 are applicable to product that is generated from the alternative method for either RNA or from DNA as described above. Step I comprises the steps of annealing an oligonucleotide comprising a 3' DNA segment (E) that is complementary to sequence (E') and a 5' RNA segment comprising sequence (G). Step II comprises extending the third primer extension product to produce a sequence (G') at its 3' end complementary to sequence (G). Step III comprises cleaving the RNA from the heteroduplex created in step II to produce a single-stranded portion of the third primer extension product corresponding to sequence (G').

[00344] Step IV comprises annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the third primer extension product complementary to sequence (G'). Step V comprises extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product. Step VI comprises cleaving the RNA from the amplified product. The product of step VI can hybridize to another amplification primer, thus allowing steps IV to VI to be repeated to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence (E) complementary to sequence (E') and the 3' end of the amplified product has sequence (D') complementary to sequence (D) and optionally sequence (F') complementary to sequence (F).

Alternative method for generating a polynucleotide bound to a solid surface

[00345] One aspect of the invention is a method for attaching a polynucleotide sequence that is representative of a sequence within a nucleic acid target molecule to a solid surface. The terms solid

IDT Exhibit 1020.0853

surface and solid support are used interchangeably herein. The polynucleotide sequence that is produced is representative of the sequence within a nucleic acid target molecule if it is either the same as, or complementary to the sequence within the target nucleic acid. Where the target nucleic acid is double stranded, the method can produce sequences that are representative of both of the strands. The polynucleotide can be, for example either DNA or RNA.

[00346] The first step of the method comprises step: (a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P), is complementary to a target nucleic acid and a 5' portion of the of the primer, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid. In some embodiments, the 3' portion of the primer that is complementary to the target nucleic acid is a specific sequence. For example, where a specific region of interest of a target nucleic acid that is known or suspected to be upstream of a specific sequence on the target nucleic acid, sequence (P) of the composite primer can be designed to hybridize to this specific sequence on the target nucleic acid such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target nucleic acid comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence (P) can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence (P) comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence such as sequence (P) at the 3' end of the primer can be useful for performing a global amplification of a nucleic acid target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target nucleic acid. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample, or to determine gene copy number in a DNA sample.

[00347] The first primer extension product comprises a 5' portion comprising sequence (A). Sequence A comprises RNA. In some embodiments, sequence (A) is RNA, and sequence (P) is DNA. In other embodiments, sequence (A) will comprise some DNA. In some embodiments, sequence (P) will comprise some RNA. In some embodiments, sequence (A) and sequence (P) are adjacent.

[00348] The method further comprises step: (b) separating or removing the first primer extension product from the target nucleic acid. The first primer extension product can be separated from the target nucleic acid by a variety of methods. In some cases the separation can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. Other methods of

IDT Exhibit 1020.0854

separating the first primer extension product from the target nucleic acid involve selectively cleaving or degrading the target nucleic acid. Where the target nucleic acid is RNA, the cleaving or degrading can be accomplished by denaturing or heating the sample to degrade RNA or with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or chemically. In some embodiments, the target nucleic acid is completely cleaved or degraded. In other embodiments, the target nucleic acid is only partly cleaved or degraded. The amount of cleavage or degradation required is that amount which will allow the extension of the second primer. In some embodiments, the cleavage or degradation is carried out partially, and the fragments of the target nucleic acid that remain can constitute the second primer for step (c).

[00349] The method further comprises step: (c) extending a second primer to produce a double-stranded product comprising a second primer extension product hybridized to the first primer extension product, wherein the second primer comprises a 3' segment complementary to a portion of the first primer extension product and 5' segment non-complementary sequence (B) to the first primer extension product, whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer and a portion of the 5' end of the second primer extension product comprises non-complementary sequence (B).

[00350] The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. The second primer is generally a tailed primer having a 3' portion which is complementary to the first primer extension product, and a 5' portion, sequence (B), which is not complementary to the first primer extension product. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA or target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target nucleic acid. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target nucleic acid (sense copies). The second primer comprises a sequence (B) that is homologous to a sequence (B) on a solid support.

[00351] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded since sequence (B) is not hybridized to the first primer extension product.

IDT Exhibit 1020.0855

[00352] The method further comprises step: (d) adding an exonuclease to the double-stranded DNA/RNA hybrid, whereby single stranded 3' nucleotides are removed from the first primer extension product. Non-limiting examples of an exonuclease include single-strand specific 3'-exonucleases such as exonuclease 1. The exonuclease should remove all of the single-stranded 3' nucleotides which are not hybridized to sequence (B). In some embodiments, the exonuclease may remove additional 3' nucleotides which are hybridized to the second primer extension product. In other embodiments, a polymerase comprising exonuclease activity may be used. Non-limiting examples include a T4 polymerase comprising 3' exonuclease activity.

[00353] The method further comprises step: (e) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product. The extension of the first primer extension product is carried out with a DNA polymerase as described herein and is generally carried out with a DNA-dependent DNA polymerase if the second primer extension product contains only DNA or with RNA-dependent DNA polymerase if the second primer extension product contains a RNA sequence . The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the second primer extension product.

[00354] The method further comprises step: (f) denaturing the first and second primer extension products. The first primer extension product can be separated from the second primer extension product by denaturation. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents.

[00355] The method further comprises step: (g) binding the sequence (B') of the first primer extension product to a third primer comprising sequence (B) bound to a solid surface, whereby the first primer extension product is attached to the solid surface. The third primer comprises an oligonucleotide with sequence (B) that is complementary to the sequence (B') of the first primer extension product and results in attaching the single stranded first primer extension product to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A) at its 5' end. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B') at or near its 3' end.

[00356] Step (g) of binding the polynucleotides to the solid surface sequence (B) can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00357] The method further comprises step: (h) extending the sequence (B) of the third primer to produce a double-stranded product comprising a third primer extension product hybridized to the first primer extension product, wherein the 5'end of the third primer comprises a sequence (B) complementary to the sequence (B') of the first primer extension product, whereby a portion of the 3' end of the third primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer.

IDT Exhibit 1020.0856

[00358] The extension of the third primer is carried out with a DNA polymerase as described herein. The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the third primer extension product.

[00359] The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00360] A schematic exemplary of an embodiment of the invention relating to an alternative method for generating polynucleotide bound to a solid surface is shown in Figure 13. The figure shows a target nucleic acid (RNA or DNA) and a chimeric RNA/DNA first primer. The primer is first annealed to the target nucleic. Step Ia (RNA target) and Ib (DNA target) illustrates extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target nucleic acid and a 5' portion, sequence (A), of the of the primer is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid, forming an RNA/DNA hybrid. The sequence complementary to a target nucleic acid can be a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step IIa (RNA target) and IIb (DNA target) represent separating the target nucleic acid from the RNA/DNA hybrid. Separation of the target nucleic acid can be accomplished thermally, chemically, or enzymatically, e.g. with RNase H. The second primer comprising a 5' sequence (B) is then annealed to the first primer extension product. Step IIIa and IIIb illustrate extending a second primer, comprising a 5' sequence (B) and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. This embodiment provides for attachment of the first primer extension product to the solid surface by creating a sequence (B'), allowing attachment to a solid surface comprising sequences (B) attached thereto. Step IV shows the removal of 3' nucleotides from the 3' region of the first primer extension product that is not hybridized to the second primer extension product. This step may be done using exonucleases or

IDT Exhibit 1020.0857

using DNA polymerase comprising exonuclease activity. Step V shows extension of the first primer extension product by DNA polymerase to generate a sequence (B'), complementary to the sequence (B) of the second primer extension product. Step VI shows the denaturation of the first and second primer extension product by methods described previously. The first primer extension product comprising a sequence (B') and a defined sequence (A) at its 3' end is useful for storage, archiving and analysis as it has a sequence (B') capable of binding to a solid surface. Such first primer extension product also comprises a sequence that is representative of (identical to or substantially identical to) a sequence in the target RNA, so analysis of this product provides information about the target RNA. Step VII shows the binding of sequence (B') of the first primer extension product to a sequence (B) on a solid surface, whereby the first primer extension product becomes bound to the solid surface. Step VIII shows extension of the immobilized sequence (B) oligonucleotide on the solid support using DNA and RNA-dependent DNA polymerase, resulting in another DNA/RNA heteroduplex.

Alternative method for generating a polynucleotide for binding to a solid surface from an RNA target

[00361] The invention provides methods, compositions and kits for copying, storing, and amplifying polynucleotides having sequences related to target ribonucleic acid (RNA) sequences. The methods provide for amplification of a single RNA species or pool of RNA species. The methods are suitable for, for example, generation of libraries, including cDNA libraries. The methods can generate single stranded RNA or DNA products, which are readily suitable for multiplex analysis by microarray technologies, as well as electrophoresis-based technologies such as differential display, and for sequencing.

[00362] The methods of the invention can copy, store, and amplify one or more species of RNA, such as a pool of RNA sequences, and is most particularly suitable for the amplification of all RNA (such as whole transcriptome or total RNA) sequences in a biological sample. Thus, one of the major advantages of the methods of the invention is the ability to copy, store, and amplify an entire pool of sequences, which is essential for the ability to analyze the gene expression profile in cells, such as the cells in a biological sample of interest. The methods of the invention have the potential of amplifying a multiplicity, a large multiplicity, and in some embodiments all RNA (such as whole transcriptome or total RNA in a sample) sequences in a sample.

[00363] Insofar as many mRNAs have a unique polyA 3'-end, the amplification initiated from the 3'-end sequence of mRNAs is most common for preparation of cDNA libraries and subsequent sequence analysis for determination of gene expression profiling or other applications. The methods of the invention are similarly suited for preparation of libraries of amplified 3'-portions of mRNAs. The sequence of the first primer used in the methods of invention can be designed to be complementary to a multiplicity, or all, of the mRNA species in the sample by using random sequences, according to methods known in the art. The methods are also useful for whole transcriptome amplification. The methods of the invention can be used for the total RNA in samples such as viral RNA.

[00364] An aspect of the invention is a method for generating a polynucleotide comprising a sequence (B') for binding to a solid surface from a RNA target comprising the step of: (a) extending

IDT Exhibit 1020.0858

a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. This extension is generally performed with an enzyme comprising RNA-dependent DNA polymerase activity.

[00365] In some embodiments, the 3' portion of the primer that is complementary to the target RNA is a specific sequence. For example, where a specific region of interest of a target RNA that is known or suspected to be upstream of a specific sequence on the target RNA, the sequence that is complementary to the target RNA of the first primer can be designed to hybridize to this specific sequence on the target RNA such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target RNA comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence that is complementary to the target RNA can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence that is complementary to the target RNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a target RNA, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target RNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample. In some embodiments more than one type of sequence that is complementary to the target RNA can be used, for instance both a primer with a random sequence and a primer, or combination of primers with a specific sequence complementary to RNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00366] The method further comprises the step of: (b) removing the target RNA from the RNA/DNA hybrid. In some embodiments, the removal of the target RNA from the RNA/DNA hybrid involves selectively cleaving or degrading the target RNA. In some cases, the complex comprising the first primer extension product and the nucleic acid can be heated in reaction conditions comprising Mg, which leads to cleaving the RNA including the RNA in the RNA/DNA hybrid. The RNA can also be removed from the RNA/DNA hybrid by denaturation, performed, for example by heating the sample (thermal methods), or by adding denaturing agents or using a combination of heating the sample and adding denaturing agents. The cleaving can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or a combination of RNase enzymes, or chemically. In some embodiments, the target RNA is completely cleaved. In

IDT Exhibit 1020.0859

other embodiments, the target RNA is only partly cleaved or degraded. The amount of cleaving required is that amount which will allow the extension of the second primer.

[00367] The method further comprises the step of: (c) extending a second primer, comprising a 3' segment complementary to a portion of the first primer extension product and a 5' segment non-complementary to the first primer extension product comprising sequence (B), to produce a double-stranded DNA product with a DNA/RNA heteroduplex at one end, wherein the double-stranded product comprises a second primer extension product hybridized to the first primer extension product and wherein a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer.

[00368] The extension of the second primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used. The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target RNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target RNA (sense copies). The second primer comprises a sequence (B) that is homologous to a sequence (B) on a solid support.

[00369] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded since sequence (B) on the second primer does not hybridize to the first primer extension product.

[00370] The method further comprises the step of: (d) adding an exonuclease to the double-stranded DNA product, whereby single stranded 3' nucleotides are removed from the 3' region of the first primer extension product that is not hybridized to the second primer extension product. Non-limiting examples of an exonuclease include single-strand specific 3'-exonucleases such as exonuclease 1. The exonuclease should remove all of the single-stranded 3' nucleotides which are not hybridized to sequence (B). In some embodiments, the exonuclease may remove additional 3' nucleotides which are hybridized to the second primer extension product. In other embodiments, a

IDT Exhibit 1020.0860

polymerase comprising exonuclease activity may be used. Non-limiting examples include a T4 polymerase comprising 3' exonuclease activity.

[00371] The method further comprises the step of: (e) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product. The extension of the first primer extension product is carried out with a DNA polymerase as described herein and is generally carried out with a DNA-dependent DNA polymerase if the second primer extension product contains only DNA or with RNA-dependent DNA polymerase if the second primer extension product contains a RNA sequence. The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the second primer extension product.

[00372] The method further comprises the step of: (f) denaturing the double-stranded DNA product. The first primer extension product can be separated from the second primer extension product by denaturation. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents.

[00373] The method further comprises the step of: (g) attaching the single-stranded first primer extension product to a solid support by annealing sequence (B') to the solid support comprising an oligonucleotide attached thereto, comprising a sequence (B). The oligonucleotide or third primer comprises an oligonucleotide sequence (B) that is complementary to the sequence (B') of the first primer extension product and results in attaching the single stranded first primer extension product to the solid surface. The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A) at its 5' end. As described above, in some embodiments, the nucleic acid bound to the bead also comprises sequence (B') at or near its 3' end.

[00374] Step (g) of binding the polynucleotides to the solid surface sequence (B) can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00375] The method further comprises the step of: (h) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the second extension product, wherein the third primer extension product comprises a 3' sequence (A'), whereby a DNA/RNA heteroduplex at one end is generated. The extension of the third primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used. The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the third primer extension product.

[00376] The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some

IDT Exhibit 1020.0861

embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00377] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in co-pending application.

[00378] One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification.

[00379] In some embodiments, the amplification comprises the steps of : (i) cleaving the RNA region from the first polynucleotide product hybridized to the third primer extension product using RNase H; (j) annealing an amplification primer to sequence (A') on the single-stranded portion of the third primer extension product, wherein the amplification primer has a DNA portion and a 5' RNA portion; (k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the solid support; (l) repeating steps (i) to (k) to produce multiple copies of an amplified product wherein the amplified product comprises sequence (B') at its 3' end; and (m) capturing the amplified product on the solid support wherein the solid support comprises sequence (B).

[00380] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the bound nucleic acids has a specific sequence (A') at its 3' end and also a specific sequence (B) at its 5' end, and where the different nucleic acids have different intervening

IDT Exhibit 1020.0862

sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target RNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target RNA.

[00381] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00382] The step of attaching or binding the polynucleotides to the solid surface through the sequence (B), step (g), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00383] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary to the third primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (A') on the third primer extension product for the hybridization of primer. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction (PCR). For example where the third primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (A') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (A') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00384] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface, e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00385] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer first primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the first primer extension products bound to the

IDT Exhibit 1020.0863

beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target RNA, such as whole transcriptome or total RNA, wherein the plurality of first primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface, e.g. sequence (A') or sequence (B), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00386] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single first primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00387] A schematic exemplary of an embodiment of the invention relating to this method for generating a polynucleotide for binding to a solid surface from an RNA target is shown in Figure 14. The figure shows a target RNA and a chimeric RNA/DNA first primer. The primer is first annealed to the target RNA. Step I illustrates extension of the first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA, to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. The sequence complementary to a target RNA can be a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step II represents separation of the target RNA from the RNA/DNA hybrid. The separation can be accomplished thermally, chemically, or enzymatically, e.g. with RNase H. The second primer comprising a 5' sequence (B) is then annealed to the first primer extension product. Step III illustrate extending a second primer, comprising a 5' sequence (B) and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. This embodiment provides for attachment of the first primer extension product to the solid surface by creating a sequence (B'), allowing attachment to a solid surface comprising sequences (B) attached thereto. Step IV shows the removal of 3' nucleotides from the 3' region of the first primer extension product that is not hybridized to the second primer extension product. This step may be done using exonucleases or using DNA polymerase comprising exonuclease activity. Step V shows extension of the first primer extension product by DNA polymerase to generate a sequence (B'), complementary to the sequence (B) of the second primer extension product. Step VI shows the denaturation of the first and second primer extension product by methods described previously. The first primer extension product

IDT Exhibit 1020.0864

comprising a sequence (B') and a defined sequence (A) at its 3' end is useful for storage, archiving and analysis as it has a sequence (B') capable of binding to a solid surface. Such first primer extension product also comprises a sequence that is representative of (identical to or substantially identical to) a sequence in the target RNA, so analysis of this product provides information about the target RNA. Step VII shows the binding of sequence (B') of the first primer extension product to a sequence (B) on a solid surface, whereby the first primer extension product becomes bound to the solid surface. Step VIII shows extension of the immobilized sequence (B) oligonucleotide on the solid support using DNA and RNA-dependent DNA polymerase, resulting in another DNA/RNA heteroduplex.

[00388] In Figure 15, Step I illustrates amplification using single primer isothermal linear amplification (SPIA) wherein RNaseH cleaves the RNA from the DNA/RNA heteroduplex, a chimeric RNA/DNA primer binds to sequence (A'), and DNA polymerase with strand displacement activity is used to extend the chimeric primer to produce amplified product of the first primer extension product. Step II illustrates amplified products that are produced from repeated rounds of SPIA from Step I of the figure. Since the amplification products are generated in close proximity to access sequence (B) on the solid support, the amplified products are captured on the solid support through hybridization of the 3' sequence (B') to sequence (B) for further manipulations as described herein such as clonal amplification.

Alternative method for generating a polynucleotide for binding to a solid surface from a DNA target

[00389] The methods of the present invention can be used to analyze the DNA (e.g. genomic DNA) samples that are important for many studies. The methods can be used for high-throughput genomic analysis, and can be used for forensic and paleoarcheology work which can be severely limited by nucleic acid sample size. The methods can be used, for example, for the genotyping of multiple loci in the study of complex diseases. The methods can also be used for the determination of genomic instability in various pathological conditions such as cancer, which is most precisely carried out in well defined cell populations, such as that obtained by laser capture micro-dissection or cell sorting. The DNA amplification technologies described herein provide global amplification of very small polynucleotide samples, for example, from one or a very few cells.

[00390] One aspect of the invention is a method for generating from a DNA target a polynucleotide comprising a sequence for binding to a solid surface comprising the steps of: (a) denaturing a double-stranded target DNA. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents, or using a combination of heating the sample and adding denaturing agents.

[00391] The method further comprises step: (b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension product hybridized to the target DNA and comprising sequence (A) at its 5' end. The enzyme that carries out step (b) is generally a DNA polymerase. In some cases a mixture of DNA polymerases can be used. This extension is generally performed with

IDT Exhibit 1020.0865

an enzyme comprising DNA-dependent DNA polymerase activity. The sequence that is complementary to the target DNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a DNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target DNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the number of gene copies in a DNA sample, or obtaining sequence information. In some embodiments, the extension of one first primer, will result in the release of a downstream first primer extension product. This can occur throughout the target DNA resulting in the release of multiple first primer extension products from the target DNA. This process can occur simultaneously on both of the strands of the double-stranded DNA target, thus creating first primer extension products complementary to sequences in both strands.

[00392] In some embodiments, the first primer extension step is carried out with a DNA polymerase capable of extension at elevated temperature that is not compatible with subsequent hybridization of the random sequence to the displaced primer-extension product. For example, Bst DNA polymerase can be used which is active at elevated temperature. The reaction can be carried out stepwise, first with incubation at a lower temperature such as about 25 °C, followed by incubation at higher temperature such as about 50 °C. In some embodiments, the first incubation is carried out below about 30 °C, and the second incubation is carried out above about 40 °C. In some embodiments, a DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer. Mixtures of DNA polymerases can also be useful.

[00393] The method further comprises step: (c) separating the first primer extension product from the target DNA. In some embodiments, the separation can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. The amount of cleaving required is that amount which will allow the extension of the second primer.

[00394] The method further comprises step: (d) annealing to the first primer extension product and extending a second primer comprising a 3' complementary DNA region that comprises a random sequence, wherein the second primer is a tailed primer comprising a 5' sequence (B), to form a double stranded product comprising a first primer extension product and a second primer extension product, whereby a double-stranded product with a DNA/RNA heteroduplex at one end is generated. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used. This step may be carried out with or without prior denaturation. If carried out without denaturation, generally, only the single stranded displaced first primer extension product will hybridize to the second primer. Generally the second primer does not comprise RNA. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer comprises a random primer sequence that randomly binds to

IDT Exhibit 1020.0866

the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target DNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target DNA, or a sequence common to a family of DNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target DNA.

[00395] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first composite primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (d). The primer extension results in a product that is at least partially double stranded since sequence (B) not does hybridize to the first primer extension product. The method produces a nucleic acid that comprises a sequence (B), allowing it to be bound to a solid surface by hybridization to its complement, which is immobilized on the solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead.

[00396] The method further comprises step: (e) adding an exonuclease to the double-stranded DNA product, whereby single stranded 3' nucleotides are removed from the 3' region of the first primer extension product that is not hybridized to the second primer extension product. Non-limiting examples of an exonuclease include single-strand specific 3'-exonucleases such as exonuclease 1. The exonuclease should remove all of the single-stranded 3' nucleotides which are not hybridized to sequence (B). In some embodiments, the exonuclease may remove additional 3' nucleotides which are hybridized to the second primer extension product. In other embodiments, a polymerase comprising exonuclease activity may be used. Non-limiting examples include a T4 polymerase comprising 3' exonuclease activity.

[00397] The method further comprises step: (f) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product. The extension of the first primer extension product is carried out with a DNA polymerase as described herein and is generally carried out with a DNA-dependent DNA polymerase if the second primer extension product contains only DNA or with RNA-dependent DNA polymerase if the second primer extension product contains a RNA sequence. The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the second primer extension product.

[00398] The method further comprises step: (g) denaturing the double-stranded DNA product. The first primer extension product can be separated from the second primer extension product by denaturation. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents.

IDT Exhibit 1020.0867

[00399] The method further comprises step: (h) attaching the single-stranded first primer extension product to a solid support by annealing sequence (B') to the solid support comprising an oligonucleotide attached thereto, comprising a sequence (B), whereby a plurality of first primer extension products become bound to the solid surface. The oligonucleotide or third primer comprises a sequence or oligo (B) that is complementary to the sequence (B') of the first primer extension product and results in attaching the single stranded first primer extension product to the solid surface. The method produces a nucleic acid that is hybridize to a sequence on the solid surface and has a specific sequence (A) at its 5' end.

[00400] Step (h) of binding the polynucleotides to the solid surface sequence (B) can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00401] The method further comprises step: (i) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the first primer extension product, wherein the third primer extension product comprises a 3' sequence (A'), whereby a DNA/RNA heteroduplex at one end is generated.

[00402] The extension of the third primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used. The primer extension results in a product that is double stranded and comprises sequences (A) and (B') on the first primer extension product and sequences (B) and (A') on the third primer extension product.

[00403] The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end (and in some embodiments also a specific sequence (B) at its 5' end), and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00404] One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a

IDT Exhibit 1020.0868

composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (A') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification.

[00405] In some embodiments, the amplification comprises the steps of: (j) cleaving the RNA from the first polynucleotide product hybridized to the amplified product using RNase H; (k) annealing an amplification primer to the single-stranded portion of the amplified product complementary to sequence (A'), wherein the amplification primer has a DNA portion and a 5' RNA portion; (l) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the bead or isolated area; (m) repeating steps (j) to (l) to produce multiple copies of an amplified product wherein the amplified product comprises sequence (B') at its 3' end; and (n) capturing the amplified product on the solid support comprising sequence (B).

[00406] In some embodiments, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (A') at its 3' end and also a specific sequence (B) at its 5' end, and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target DNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target DNA.

[00407] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00408] The step of binding the polynucleotides to the solid surface through the sequence (B), step (h), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00409] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary to the third primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification. Generally, the amplification is carried out using the sequence (A') on the third primer extension product for the hybridization of primer. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction (PCR). For example where the third primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (A') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (A') and the other designed to hybridize to sequence (B'), the complement of sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

IDT Exhibit 1020.0869

[00410] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface, e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00411] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer first primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the second primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target DNA, such as genomic DNA, wherein the plurality of second primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target DNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00412] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single first primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00413] A schematic exemplary of an embodiment of the invention relating to method for generating a polynucleotide comprising a sequence (B) for binding to a solid surface from a DNA target is shown in Figure 16. Step I represents denaturing a double-stranded target DNA, for example by raising the temperature. Steps II illustrate annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end. The enzyme that carries out Step II is generally a DNA polymerase. In some cases a mixture of DNA polymerases can be used. In some embodiments, a DNA polymerase with strand displacement activity is used such that a growing first primer extension product can displace a

IDT Exhibit 1020.0870

downstream first primer extension product, producing a plurality of first primer extension products, representing different regions of the sequence of the target DNA are produced. Step III illustrates annealing a second primer to the first primer extension product that has been denatured from the target DNA. Step IV illustrates extending a second primer comprising a 3' DNA region that comprises a random sequence, wherein the primer is a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product. Step V shows the removal of single stranded 3' nucleotides from the 3' region of the first primer extension product that is not hybridized to the second primer extension product. This step may be done using exonucleases. Step VI shows extension of the first primer extension product by DNA polymerase to generate a sequence (B'), complementary to sequence (B) of the second primer extension product. Step VII shows the denaturation of the first and second primer extension product by methods described previously. Step VIII shows the binding of sequence (B') of the first primer extension product to a sequence (B) on a solid surface, whereby the first primer extension product becomes hybridized to the solid surface. The immobilized sequence (B) oligonucleotide on the solid support is extended using DNA and RNA-dependent DNA polymerase, resulting in another DNA/RNA heteroduplex, as illustrated in step IX. Amplification of the first primer extension product takes place using repeated cycles of single primer isothermal linear amplification (SPIA), which comprises the following steps not shown in the figure: RNaseH cleaves the RNA from the DNA/RNA heteroduplex, a chimeric RNA/DNA primer binds to sequence (A'), and DNA polymerase with strand displacement activity is used to extend the chimeric primer to produce amplified product of the first primer extension product. As shown in step X, since the amplification products are generated in close proximity to access sequence (B) on the solid support, the amplified products are captured on the solid support through hybridization of the 3' sequence (B') to sequence (B) for further manipulations as described herein such as clonal amplification.

Alternative method for generating a polynucleotide having a defined 3' and 5' sequences from an RNA target

[00414] One aspect of the invention is a method for generating a polynucleotide having defined 3' and 5' sequences from a RNA target. The method utilizes a composite RNA/DNA oligonucleotide to generate an oligonucleotide extension product comprising sequences (A) and (C), which will allow extension of the third primer on the solid surface such that the third primer extension product comprises a sequence (C') at its 3' end than can be used as a site for isothermal amplification in a manner such that the sequence (A) is present at or near the 5' end of the amplified product produced in this amplification.

[00415] The method comprises the steps: (a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. This extension is generally performed with an enzyme comprising RNA-dependent DNA polymerase activity. In some embodiments, the 3' portion of the primer that is complementary to the target RNA is a specific sequence. For example, where a specific region of interest of a target RNA that is

IDT Exhibit 1020.0871

known or suspected to be upstream of a specific sequence on the target RNA, the sequence that is complementary to the target RNA of the first primer can be designed to hybridize to this specific sequence on the target RNA such that extension of the primer results in producing a first primer extension product that is complementary to such upstream region. The specific sequence may be common to a family of target RNA. A combination of primers with various specific sequences at the 3' end can also be useful. In some embodiments, such as where the target RNA comprises mRNA, and the mRNA comprises a plurality of sequences, each having a 3' poly-A segment; the specific sequence that is complementary to the target RNA can comprise a sequence that will hybridize to the poly-A region of the mRNA, thus allowing the extension of the first primer to produce a plurality of first primer extension products, each of which is complementary to the region of an mRNA molecule adjacent to the poly-A region. In some embodiments, the sequence that is complementary to the target RNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a RNA target, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target RNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the level of expression in an mRNA sample. In some embodiments more than one type of sequence that is complementary to the target RNA can be used, for instance both a primer with a random sequence and a primer, or combination of primers with a specific sequence complementary to RNA can be used. In some embodiments, multiple primers comprising different specific sequences can be used.

[00416] The method further comprises step: (b) removing the target RNA from the RNA/DNA hybrid. In some embodiments, the cleaving of the target RNA from the RNA/DNA hybrid involves selectively cleaving or degrading the target RNA. In some cases the complex comprising the first primer extension product and the nucleic acid can be heated in reaction conditions comprising Mg, which leads to cleaving the RNA including the RNA in the RNA/DNA hybrid. The RNA can also be removed from the RNA/DNA hybrid by denaturation, performed, for example by heating the sample (thermal methods), or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. The cleaving can be accomplished with an enzyme that cleaves RNA from an RNA/DNA hybrid such as RNase H, or a combination of RNase enzymes, or cleaving can be accomplished chemically or by heating. In some embodiments, the target RNA is completely cleaved. In other embodiments, the target RNA is only partly cleaved or degraded. The amount of cleaving required is that amount which will allow the extension of the second primer.

[00417] The method further comprises step: (c) extending a second primer, comprising a 3' segment complementary to a portion of the first primer extension product and a 5' segment non-complementary to the first primer extension product comprising sequence (B), to produce a double-stranded product with a DNA/RNA heteroduplex at one end; wherein the double-stranded product comprises a second primer extension product hybridized to the first primer extension product, and

IDT Exhibit 1020.0872

whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer.

[00418] The extension of the second primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used.

[00419] The second primer can comprise RNA, DNA, or can be a composite primer comprising both RNA and DNA. In some embodiments, the second primer can comprise a specific primer sequence that is designed to hybridize to a specific sequence in the first primer extension product. In some embodiments the second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target RNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target RNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target RNA, or a sequence common to a family of RNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target RNA (sense copies).

[00420] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (c). The primer extension results in a product that is at least partially double stranded since sequence (B) does not hybridize to the first primer extension product. The product further comprises a DNA-RNA heteroduplex region. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead.

[00421] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in copending application.

IDT Exhibit 1020.0873

[00422] The method further comprises step: (d) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single-stranded. The cleaving of RNA can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (c).

[00423] The method further comprises step: (e) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA sequence (A) that is complementary to sequence (A') and a 5'-RNA segment comprising sequence (C) that is non-complementary to the second primer extension product. The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 5' DNA segment is complementary to all of sequence (A'), in other embodiments, the 5' DNA segment is complementary to portion of sequence (A'). In some embodiments, 5' RNA segment comprising sequence (C) is partly complementary to sequence (A').

[00424] The method further comprises step: (f) extending the oligonucleotide at the 3' segment to form an oligonucleotide extension product hybridized to the second primer extension product. In some embodiments, the oligonucleotide is extended from its 3' end to produce an oligonucleotide extension product hybridized to the second primer extension product and displaces the first primer extension product. The second primer comprises a sequence (B), such that the oligonucleotide extension product will comprise a sequence (B') at or near its 3' end that is complementary to sequence (B). An optional step that may be added using a DNA polymerase that has RNA dependent DNA polymerase activity is extension of the second primer extension product to create a heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C). This step creates an RNA/DNA heteroduplex region.

[00425] The method further comprises step: (g) denaturing the double-stranded DNA product. The first primer extension product can be separated from the second primer extension product by denaturation. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents.

[00426] The method further comprises step: (h) attaching the single-stranded first primer extension product to a solid support by annealing sequence (B') to the solid support comprising a sequence (B). The oligonucleotide bound to the solid surface comprises sequence (B) such that binding of sequence (B') of the oligonucleotide primer extension product results in attaching the oligonucleotide primer extension product to the solid surface. The method produces a nucleic acid that is hybridized to sequence (B) on a solid surface that has a specific sequence (A) and (C) at its 5' end.

[00427] The method further comprises step: (i) extending sequence (B) on the solid support to produce a third primer extension product, comprising a 3' sequence (A') and (C'), whereby a DNA/RNA heteroduplex at one end is generated.

[00428] The extension of the third primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used.

IDT Exhibit 1020.0874

[00429] The primer extension results in a product that is double stranded and comprises sequences (B'), (A), (C) on the oligonucleotide primer extension product and sequences (B), (A') and (C') on the third primer extension product.

[00430] In some embodiments, the sample comprising the target RNA is in a sample that also comprises DNA. In such cases, it can be advantageous to add a selective DNA dependent DNA polymerase inhibitor such as actinomycin such that it is present during step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a). The presence of a DNA dependent DNA polymerase inhibitor such as actinomycin is particularly advantageous when a first primer comprising a random sequence is used, as the inhibitor allows for the selective creation of first primer extension products to RNA without the need of separating the RNA from the DNA. This is also advantageous when the priming is carried out at specific target sequences since the sequence may be the same on the DNA when the DNA and RNA in the sample represent total nucleic acid from the same biological entity, for example, human tissue, animal tissue, and the like. The use of DNA dependent DNA polymerase inhibitors such as actinomycin is described in co-pending application.

[00431] The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') and (C') at its 3' end and a sequence (B) at or near its 5' end. The specific, or universal, sequence (A') or (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. The specific sequence (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (C') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. When the sequence (C') acts as a site to which a composite amplification primer binds, the amplified product that is produced has the sequence (A) and a portion of sequence (C) at its 5' end. The third primer comprises the sequence (B) and the amplified product also has the sequence (B'), complementary to (B) at or near its 3' end. Thus the method produced amplified product with defined sequences at or near both its 3' and 5' ends.

[00432] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids hybridized to a solid surface is created in which each of the nucleic acids has a specific sequence (A) and (C) at its 5' end and also a specific sequence (B') at its 3' end, and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target RNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target RNA.

[00433] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

IDT Exhibit 1020.0875

[00434] The step of binding the polynucleotides to the solid surface through sequence (B'), step (i), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00435] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary to the third primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification.

[00436] Generally, the amplification is carried out using the sequence (C') on the third primer extension product for the hybridization of a primer such as a composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the third primer extension product comprises both a sequence (B) at or near its 5' end and a sequence (C') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (C') and the other designed to hybridize to sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00437] In some embodiments, the amplification is performed by a method comprising the following steps: (j) cleaving the RNA from the heteroduplex polynucleotide product hybridized to the amplified product using RNase H to produce a single-stranded portion of the third primer extension product corresponding to sequence (C'); (k) annealing an amplification primer to the single-stranded portion of the third primer extension product complementary to sequence (C'), wherein the amplification primer has a DNA portion and a 5' RNA portion; (l) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the solid support; (m) repeating steps (j) to (l) to produce multiple copies of the amplified product comprising sequences (A) and (B'); and (n) capturing the amplified product on the solid support wherein the solid support comprises sequence (B).

[00438] This amplification method, utilizing a sequence (B) and (C'), allows for the production of an amplified product comprising a sequence (B') at or near its 3' end that is substantially complementary to sequence (B), and a sequence (A) near its 5' end that is complementary to sequence (A'), thus producing an amplified polynucleotide product with defined 3' and 5' ends.

[00439] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other

IDT Exhibit 1020.0876

portion of the solid surface by hybridization to a complementary sequence bound to such surface, e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00440] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer oligonucleotide primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the third primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target RNA, such as whole transcriptome or total RNA, wherein the plurality of third primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00441] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single third primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a second primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00442] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B) at their 5' ends, whereby upon the amplification of step (n) multiple copies of amplified product comprising sequence (B') at their 3' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer third primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of third primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (B') on the amplified product to the sequence (B) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it. Where these sequences are representative of the target RNA, the plurality of beads can constitute a library representative of such RNA.

[00443] Since the amplification products are generated in close proximity to access sequence (B) on the solid support, the amplified products are captured on the solid support through hybridization of the 3' sequence (B') to sequence (B) for further manipulations as described herein. The (B) sequences on the beads can be extended along the amplified product by a DNA polymerase or

IDT Exhibit 1020.0877

mixture of polymerases to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (A') near their 3' ends. This method provides for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00444] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (A') and a portion of sequence (C'). Here, the portion of sequence (C') is the DNA portion of the chimeric amplification primer (C) that does is generally not cleaved by RNase H and therefore becomes incorporated into the amplified product. This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (A') that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein.

[00445] The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

[00446] In some embodiments the methods of the invention provide for performing bridge PCR comprising making amplified product as described above with defined 3' and 5' ends, and further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the defined 3' and 5' sequences, for example, A and B' sequences, on the amplified product in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

[00447] In some embodiments the methods of the invention provide for making amplified product as described above with defined 3' and 5' ends and further performing rolling circle amplification comprising performing the steps of: (o) hybridizing the amplified product to a nucleic acid sequence comprising regions complementary to A and B' sequences in close proximity; (p) optionally extending the gap with a DNA polymerase enzyme; (q) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

[00448] In some embodiments, the rolling circle amplification uses primers complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product. In some cases, such a primer can be an oligonucleotide attached to a solid surface, thus resulting in amplified product bound to the surface

[00449] In some embodiments the methods of the invention provide for performing PCR comprising making amplified product as described above with defined 3' and 5' ends, further comprising the

IDT Exhibit 1020.0878

steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

[00450] In some embodiments the methods of the invention provide for performing strand displacement amplification (SDA) comprising making amplified product as described above with defined 3' and 5' ends, wherein the defined 3' and 5' ends, for example, sequences (A) and (B'), in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

[00451] A schematic exemplary of an embodiment of the invention relating to generating a polynucleotide having a defined 3' and 5' sequences is shown in Figure 17. The figure shows a target RNA and a chimeric RNA/DNA first primer. The primer is first annealed to the target RNA. Step I illustrates extension of the first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA, to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid. The sequence complementary to a target RNA can be a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step II represents separation of the target RNA from the RNA/DNA hybrid. The separation can be accomplished thermally, chemically, or enzymatically, e.g. with RNase H. The second primer comprising a 5' sequence (B) is then annealed to the first primer extension product. Step III illustrate extending a second primer, comprising a 5' sequence (B) and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. In step IV, cleavage of the RNA from the first primer extension product in the DNA-RNA heteroduplex occurs such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. As shown, the cleavage is performed using RNase H. Chemical and thermal means can alternatively be employed. Step V illustrates annealing to the second primer extension product a chimeric oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C). Step VI is an optional step and may occur when there is DNA polymerase comprising RNA-dependent DNA polymerase activity. In this step the second primer extension product is extended along sequence C. Step VII illustrates extension of the oligonucleotide at the 3' end, generating an oligonucleotide extension product which is hybridized to the second primer extension product and comprises a sequence (B'), complementary to sequence (B) on the second primer extension product. The first primer extension product is displaced during 3'-extension of the oligonucleotide. Step VIII illustrates the denaturation of the chimeric oligonucleotide extension product from the second primer extension product. Step IX illustrates binding of the chimeric oligonucleotide extension product to a third primer comprising sequence (B) on the solid surface. Step X illustrates extension of the third primer to create a strand complementary to the chimeric oligonucleotide extension product comprising sequence (C') and (A').

IDT Exhibit 1020.0879

[00452] In Figure 18, Step I illustrates the steps of: cleaving the RNA from the DNA-RNA heteroduplex created in step X above to produce a single-stranded portion of the third primer extension product corresponding to sequence (C'); annealing an amplification primer to sequence (C'), wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension product complementary to sequence (C'); extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product. These steps can be repeated to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence complementary to sequence (A'). In the embodiment illustrated, the amplified product shown comprises a defined sequences on both the 5' and 3' ends. Step II illustrates that the amplification products are generated in close proximity to the sequence (B) immobilized on the bead and will thus allow the amplification product to be captured on the bead via hybridization of its sequence (B'). DNA polymerase extends immobilized sequence (B) along the hybridized amplification product to generate bound nucleic acid comprising specific sequences (B) and (A'). Also shown on the amplified product is a portion of sequence (C') that is the DNA portion of the chimeric amplification primer (C). This portion does not generally become cleaved by RNase H and therefore becomes incorporated into the amplified product.

Alternative method for generating a polynucleotide having a defined 3' and 5' sequences from a DNA target

[00453] The method utilizes a composite RNA/DNA oligonucleotide to generate an oligonucleotide extension product comprising sequences (A) and (C), which will allow extension of the third primer on the solid surface such that the third primer extension product comprises a sequence (C') at its 3' end than can be used as a site for isothermal amplification in a manner such that the sequence (A) is present at or near the 5' end of the amplified product produced in this amplification, and where a second primer comprising sequence (B) is used, amplified products with defined sequences at both the 3' and 5' ends can be produced.

[00454] The method comprises the step: (a) denaturing a double-stranded target DNA. Double stranded DNA can be denatured, for example by heating, or by the addition of denaturing agents.

[00455] The method further comprises step: (b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a first primer extension product hybridized to the target DNA and comprising sequence (A) at its 5' end. This extension is generally performed with an enzyme comprising DNA-dependent DNA polymerase activity. The sequence that is complementary to the target DNA comprises a random sequence, such that the extension of the first primer results in a plurality of first primer extension products complementary to the sequences adjacent to the sequence where each random species hybridizes. The use of a random sequence at the 3' end of the primer can be useful for performing a global amplification of a target DNA, generating a plurality of sequences which together can represent, for example substantially the whole sequence of the target DNA. In some embodiments, the relative amounts of the various sequences can be used to quantitate the relative amount of a given sequence in a sample, for example to determine the number of gene copies in a

IDT Exhibit 1020.0880

DNA sample, or obtaining sequence information. In some embodiments, the extension of one first primer, will result in the release of a downstream first primer extension product. This can occur throughout the target DNA resulting in the release of multiple first primer extension products from the target DNA. This process can occur simultaneously on both of the strands of the double-stranded DNA target, thus creating first primer extension products complementary to sequences in both strands.

[00456] In some embodiments, the first primer extension step is carried out with a DNA polymerase capable of extension at elevated temperature that is not compatible with subsequent hybridization of the random sequence to the displaced primer-extension product. For example, Bst DNA polymerase can be used which is active at elevated temperature. The reaction can be carried out stepwise, first with an incubation at a lower temperature such as about 25 °C, followed by an incubation at higher temperature such as about 50 °C. In some embodiments, the first incubation is carried out below about 30 °C, and the second incubation is carried out above about 40 °C. In some embodiments, a DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer. Mixtures of DNA polymerases can also be useful.

[00457] The method further comprises step: (c) separating the first primer extension product from the target DNA. In some embodiments, the separation can be affected by denaturing the complex comprising the first primer extension product and the nucleic acid. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents. The amount of cleaving required is that amount which will allow the extension of the second primer.

[00458] The method further comprises step: (d) annealing to the first primer extension product and extending a second primer comprising a 3' complementary DNA region that comprises a random sequence, wherein the second primer is a tailed primer comprising a 5' sequence (B), to form a double-stranded product comprising a first primer extension product and a second primer extension product, whereby a double-stranded product with a DNA/RNA heteroduplex at one end is generated. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used.

[00459] This step may be carried out with or without prior denaturation. If carried out without denaturation, generally, only the single stranded displaced first primer extension product will hybridize to the second primer. Generally the second primer does not comprise RNA. The extension of the second primer is carried out with a DNA polymerase as described herein. The second primer comprises a random primer sequence that randomly binds to the first primer extension product. Extension of the second primer comprising a random sequence produces a plurality of second primer extension products. The use of a random sequence at the 3' end of the primer is useful, for example, in performing global amplification of a target DNA, whereby a plurality of second primer extension products are produced which is representative of the sequence of the target DNA. In some embodiments, for example where the first primer is designed to hybridize to a specific sequence on a target DNA, or a sequence common to a family of DNA targets, random priming by the second primer ensures amplification of the entire selected target or family of selected

IDT Exhibit 1020.0881

targets. In this embodiment, the second primer extension products comprise sequences which are the same or substantially the same as the sequences in the target DNA.

[00460] The second primer extension product is extended such that the 3' portion of the second primer extension product comprises a sequence (A') which is complementary to sequence (A) of the first composite primer. Since sequence (A) on the first primer extension product comprises RNA, both DNA dependent DNA polymerase activity and RNA dependent DNA polymerase activity are used in step (d). The primer extension results in a product that is at least partially double stranded since sequence (B) does not hybridize to the first primer extension product. The method produces a nucleic acid that comprises a sequence (B), allowing it to be bound to a solid surface by hybridization to its complement, which is immobilized on the solid surface and that has a specific sequence (A') at its 3' end. The specific, or universal, sequence (A') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead.

[00461] The method further comprises step: (e) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. The cleaving of RNA can be performed, for example by treatment with RNase H, which will selectively cleave the RNA portion of the DNA/RNA partial heteroduplex formed in step (d).

[00462] The method further comprises step: (f) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C).

[00463] The oligonucleotide comprises at least one DNA and at least one RNA portion. In some embodiments the 5' DNA segment is complementary to all of sequence (A'), in other embodiments, the 5' DNA segment is complementary to portion of sequence (A'). In some embodiments, 5' RNA segment comprising sequence (C) is partly complementary to sequence (A').

[00464] The method further comprises step: (g) extending the oligonucleotide along the second primer extension product to form an oligonucleotide extension product comprising a sequence (B'), complementary to sequence (B) on the second primer extension product. In some embodiments, the oligonucleotide is extended from its 3' end to produce an oligonucleotide extension product hybridized to the second primer extension product and displaces the first primer extension product. The second primer comprises a sequence (B), such that the oligonucleotide extension product will comprise a sequence (B') at or near its 3' end that is complementary to sequence (B). An optional step that may be added using a DNA polymerase that has RNA dependent DNA polymerase activity is extension of the second primer extension product to create a heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C). This step creates an RNA/DNA heteroduplex region.

[00465] The method further comprises step: (h) denaturing the double-stranded DNA product. The oligonucleotide primer extension product can be separated from the second primer extension product by denaturation. Denaturation can be performed, for example by heating the sample, or by adding a denaturing agent, or using a combination of heating the sample and adding denaturing agents.

[00466] The method further comprises step: (i) attaching the single-stranded oligonucleotide extension product to solid support by annealing sequence (B') to the bead or isolated area

IDT Exhibit 1020.0882

comprising a sequence (B). The third primer comprises an oligonucleotide sequence (B) that is complementary to the sequence (B') of the oligonucleotide primer extension product and results in attaching the single stranded oligonucleotide primer extension product to the solid surface. The method produces a nucleic acid that is hybridized to a solid surface that has a specific sequence (A) at its 5' end.

[00467] Step (i) of binding the polynucleotides to the solid surface sequence (B) can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00468] The method further comprises the step of: (j) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the oligonucleotide extension product, comprising a 3' sequence (A') and (C'), whereby a DNA/RNA heteroduplex at one end is generated. The extension of the third primer is carried out with a DNA polymerase as described herein. In some embodiments, a DNA polymerase comprising both DNA and RNA dependent DNA polymerase activities is used here. In other embodiments, both a RNA dependent DNA polymerase and a DNA dependent DNA polymerase are used. The primer extension results in a product that is double stranded and comprises sequences (B'), (A), (C) on the first primer extension product and sequences (B), (A') and (C') on the third primer extension product.

[00469] The method produces a nucleic acid that is bound to a solid surface that has a specific sequence (A') and (C') at its 3' end and a sequence (B) at or near its 5' end. The specific, or universal, sequence (A') or (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. The specific, or universal, sequence (C') can be a site for primer hybridization and further analysis or amplification of the nucleic acid bound to the bead. One aspect of the invention comprises amplification of the nucleic acid bound to the bead. In some embodiments, the amplification is carried out using isothermal amplification using a composite RNA/DNA primer, RNase H, and a polymerase with strand displacement activity. For this embodiment, the sequence (C') acts as the site to which the composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids bound to a solid surface is created in which each of the nucleic acids has a specific sequence (C') at its 3' end and also a specific sequence (B) at its 5' end, and where the different nucleic acids have different intervening sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target nucleic acid. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target nucleic acid.

[00470] In some embodiments, for example where random sequences at the 3' end of the first and/or second primer are used, a plurality of different nucleic acids hybridized to a solid surface is created in which each of the nucleic acids has a specific sequence (A) and (C) at its 5' end and also a specific sequence (B') at its 3' end, and where the different nucleic acids have different intervening

IDT Exhibit 1020.0883

sequences, wherein the intervening sequences are identical to or substantially identical to the sequences in the target DNA. The set of bound nucleic acids thus generated can be analyzed, for example, by sequencing in order to provide information about the sequence of the target DNA.

[00471] The solid surface can be any of a variety of surfaces, some described in more detail below. The solid surface can be, for example a planar surface, for example, a planar array. In some embodiments the solid surface comprises a plurality of beads. In some embodiments the beads are magnetic.

[00472] The step of binding the polynucleotides to the solid surface through sequence (B'), step (i), can be carried out such that only one nucleic acid is bound to an isolated area of a surface or only one nucleic acid is bound to a single bead. This isolated binding of nucleic acids can be used for clonal amplification of the specific bound nucleic acid in that area or on that bead. Such bound, isolated nucleic acids can also be stored and archived for later analysis, for example by sequencing. The bound, isolated nucleic acids can be amplified, stored, and analyzed multiple times.

[00473] In some embodiments, the method further comprises treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary to the third primer extension product. This step comprises carrying out an amplification reaction wherein the bound nucleic acid acts as a template for the amplification.

[00474] Generally, the amplification is carried out using the sequence (C') on the third primer extension product for the hybridization of a primer such as a composite RNA/DNA amplification primer hybridizes, allowing for amplification. In some embodiments the amplification is an isothermal amplification reaction comprising a composite RNA/DNA primer, RNase H, and a DNA polymerase with strand displacement activity. In some embodiments, the amplification is carried out using polymerase chain reaction, (PCR). For example where the second primer extension product comprises both as sequence (B) at or near its 5' end and a sequence (C') at or near its 3' end, a set of primers, one designed to hybridize to all or a portion of the sequence (C') and the other designed to hybridize to sequence (B), can be used to carry out a PCR reaction to exponentially produce double stranded amplified product.

[00475] In some embodiments, the amplification is performed by a method comprising the following steps: (k) cleaving the RNA from the heteroduplex polynucleotide product hybridized to the amplified product using RNase H to produce a single-stranded portion of the second primer extension product corresponding to sequence (C'); (l) annealing an amplification primer to the single-stranded portion of the amplified product complementary to sequence (C'), wherein the amplification primer has a DNA portion and a 5' RNA portion; (m) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the amplified product on the bead or isolated area; and (n) repeating steps (k) to (m) to produce multiple copies of the second polynucleotide product comprising sequences (A) and (B'). This amplification method, utilizing a sequence (B) and (C'), allows for the production of an amplified product comprises a sequence (B') at or near its 3' end that is substantially complementary to sequence (B), and a sequence (A) near its 5' end that is complementary to sequence (A'), thus producing an amplified polynucleotide product with defined 3' and 5' ends.

IDT Exhibit 1020.0884

[00476] In some embodiments the amplification is carried out such that the amplified product is not attached to the substrate, but is freely dissolved in the solution. In other embodiments, the amplification is carried out such that the amplified product remains bound to the substrate, for example by performing solid phase PCR such as bridge PCR. In yet other embodiments, an amplified product is generated that may float freely in solution, but which comprises a sequence, for example sequence (A) or sequence (B'), that allows it to be captured to another solid surface or other portion of the solid surface by hybridization to a complementary sequence bound to such surface, e.g. sequence (A') or sequence (B). In some embodiments, the amplified product is a single-stranded product and, because it is generated at the solid surface, the amplified product readily captured by complementary sequences, e.g. sequence (B), bound to the surface.

[00477] In one aspect of the invention, a plurality of beads is used, and the methods described above are carried out such that on average, one or fewer oligonucleotide primer extension product molecules are bound per bead. The beads are dispersed into an aqueous solution, and a plurality of microreactors, e.g. droplets, are produced such that on average one or fewer beads is contained within each of the plurality of microreactors. The amplification of the third primer extension products bound to the beads is then carried out such that the clonal amplification of each of the plurality of second primer extension products in the separate microreactors is achieved. This clonal amplification in microreactors can be performed on a sample of target DNA, such as whole transcriptome or total DNA, wherein the plurality of third primer extension products comprise sequences that correspond to most, to substantially all, or to all of the sequences in the target RNA. In some embodiments, the amplified products are captured by bead having attached thereto a plurality of oligonucleotides comprising complementary sequences bound to such surface (e.g. sequence (A') or sequence (B)), which are complementary to sequence (A) or sequence (B') on the amplified product.

[00478] In some embodiments, the plurality of beads, produced as described above, with each bead comprising a single oligonucleotide primer extension product can comprise a library. These libraries can be stored, then later clonally amplified. In some embodiments, a library of beads can comprise a plurality of beads wherein each bead had multiple copies of a single amplification product generated from a third primer extension product. These libraries can be analyzed, for example by sequencing. The libraries can be stored, and later analyzed. In some embodiments the libraries can be stored, then analyzed multiple times.

[00479] In some embodiments, a bead or isolated area of the solid surface comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B) at their 5' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (B') at their 3' end are hybridized to the bead or isolated area. For example, where beads are used, a plurality of beads in a plurality of microreactors wherein, the plurality of beads has, on average one or fewer oligonucleotide primer extension products bound to it and there are, on average, one or fewer beads in each microreactor, a clonal amplification of the plurality of third primer extension products can be carried out, and the amplified products in each of the microreactors will bind to the bead through the sequence (B') on the amplified product to the sequence (B) on the beads. This approach produces a plurality of beads, each with multiple copies of a different sequence bound to it.

IDT Exhibit 1020.0885

Where these sequences are representative of the target DNA, the plurality of beads can constitute a library representative of such DNA.

[00480] After the amplified products are bound to the beads by hybridization, the (B) sequences on the beads can be extended along the amplified product by a DNA polymerase or mixture of polymerases to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and also comprise sequence (A') near their 3' ends. This method provides for the production of beads with polynucleotides complementary to amplified product covalently attached to the beads. Covalently attached polynucleotides such as those produce here are more robust than nucleotides that are attached only by hybridization to the beads. Thus, the covalently attached polynucleotides can be more stable and can be used with analysis methods and sequencing methods that have harsher conditions which would result in the displacement of polynucleotides bound only by hybridization.

[00481] In some embodiments, the amplified product is removed from the covalently bound polynucleotide to render the polynucleotide single stranded. Such single stranded covalently bound polynucleotides comprise a specific sequence at their 3' ends comprising sequence (A') and a portion of sequence (C'). Here, the portion of sequence (C') is the DNA portion of the chimeric amplification primer (C) that does is generally not cleaved by RNase H and therefore becomes incorporated into the amplified product. This specific sequence at the 3' end of the covalently bound polynucleotide can act as a hybridization site for a primer complementary to sequence (A') that can act as a primer to carry out sequencing by any of a variety of sequencing methods, for example, those described herein.

[00482] The sequencing methods can comprise the use of cleavable labeled terminators. The sequencing method can comprise pyrophosphate detection. The sequencing method can comprise an isothermal sequencing method, for example using chimeric primers, RNase H, and a polymerase with strand displacement activity. The sequencing method can also comprise cycle sequencing.

[00483] In some embodiments the methods of the invention provide for performing bridge PCR comprising making amplified product as described above with defined 3' and 5' ends, and further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the defined 3' and 5' sequences, for example, A and B' sequences, on the amplified product in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

[00484] In some embodiments the methods of the invention provide for making amplified product as described above with defined 3' and 5' ends and further performing rolling circle amplification comprising performing the steps of: (o) hybridizing the amplified products to a target nucleic acid comprising regions complementary to A and B' sequences in close proximity; (p) optionally extending the gap with a polymerase enzyme; and (q) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

[00485] In some embodiments, the rolling circle amplification uses primers complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the

IDT Exhibit 1020.0886

amplified product. In some cases, such a primer can be an oligonucleotide attached to a solid surface, thus resulting in amplified product bound to the surface

[00486] In some embodiments the methods of the invention provide for performing PCR comprising making amplified product as described above with defined 3' and 5' ends, further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

[00487] In some embodiments the methods of the invention provide for performing strand displacement amplification (SDA) comprising making amplified product as described above with defined 3' and 5' ends, wherein the defined 3' and 5' ends, for example, sequences (A) and (B'), in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

[00488] A schematic exemplary of an embodiment of the invention relating to generating a polynucleotide having a defined 3' and 5' sequences is shown in Figure 19. The figure shows a double stranded target DNA that is denatured in step I. In step II, a chimeric RNA/DNA first primer is first annealed to the target DNA and is extended to form a first primer extension product hybridized to the target DNA, forming a RNA/DNA hybrid. The first primer comprises a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target DNA. The sequence complementary to a target DNA can be a specific sequence, a sequence that will hybridize to Poly-A, a sequence common to a plurality of regions (consensus sequence), or a random sequence. Step III represents separation of the target DNA from the RNA/DNA hybrid. The separation can be accomplished thermally, chemically, or enzymatically, e.g. with RNase H. A second primer comprising a 5' sequence (B) is then annealed to the first primer extension product. Step IV illustrate extension of the second primer, comprising a 5' sequence (B) and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer. In step V, cleavage of the RNA from the first primer extension product in the DNA-RNA heteroduplex occurs such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded. As shown, the cleavage is performed using RNase H. Chemical and thermal means can alternatively be employed. Step VI illustrates annealing to the second primer extension product a chimeric oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C). Step VII is an optional step and may occur when there is DNA polymerase comprising RNA-dependent DNA polymerase activity. In this step the second primer extension product is extended along sequence C. Step VIII illustrates extension of the oligonucleotide at the 3' end, generating an oligonucleotide extension product which is hybridized to the second primer extension product and comprises a sequence (B'), complementary to sequence (B) on the second primer extension product. The first primer extension product is displaced during 3'-extension of the oligonucleotide. Step IX illustrates the denaturation of the chimeric

IDT Exhibit 1020.0887

oligonucleotide extension product from the second primer extension product. Step X illustrates binding of the chimeric oligonucleotide extension product to a third primer comprising sequence (B) on the solid surface. Step XI illustrates extension of the third primer to create a strand complementary to the chimeric oligonucleotide extension product comprising sequence (C') and (A'). Amplification of the oligonucleotide extension product comprising specific sequences (B) and (A') at the 3' and 5' ends, respectively, take place by carrying out the steps illustrated in Figure 18.

[00489] In one aspect of the present invention, methods are provided for amplifying a target nucleic acid or its complement on a solid support to form a plurality of amplified products comprising clonally amplifying said target sequence or its complement by linear amplification. In some embodiments, the method comprises amplification using a single primer. In some embodiments, the method comprises amplification from a double-stranded nucleic acid having a single-stranded 3' overhang at one end. In some embodiments, the method comprises amplification using a DNA-RNA chimeric primer. In some embodiments, the method comprises a combination of linear amplification using a single primer from a double-stranded nucleic acid having a single-stranded 3' overhang at one end using a DNA-RNA chimeric primer. In some embodiments, the target nucleic acid sequence is coupled to a solid support. In some cases, the amplification is isothermal. In some cases, the solid support is a bead. In some cases, the amplification results in at least 10,000; 100,000; one million or more copies of the target sequence, its complement, or a portion thereof. In some cases, the solid surface comprises a plurality of primers of substantially identical sequence. In some cases, the solid surface consists of a plurality of primers of substantially identical sequence. In some cases, the target nucleic acid sequence is a linear template. In some cases, the target nucleic acid sequence is greater than about 150, 200, 250, 300, 400, 500, 1kb, 2kb, 5kb, 10kb or more in length. In some cases, the target nucleic acid is less than about 1 megabase, 100kb, 50kb, 10kb, 5kb, or less in length.

[00490] In one aspect of the present invention, methods are provided for clonally amplifying a target nucleic acid sequence or its complement by delivering the target nucleic acid sequence into an emulsion and performing linear amplification of the target nucleic acid sequence inside the emulsion. In some cases, the method further comprises the step of forming the emulsion first. In some cases, the emulsion is formed around the target nucleic acid and/or the target nucleic acid in a reaction mixture comprising amplification reagents including but not limited to buffers; one or more enzymes such as a DNA polymerase with substantial strand displacement activity, exonuclease, and RNase H; salts; primers including chimeric primers, amplification primers, and all-DNA primers, oligonucleotides including chimeric oligonucleotides; and dNTPs. In some cases, the method may further comprise amplifying the target nucleic acid in the presence of a solid surface such as a bead, a substantially planar array, an isolated surface, or a well in a plate. In some cases the amplification results in a plurality of non-multimerized individual amplification products. These are physically separated or chemically separated amplified products. In some cases, the amplification results in a lower error rate than PCR such as for example fewer errors than 2 in every 100,000 nucleotides incorporated into amplified product.

IDT Exhibit 1020.0888

General Techniques

[00491] The practice of the invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M. J. Gait, ed., 1984); "Animal Cell Culture" (R. I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Current Protocols in Molecular Biology" (F. M. Ausubel et al., eds., 1987, and periodic updates); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994).

[00492] The terms "Polynucleotide," or "nucleic acid," as used interchangeably herein, refer to polymers of nucleotides of any length, and include DNA and RNA. The nucleotides can be deoxyribonucleotides, ribonucleotides, modified nucleotides or bases, and/or their analogs, or any substrate that can be incorporated into a polymer by DNA or RNA polymerase. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and their analogs. If present, modification to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications include, for example, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, cabamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example, proteins (e.g., nucleases, toxins, antibodies, signal peptides, ply-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide(s). Further, any of the hydroxyl groups ordinarily present in the sugars may be replaced, for example, by phosphonate groups, phosphate groups, protected by standard protecting groups, or activated to prepare additional linkages to additional nucleotides, or may be conjugated to solid supports. The 5' and 3' terminal OH can be phosphorylated or substituted with amines or organic capping groups moieties of from 1 to 20 carbon atoms. Other hydroxyls may also be derivatized to standard protecting groups. Polynucleotides can also contain analogous forms of ribose or deoxyribose sugars that are generally known in the art, including, for example, 2'-O-methyl-, 2'-O-allyl, 2'-fluoro- or 2'-azido-ribose, carbocyclic sugar analogs, .alpha.-anomeric sugars, epimeric sugars such as arabinose, xyloses or lyxoses, pyranose sugars, furanose sugars, sedoheptuloses, acyclic analogs and abasic nucleoside analogs such as methyl riboside. One or more phosphodiester linkages may be replaced by alternative linking groups. These alternative linking groups include, but are not limited to, embodiments wherein phosphate is replaced by P(O)S("thioate"), P(S)S ("dithioate"), "(O)NR.sub.2 ("amidate"), P(O)R, P(O)OR', CO or CH.sub.2 ("formacetal"), in which each R or R' is independently H or substituted or unsubstituted alkyl (1-20 C) optionally containing an ether (--O--) linkage, aryl, alkenyl, cycloalkyl, cycloalkenyl or araldyl.

IDT Exhibit 1020.0889

Not all linkages in a polynucleotide need be identical. The preceding description applies to all polynucleotides referred to herein, including RNA and DNA.

[00493] "Oligonucleotide," as used herein, generally refers to short, generally single stranded, generally synthetic polynucleotides that are generally, but not necessarily, less than about 200 nucleotides in length. The oligonucleotide(s) are generally comprised of a sequence of at least 5 nucleotides, generally from about 10 to about 100 nucleotides, about 20 to about 50 nucleotides, and often about 10 to about 30 nucleotides in length. The oligonucleotides of the invention can be DNA, RNA, DNA-RNA, or other polynucleotide. The terms oligo or sequence may be used interchangeable herein.

[00494] Various techniques can be employed for preparing an oligonucleotide utilized in the present invention. Such oligonucleotide can be obtained by biological synthesis or by chemical synthesis. For short sequences (up to about 100 nucleotides) chemical synthesis will frequently be more economical as compared to the biological synthesis. In addition to economy, chemical synthesis provides a convenient way of incorporating low molecular weight compounds and/or modified bases during the synthesis step. Furthermore, chemical synthesis is very flexible in the choice of length and region of the target polynucleotide binding sequence. The oligonucleotide can be synthesized by standard methods such as those used in commercial automated nucleic acid synthesizers. Chemical synthesis of DNA on a suitably modified glass or resin can result in DNA covalently attached to the surface. This may offer advantages in washing and sample handling. For longer sequences standard replication methods employed in molecular biology can be used such as the use of M13 for single stranded DNA as described by J. Messing (1983) Methods Enzymol, 101, 20-78.

[00495] In the present invention, nucleoside triphosphates are incorporated by a polymerase enzyme in the extension of the primer to produce an extension product. Nucleoside triphosphates are generally nucleosides having a 5'-triphosphate substituent. The nucleosides are pentose sugar derivatives of nitrogenous bases of either purine or pyrimidine derivation, covalently bonded to the 1'-carbon of the pentose sugar, which is usually a deoxyribose or a ribose. The purine bases include adenine(A), guanine (G), inosine (I), and derivatives and analogs thereof. The pyrimidine bases include cytosine (C), thymine (T), uracil (U), and derivatives and analogs thereof. Nucleoside triphosphates include deoxyribonucleoside triphosphates such as the four common triphosphates dATP, dCTP, dGTP and dTTP and ribonucleoside triphosphates such as the four common triphosphates rATP, rCTP, rGTP and rUTP. The term "nucleoside triphosphates" also includes derivatives and analogs thereof, which are exemplified by those derivatives that are recognized in a similar manner to the underivatized nucleoside triphosphates. Examples of such derivatives or analogs, by way of illustration and not limitation, are those which are biotinylated, amine modified, alkylated, and the like and also include phosphorothioate, phosphite, ring atom modified derivatives, and the like.

[00496] As used herein, the term "nucleotide" generally refers to a base-sugar-phosphate combination that is the monomeric unit of nucleic acid polymers, i.e., DNA and RNA. In some aspects of the invention modified nucleotides are used, for example, where a nucleotide is connected to a ligand. A modified nucleotide is generally the unit in a nucleic acid polymer that results from

IDT Exhibit 1020.0890

the incorporation of a modified nucleoside triphosphate during an amplification reaction and therefore becomes part of the nucleic acid polymer.

[00497] As used herein, a nucleoside is generally a base-sugar combination or a nucleotide lacking a phosphate moiety.

[00498] A "primer" is generally a nucleotide sequence (i.e. a polynucleotide), generally with a free 3'-OH group, that hybridizes with a template sequence (such as a target RNA, or a primer extension product) and is capable of promoting polymerization of a polynucleotide complementary to the template. A "primer" can be, for example, an oligonucleotide. It can also be, for example, a sequence of the template (such as a primer extension product or a fragment of the template created following RNase cleavage of a template-DNA complex) that is hybridized to a sequence in the template itself (for example, as a hairpin loop), and that is capable of promoting nucleotide polymerization. Thus, a primer can be an exogenous (e.g., added) primer or an endogenous (e.g., template fragment) primer.

[00499] The primers of the invention are usually oligonucleotide primers. A primer is generally an oligonucleotide that is employed in an extension on a polynucleotide template. The oligonucleotide primer is usually a synthetic nucleotide that is single stranded, containing a sequence at its 3'-end that is capable of hybridizing with a sequence of the target polynucleotide. Normally, the 3' region of the primer that hybridizes with the target nucleic acid has at least 80%, preferably 90%, more preferably 95%, most preferably 100%, complementarity to a sequence or primer binding site. The number of nucleotides in the hybridizable sequence of a specific oligonucleotide primer should be such that stringency conditions used to hybridize the oligonucleotide primer will prevent excessive random non-specific hybridization. Usually, the number of nucleotides in the hybridizing portion of the oligonucleotide primer will be at least as great as the defined sequence of the target polynucleotide, namely, at least ten nucleotides, at least about 15 nucleotides and generally from about 10 to about 200, usually about 20 to about 50 nucleotides.

[00500] A "random primer," as used herein, is a primer that comprises a sequence that is designed not necessarily based on a particular or specific sequence in a sample, but rather is based on a statistical expectation (or an empirical observation) that the sequence of the random primer is hybridizable (under a given set of conditions) to one or more sequences in the sample. The random primers used herein are generally tailed random primers comprising a 3' segment that acts as a random primer to the target polynucleotide, and a 5' sequence that generally does not hybridize to the target polynucleotide. The sequence of a random primer (or its complement) may or may not be naturally-occurring, or may or may not be present in a pool of sequences in a sample of interest. The amplification of a plurality of polynucleotides, e.g. DNA or RNA species in a single reaction mixture would generally, but not necessarily, employ a multiplicity, preferably a large multiplicity, of random primers. As is well understood in the art, a "random primer" can also refer to a primer that is a member of a population of primers (a plurality of random primers) which collectively are designed to hybridize to a desired and/or a significant number of target sequences. A random primer may hybridize at a plurality of sites on a nucleic acid sequence. The use of random primers provides a method for generating primer extension products complementary to a target polynucleotide which does not require prior knowledge of the exact sequence of the target. In some embodiments one portion of a primer is random, and another portion of the primer comprises a defined sequence. For

IDT Exhibit 1020.0891

example, in some embodiments, a 3' portion of the primer will comprise a random sequence, while the 5' portion of the primer comprises a defined sequence. In some embodiments a 3' random portion of the primer will comprise DNA, and a 5' portion defined portion of the primer will comprise RNA, in other embodiments, both the 3' and 5' portions will comprise DNA.

[00501] Composite primers are employed in certain embodiments of the invention. Composite primers are primers that are composed of RNA and DNA portions. In some aspects, the composite primer is a tailed composite primer comprising, for example, a 3' DNA portion and a 5' RNA portion. In the tailed composite primer, a 3' portion, all or a portion of which comprises DNA is complementary to a polynucleotide; and a 5' portion, all or a portion of which comprises RNA, is not complementary to the polynucleotide and does not hybridize to the polynucleotide under conditions in which the 3' portion of the tailed composite primer hybridizes to the polynucleotide target. When the tailed composite primer is extended with a DNA polymerase, a primer extension product with a 5' RNA portion comprising a defined sequence can be created. This primer extension product can then have a second primer anneal to it, which can be extended with a DNA polymerase to create a double stranded product with an RNA/DNA heteroduplex comprising a defined sequence at one end. The RNA portion can be selectively cleaved from the partial heteroduplex to create a double stranded DNA with a 3' single stranded overhang which can be useful for a various aspects of the present invention including allowing for isothermal amplification using a composite amplification primer.

[00502] In other aspects, the composite primer is an amplification composite primer. In the amplification composite primer, both the RNA and the DNA portions are generally complementary and hybridize to a sequence in the polynucleotide to be copied or amplified. In some embodiments, a 3' portion of the amplification composite primer is DNA and a 5' portion of the composite amplification primer is RNA. The composite amplification primer is designed such that the primer is extended from the 3' DNA portion to create a primer extension product. The 5 'RNA portion of this primer extension product, in a DNA-RNA heteroduplex is susceptible to cleavage by RNase H, thus freeing a portion of the polynucleotide to the hybridization of an additional composite amplification primer. The extension of the additional composite primer by a DNA polymerase with strand displacement activity releases the primer extension product from the original primer and creates another copy of the sequence of the polynucleotide. Repeated rounds of primer hybridization, primer extension with strand displacement DNA synthesis, and RNA cleavage create multiple copies of the sequence of the polynucleotide. Composite primers are described in more detail below.

[00503] Polymerases are used in the methods of the invention, for example to extend primers to produce extension products. A polymerase, or nucleotide polymerase, is a catalyst, usually an enzyme, for forming an extension of a polynucleotide along a DNA or RNA template where the extension is complementary thereto. The nucleotide polymerase is a template dependent polynucleotide polymerase and utilizes nucleoside triphosphates as building blocks for extending the 3'-end of a polynucleotide to provide a sequence complementary with the polynucleotide template. Usually, the catalysts are enzymes, such as DNA polymerases, for example, prokaryotic DNA polymerase (I, II, or III), T4 DNA polymerase, T7 DNA polymerase, Klenow fragment, Bst DNA

IDT Exhibit 1020.0892

polymerase, reverse transcriptase, Bca polymerase, Vent DNA polymerase, Pfu DNA polymerase, Taq DNA polymerase, and the like, derived from any source such as cells, bacteria, such as E. coli, plants, animals, virus, thermophilic bacteria, and so forth. RNA polymerases include T7 RNA polymerase, AMV polymerase, Q-beta-replicase, and so forth. In some cases, Bst DNA Polymerase Large Fragment can be used. Bst DNA polymerase Large Fragment is the portion of the *Bacillus stearothermophilus* DNA Polymerase protein that contains the 5′ → 3′ polymerase activity, but lacks the 5′ →3′ exonuclease domain. Where the polymerase forms an extension product on a DNA template, it is referred to herein as a DNA dependent polymerase. Where the polymerase forms an extension product on a RNA template, it is referred to herein as a RNA dependent polymerase.

[00504] A "labeled dNTP," or "labeled rNTP," as used herein, refers, respectively, to a dNTP or rNTP, or analogs thereof, that is directly or indirectly attached with a label. For example, a "labeled" dNTP or rNTP, may be directly labeled with, for example, a dye and/or a detectable moiety, such as a member of a specific binding pair (such as biotin-avidin). A "labeled" dNTP or rNTP, may also be indirectly labeled by its attachment to, for example, a moiety to which a label is/can be attached. A dNTP or rNTP, may comprise a moiety (for example, an amine group) to which a label may be attached following incorporation of the dNTP or rNTP into an extension product. Useful labels in the present invention include fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, green fluorescent protein and the like), radioisotopes (e.g., ^{3}H, ^{35}S, ^{32}P, ^{33}P, ^{125}I, or ^{14}C), enzymes (e.g., LacZ, horseradish peroxidase, alkaline phosphatase,), digoxigenin, and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads.

[00505] The term "double stranded product" is used herein to refer to products that are produced by the extension of a primer. It is understood that the products are at least partially double stranded, for example, in the region comprising the primer extension product and its complement. The double stranded product need not be completely double-stranded, and may have single stranded regions. It is also understood that the double stranded product can have heteroduplex regions in which one strand comprises RNA and the complementary strand comprised DNA in that region.

Amplification

[00506] Some aspects of the invention comprise the amplification of polynucleotide molecules or sequences within the polynucleotide molecules. Amplification generally refers to a method that results in the formation of one or more copies of a nucleic acid or polynucleotide molecule or in the formation of one or more copies of the complement of a nucleic acid or polynucleotide molecule. Amplifications can be used in the invention, for example, to amplify or analyze a polynucleotide bound to a solid surface. The amplifications can be performed, for example, after archiving the samples in order to analyze the archived polynucleotide.

[00507] In some aspects of the invention, exponential amplification of nucleic acids or polynucleotides is used. These methods often depend on the product catalyzed formation of multiple copies of a nucleic acid or polynucleotide molecule or its complement. The amplification products are sometimes referred to as "amplicons." One such method for the enzymatic amplification of specific double stranded sequences of DNA is known as the polymerase chain reaction (PCR). This

IDT Exhibit 1020.0893

in vitro amplification procedure is based on repeated cycles of denaturation, oligonucleotide primer annealing, and primer extension by thermophilic template dependent polynucleotide polymerase, resulting in the exponential increase in copies of the desired sequence of the polynucleotide analyte flanked by the primers. The two different PCR primers, which anneal to opposite strands of the DNA, are positioned so that the polymerase catalyzed extension product of one primer can serve as a template strand for the other, leading to the accumulation of a discrete double stranded fragment whose length is defined by the distance between the 5' ends of the oligonucleotide primers.

[00508] Another method for amplification involves amplification of a single stranded polynucleotide using a single oligonucleotide primer. The single stranded polynucleotide that is to be amplified contains two non-contiguous sequences that are complementary to one another and, thus, are capable of hybridizing together to form a stem-loop structure. This single stranded polynucleotide already may be part of a polynucleotide analyte or may be created as the result of the presence of a polynucleotide analyte.

[00509] Another method for achieving the result of an amplification of nucleic acids is known as the ligase chain reaction (LCR). This method uses a ligase enzyme to join pairs of preformed nucleic acid probes. The probes hybridize with each complementary strand of the nucleic acid analyte, if present, and ligase is employed to bind each pair of probes together resulting in two templates that can serve in the next cycle to reiterate the particular nucleic acid sequence.

[00510] Another method for achieving a nucleic acid amplification is the nucleic acid sequence based amplification (NASBA). This method is a promoter-directed, enzymatic process that induces in vitro continuous, homogeneous and isothermal amplification of a specific nucleic acid to provide RNA copies of the nucleic acid. The reagents for conducting NASBA include a first DNA primer with a 5' tail comprising a promoter, a second DNA primer, reverse transcriptase, RNAse-H, T7 RNA polymerase, NTPs and dNTPs.

[00511] Another method for amplifying a specific group of nucleic acids is the Q-beta-replicase method, which relies on the ability of Q-beta-replicase to amplify its RNA substrate exponentially. The reagents for conducting such an amplification include "midi-variant RNA" (amplifiable hybridization probe), NTPs, and Q-beta-replicase.

[00512] Another method for amplifying nucleic acids is known as 3SR and is similar to NASBA except that the RNAse-H activity is present in the reverse transcriptase. Amplification by 3SR is an RNA specific target method whereby RNA is amplified in an isothermal process combining promoter directed RNA polymerase, reverse transcriptase and RNase H with target RNA.

[00513] Another method for amplifying nucleic acids is the Transcription Mediated Amplification (TMA) used by Gen-Probe. The method is similar to NASBA in utilizing two enzymes in a self-sustained sequence replication. See U.S. Pat. No. 6,946,254.

[00514] Another method for amplification of nucleic acids is Strand Displacement Amplification (SDA) (Westin et al 2000, Nature Biotechnology, 18, 199-202; Walker et al 1992, Nucleic Acids Research, 20, 7, 1691-1696), which is an isothermal amplification technique based upon the ability of a restriction endonuclease such as HincII or BsoBI to nick the unmodified strand of a hemiphosphorothioate form of its recognition site, and the ability of an exonuclease deficient DNA polymerase such as Klenow exo minus polymerase, or Bst polymerase, to extend the 3'-end at the

IDT Exhibit 1020.0894

nick and displace the downstream DNA strand. Exponential amplification results from coupling sense and antisense reactions in which strands displaced from a sense reaction serve as targets for an antisense reaction and vice versa.

[00515] Another method for amplification of nucleic acids is Rolling Circle Amplification (RCA) (Lizardi et al. 1998, Nature Genetics, 19:225-232). RCA can be used to amplify single stranded molecules in the form of circles of nucleic acids. In its simplest form, RCA involves the hybridization of a single primer to a circular nucleic acid. Extension of the primer by a DNA polymerase with strand displacement activity results in the production of multiple copies of the circular nucleic acid concatenated into a single DNA strand.

[00516] In some embodiments of the invention, RCA is coupled with ligation. For example, a single oligonucleotide can be used both for ligation and as the circular template for RCA. This type of polynucleotide can be referred to as a "padlock probe" or a "RCA probe". For a padlock probe, both termini of the oligonucleotide contains sequences complementary to a domain within a nucleic acid sequence of interest. The first end of the padlock probe is substantially complementary to a first domain on the nucleic acid sequence of interest, and the second end of the padlock probe is substantially complementary to a second domain, adjacent to the first domain near the first domain. Hybridization of the oligonucleotide to the target nucleic acid results in the formation of a hybridization complex. Ligation of the ends of the padlock probe results in the formation of a modified hybridization complex containing a circular polynucleotide. In some cases, prior to ligation, a polymerase can fill in the gap by extending one end of the padlock probe. The circular polynucleotide thus formed can serve as a template for RCA that with the addition of a polymerase results in the formation of an amplified product nucleic acid. The methods of the invention described herein, can produce amplified products with defined sequences on both the 5' and 3' ends. Such amplified products can be used as padlock probes.

[00517] Some aspects of the invention utilize the linear amplification of nucleic acids or polynucleotides. Linear amplification generally refers to a method that involve the formation of one or more copies of the complement of only one strand of a nucleic acid or polynucleotide molecule, usually a nucleic acid or polynucleotide analyte. Thus, the primary difference between linear amplification and exponential amplification is that the latter is auto catalyzed, that is, the product serves to catalyze the formation of more product, whereas in the former process the starting sequence catalyzes the formation of product but is not itself replicated. In linear amplification the amount of product formed increases as a linear function of time as opposed to exponential amplification where the amount of product formed is an exponential function of time.

[00518] In some embodiments, amplification methods can be solid-phase amplification, polony amplification, colony amplification, emulsion PCR, bead RCA, surface RCA, surface SDA, etc., as will be recognized by one of skill in the art. In some embodiments, amplification methods that results in amplification of free DNA molecules in solution or tethered to a suitable matrix by only one end of the DNA molecule can be used. Methods that rely on bridge PCR, where both PCR primers are attached to a surface (see, e.g., Adessi et al., Nucleic Acids Research (2000): 28(20): E87) can be used. In some cases the methods of the invention can create a "polymerase colony technology", or "polony", referring to a multiplex amplification that maintains spatial clustering of

IDT Exhibit 1020.0895

identical amplicons (see Harvard Molecular Technology Group and Lipper Center for Computational Genetics website). These include, for example, in situ polonies (Mitra and Church, Nucleic Acid Research 27, e34, Dec. 15, 1999), in situ rolling circle amplification (RCA) (Lizardi et al., Nature Genetics 19, 225, July 1998), bridge PCR (U.S. Pat. No. 5,641,658), picotiter PCR (Leamon et al., Electrophoresis 24, 3769, November 2003), and emulsion PCR (Dressman et al., PNAS 100, 8817, Jul. 22, 2003). The methods of the invention provide new methods for generating and using polonies.

[00519] In some embodiments, amplification methods such as amplification that maintains spatial clustering of identical amplicons, amplification that produces in situ polonies, picotiter amplification, amplification inside microdroplets, amplification inside the aqueous phase of microdroplets of an oil and water emulsion or the like, may provide clonal amplification. "Clonal amplification" as used herein is the amplification of a target sequence or a set of target sequences in a spatially or physically separated manner such that the amplified products are spatially or physically separated from the amplified products of other target sequences. For example, some of the methods of the present invention provide for amplification of a set of target sequences attached to, bound to, or hybridized to a solid surface such as a set of beads, isolated surfaces, or wells in a plate such that each surface on average contains one or fewer target sequences. In some cases, the amplification is carried out under conditions that the amplified products of the target sequence of a given surface are attached to, bound to, or hybridized to that surface. Further, in some cases, the given surface comprises few, substantially no, or no other amplified products from other target sequences. Similarly, clonal amplification may be carried out by separating the target sequences and their products from other target sequences and their products in other ways such as in a set of microreactors such as microdroplets such as in an emulsion, in a set of wells, or by dilution. Additionally clonal amplification may be performed using a combination of compositions and techniques such as by amplifying a set of target sequences on a solid surface or set of solid surfaces such as a bead or set of beads in a set of microdroplets in an emulsion.

Single Primer Isothermal Amplification Using a Complex Comprising an RNA/DNA Partial Heteroduplex as a Template

[00520] In some aspects of the invention, the amplification method that is used is a single primer isothermal amplification using a complex comprising an RNA/DNA partial heteroduplex as a template. In this method, termed single primer isothermal amplification, a complex comprising an RNA/DNA partial heteroduplex is a substrate for further amplification as follows: an enzyme which cleaves RNA sequence from an RNA/DNA hybrid (such as RNase H) cleaves RNA from the partial heteroduplex, leaving a partially double stranded polynucleotide complex comprising a 3' single stranded DNA sequence. The 3' single stranded sequence (formed by cleavage of RNA in the complex comprising an RNA/DNA partial heteroduplex) is generally the complement of the amplification composite primer, and thus forms a specific binding site for a composite primer. Extension of a bound composite primer by a DNA-dependent DNA polymerase with strand displacement activity produces a primer extension product, which displaces the previously bound

IDT Exhibit 1020.0896

cleaved primer extension product, whereby polynucleotide (generally, DNA) product accumulates. See, for example, U.S. Pat. Nos. 6,251,639 and 6,692,918.

[00521] Amplification using a complex comprising an RNA/DNA partial heteroduplex as a template for further amplification (also termed single primer isothermal amplification) generally occurs under conditions permitting composite primer hybridization, primer extension by a DNA polymerase with strand displacement activity, cleavage of RNA from an RNA/DNA hybrid and strand displacement. In so far as the composite primer hybridizes to the 3' single stranded portion (of the partially double stranded polynucleotide which is formed by cleaving RNA in the complex comprising an RNA/DNA partial heteroduplex) comprising, generally, the complement of at least a portion of the composite primer sequence, composite primer hybridization may be under conditions permitting specific hybridization.

[00522] In some embodiments, the methods of the invention result in amplification of a multiplicity, a large multiplicity, or a very large multiplicity of template polynucleotide sequences. In some embodiments, essentially all of the template polynucleotide present in the initial sample (e.g., all of the mRNA or all of the genomic DNA) is amplified. In other embodiments, at least 1, at least 5, at least 10, at least 20. at least 50, at least 100, at least 200, at least 300, or more distinct sequences (such as a gene or other subsegment of a polynucleotide, transcripts of a nucleic acid target, a marker (such as a SNP or other polymorphism) are amplified, as assessed, e.g., by analysis of marker sequences known to be present in the template sample under analysis, using methods known in the art. Template polynucleotide sequences that are amplified may be present on the same polynucleotide (e.g., a chromosome or portion of a chromosome for genomic DNA template or on the same RNA for RNA template) or on different template polynucleotides (e.g., different chromosome or portions of chromosomes for DNA template, or different RNAs for RNA template). In some case, amplification of genomic DNA is exemplified herein, it will be understood by those of skill in the art, however, that the global amplification methods of the invention are suitable for amplification of any pool or subset of polynucleotides.

[00523] In some embodiments, the methods of the invention are used to globally amplify double stranded DNA target. It is understood that in these cases, the amplified product generally is a mixture of sense and antisense copies of the template DNA. In some embodiments, the methods of the invention are used to globally amplify a single stranded DNA or RNA target. In these cases, the amplification product will generally be a copy of either the target polynucleotide (sense copy) or of the complement to the target nucleotide (antisense copy). Whether the sense or antisense copy is produced will depend on the method, as will be understood by one of ordinary skill in the art. In some embodiments, the amplification product of different senses can be annealed to form a double stranded (or partially double stranded) complex. In other embodiments, they can be prevented from annealing (or subsequently denatured) to produce a mixture of single stranded amplification products. The amplified products may be of differing lengths.

[00524] As illustrated in these embodiments, all steps are isothermal (in the sense that thermal cycling is not required), although the temperatures for each of the steps may or may not be the same. It is understood that various other embodiments may be practiced, given the general description

IDT Exhibit 1020.0897

provided above. For example, as described and exemplified herein, certain steps may be performed as temperature is changed (e.g., raised, or lowered).

[00525] For simplicity, the isothermal amplification methods of the invention are described as two distinct steps or phases, above. It is understood that the two phases may occur simultaneously in some embodiments (for example, if the enzyme that cleaves RNA from RNA/DNA hybrid is included in the first reaction mixture).

[00526] Although generally only one composite primer is described above, it is further understood that the amplification methods may be performed in the presence of two or more different composite primers that randomly prime template polynucleotide. In addition, the amplification polynucleotide products of two or more separate amplification reactions conducted using two or more different composite primers that randomly prime template polynucleotide can be combined.

Extension of primers

[00527] The methods of the present invention involve the extension of primers. In general, primers hybridize to, and are extended along (chain extended), a sequence within the target polynucleotide and, thus, the target polynucleotide acts as a template. The extended primers are chain "extension products." The sequence over which the primer is extended may lie between two defined sequences but need not be. In general, the primers hybridize with a sequence within the target polynucleotide. The target sequence usually contains from about 30 to 5,000 or more nucleotides, often 50 to 1,000 nucleotides. The target polynucleotide may be a fraction of a larger molecule or it may be substantially the entire molecule (polynucleotide analyte).

Composite Primers

[00528] Generation of primers suitable for extension by polymerization is well known in the art, such as described in PCT Pub. No. WO99/42618 (and references cited therein). The composite primer comprises a combination of RNA and DNA, with the 3'-end nucleotide being a nucleotide suitable for nucleic acid extension. The 3'-end nucleotide can be any nucleotide or analog that when present in a primer, is extendable by a DNA polymerase when hybridized to a polynucleotide template. Generally, the 3'-end nucleotide has a 3'-OH. Suitable primers include those that comprise at least one portion of RNA and at least one portion of DNA. For example, composite primers can comprise a 5'-RNA portion and a 3'-DNA portion (in which the RNA portion is adjacent to the 3'-DNA portion); or 5'- and 3'-DNA portions with an intervening RNA portion. Accordingly, in one embodiment, the composite primer comprises a 5' RNA portion and a 3'-DNA portion, preferably wherein the RNA portion is adjacent to the 3'-DNA portion. In another embodiment, the composite primer comprises 5'- and 3'-DNA portions with at least one intervening RNA portion (i.e., an RNA portion between the two DNA portions). In yet another embodiment, the composite primer of the invention comprises a 3'-DNA portion and at least one intervening RNA portion (i.e., an RNA portion between DNA portions).

Composite Amplification Primers

IDT Exhibit 1020.0898

[00529] Composite amplification primers are RNA/DNA composite primers that can be used to create multiple copies of (amplify) a polynucleotide sequence isothermally using RNA cleavage, and DNA polymerase activity with strand displacement. Amplification with such primers is described, for example in U. S. Patents 6,251,639, 6,692,918, and 6,946,251. The composite amplification primer comprises sequences capable of hybridizing to a portion of a DNA template, and most often comprises sequences hybridizable to a defined 3'-portion of the DNA.

[00530] A composite amplification primer comprises at least one RNA portion that is capable of (a) binding (hybridizing) to a sequence on a DNA template independent of hybridization of the DNA portion(s) to a sequence on the same extension product; and being cleaved with a ribonuclease when hybridized to the DNA template. The composite amplification primers bind to the DNA template to form a partial heteroduplex in which only the RNA portion of the primer is cleaved upon contact with a ribonuclease such as RNase H, while the DNA template remains intact, thus enabling annealing of another composite primer.

[00531] The composite amplification primers also comprise a 3' DNA portion that is capable of hybridization to a sequence on the DNA template such that its hybridization to the DNA is favored over that of the nucleic acid strand that is displaced from the DNA template by the DNA polymerase. Such primers can be rationally designed based on well known factors that influence nucleic acid binding affinity, such as sequence length and/or identity, as well as hybridization conditions. In a preferred embodiment, hybridization of the 3' DNA portion of the composite primer to its complementary sequence in the second strand cDNA is favored over the hybridization of the homologous sequence in the 5' end of the displaced strand to the second strand cDNA.

Ligands and Receptors

[00532] The present invention utilizes ligands to bind polynucleotides to a solid surface. A ligand is a compound which can bind to a receptor. .The ligand and receptor (antiligand) can be members of a specific binding pair of two different molecules. In some cases, the receptor and or ligand have one or more areas on the surface or in a cavity which gives rise to specific binding. The ligand can be complementary with a particular spatial and polar organization of the receptor. The specific binding pair may be members of an immunological pair such as antigen-antibody, or may be operator-repressor, nuclease-nucleotide, biotin-avidin, hormones-hormone receptors, nucleic acid duplexes, IgG-protein A, DNA-DNA, DNA-RNA, and the like. Examples of ligands and/or receptors include, but are not limited to, agonists and antagonists for cell membrane receptors, toxins and venoms, viral epitopes, hormones such as steroids, hormone receptors, peptides, enzymes and other catalytic polypeptides, enzyme substrates, cofactors, drugs including small organic molecule drugs, opiates, opiate receptors, lectins, sugars, saccharides including polysaccharides, proteins, and antibodies including monoclonal antibodies and synthetic antibody fragments, cells, cell membranes and moieties therein including cell membrane receptors, and organelles. Examples of ligand-receptor pairs include antibody-antigen; lectin-carbohydrate; peptide-cell membrane receptor; protein A-antibody; hapten-antihapten; digoxigenin-anti-digoxigenin; avidin and biotin, enzyme-cofactor and enzyme-substrate.

IDT Exhibit 1020.0899

[00533] In one embodiment, the receptor may comprise an antibody. As used herein, the term "antibody" refers to an immunoglobulin molecule or a fragment of an immunoglobulin molecule having the ability to specifically bind to a particular antigen. The antibody may be an anti-receptor antibody specific for the receptor used in the assay. Thus, the antibody may be capable of specifically binding the receptor as the antigen. Antibodies and methods for their manufacture are well known in the art of immunology. The antibody may be produced, for example, by hybridoma cell lines, by immunization to elicit a polyclonal antibody response, or by recombinant host cells that have been transformed with a recombinant DNA expression vector that encodes the antibody. Illustrative receptors include naturally occurring receptors, e.g., thyroxine binding globulin, antibodies, enzymes, Fab fragments, lectins, nucleic acids, repressors, protection enzymes, protein A, complement component C1q, DNA binding proteins or ligands and the like. In some embodiments, the receptor is an antibody, and the ligand is a molecule or portion of a molecule that is recognized by the antibody such as an epitope, or a hapten.

[00534] In some embodiments it is desirable that the ligand be a small organic molecule, for example, a compound of molecular weight less than about 1500 g/mol, generally between about 100 to about 1000 g/mol, often between about 300 to about 600 g/mol such as biotin or other haptens. The small organic molecule can provide a means for attachment of a nucleotide sequence to a label or to a support.

[00535] Various anti-ligands and ligands can be used (as labels themselves or as a means for attaching a label). In the case of a ligand that has a natural anti-ligand, such as biotin, thyroxine and cortisol, the ligand can be used in conjunction with labeled anti-ligands.

[00536] Attaching the receptors that are specific binders to the ligands to a support or surface may be accomplished by well-known techniques, commonly available in the literature. See, for example, "Immobilized Enzymes," Ichiro Chibata, Halsted Press, New York (1978) and Cuatrecasas, J. Biol. Chem., 245:3059 (1970).

[00537] The ligand and receptor can be members of capture pairs. For example, capture pairs can employ reversible (e. g. , cleavable) or irreversible linkages. Non-limiting examples of reversible linkages include thiol-thiol, digoxigenin/anti- digoxigenin, and linkages using VECTREX (E Avidin DLA (Vector Laboratories, Burlingame, CA), CaptAvidinTM, NeutrAvidinN, and D-desthiobiotin (Molecular Probes, Inc., Eugene, OR).

[00538] In some embodiments, the ligand-receptor pair can be pairs of reactive molecules that can react to form a covalent bond, thus binding the polynucleotide to the surface. For example, the ligand can comprise an amine group bound to the polynucleotide that can be reacted with a functional group attached to the solid surface to form a covalent bond. The amine on the polynucleotide can react, for example with an activated carbonyl group attached to the surface, e.g. an N-hydroxy succinimide (NHS) ester. Functional groups such as N-acylimidazole, 2- or 3-bromoacrylate, cyanuric chloride, disulfide, N- hydroxysuccinimide ester, hydrazide, iodoacetyl, imidoester, isocyanate, isothiocyanate, maleimide, succinimidyl carbonate, acyl chloride, and sulfonyl chloride can be used in order to form a covalent linkage to the solid surface.

DNA Polymerase, and an Agent Capable of Cleaving an RNA-DNA Hybrid

IDT Exhibit 1020.0900

[00539] The isothermal amplification methods of the invention employ the following enzymes: an RNA-dependent DNA polymerase, a DNA-dependent DNA polymerase, and an agent capable of cleaving an RNA strand of an RNA-DNA hybrid (for example, a ribonuclease such as RNase H). One or more of these activities may be found and used in a single enzyme. For example, RNase H activity may be supplied by an RNA-dependent DNA polymerase (such as reverse transcriptase) or may be provided in a separate enzyme. Reverse transcriptases useful for this method may or may not have RNase H activity. Many reverse transcriptases, such as those from avian myeloblastosis virus (AMV-RT), and Moloney murine leukemia virus (MMLV-RT) comprise more than one activity (for example, polymerase activity and ribonuclease activity) and can function in the formation of the double stranded cDNA molecules. However, in some instances, it is preferable to employ a reverse transcriptase which lacks or has reduced levels RNase H activity. Reverse transcriptase devoid of RNase H or with reduced levels of RNase H activity are known in the art, including those comprising a mutation of the wild type reverse transcriptase where the mutation eliminates the RNase H activity. In these cases, the addition of an RNase H from other sources, such as that isolated from E. coli, can be employed for the formation of the double stranded cDNA. The RNA-dependent DNA polymerase activity and DNA-dependent DNA polymerase activity may be provided by the same enzyme (for example, Bst polymerase), or these activities may be provided in separate enzymes. DNA polymerases with strand displacement activity are also useful.

[00540] One aspect of the invention is the formation of a complex comprising an RNA/DNA partial heteroduplex. This process generally utilizes the enzymatic activities of an RNA-dependent DNA polymerase, a DNA-dependent DNA polymerase. Generally, RNA in the RNA/DNA partial heteroduplex is cleaved by an agent (such as an enzyme, such as a ribonuclease) capable of cleaving RNA from an RNA/DNA hybrid, generating a 3' single stranded portion with sequences that are complementary to RNA in a composite primer (and thus forming a binding site for a composite primer).

[00541] RNA-dependent DNA polymerases for use in the methods and compositions of the invention are capable of effecting extension of a primer according to the methods of the invention. Accordingly, a preferred RNA-dependent DNA polymerase is one that is capable of extending a nucleic acid primer along a nucleic acid template that is comprised at least predominantly of ribonucleotides. Suitable RNA-dependent DNA polymerases for use in the methods and compositions of the invention include reverse transcriptase and, for example, a DNA polymerase that possesses both DNA-dependent and RNA-dependent DNA polymerase activity, such as Bst DNA polymerase.

[00542] DNA-dependent DNA polymerases for use in the methods and compositions of the invention are capable of effecting extension of the composite primer according to the methods of the invention. Accordingly, a preferred polymerase is one that is capable of extending a nucleic acid primer along a nucleic acid template that is comprised at least predominantly of deoxynucleotides. The formation of the complex comprising the RNA/DNA partial heteroduplex can be carried out by a DNA polymerase which comprises both RNA-dependent DNA polymerase and DNA-dependent DNA polymerase activities (such as Bst DNA polymerase, or a reverse transcriptase). Amplification of an RNA sequence according to methods of the invention involves the use of a DNA polymerase

IDT Exhibit 1020.0901

that is able to displace a nucleic acid strand from the polynucleotide to which the displaced strand is bound, and, generally, the more strand displacement capability the polymerase exhibits (i.e., compared to other polymerases which do not have as much strand displacement capability) is preferable. Preferably, the DNA polymerase has high affinity for binding at the 3'-end of an oligonucleotide hybridized to a nucleic acid strand. Preferably, the DNA polymerase does not possess substantial nicking activity. Generally, the DNA polymerase preferably has little or no 5' to 3' exonuclease activity so as to minimize degradation of primer, or primer extension products. Generally, this exonuclease activity is dependent on factors such as pH, salt concentration, whether the template is double stranded or single stranded, and so forth, all of which are familiar to one skilled in the art. Mutant DNA polymerases in which the 5' to 3' exonuclease activity has been deleted, are known in the art and are suitable for the amplification methods described herein. Mutant DNA polymerases which lack both 5' to 3' nuclease and 3' to 5' nuclease activities have also been described, for example, exo $^{-/-}$-Klenow DNA polymerase. It is preferred that the DNA polymerase displaces primer extension products from the template nucleic acid in at least about 25%, more preferably at least about 50%, even more preferably at least about 75%, and most preferably at least about 90%, of the incidence of contact between the polymerase and the 5' end of the primer extension product. In some embodiments, the use of thermostable DNA polymerases with strand displacement activity is preferred. Such polymerases are known in the art, such as described in U.S. Pat. No. 5,744,312 (and references cited therein). Generally, the DNA polymerase has little to no proofreading activity

[00543] Suitable DNA polymerases for use in the methods and compositions of the invention include those disclosed in U.S. Pat. Nos. 5,648,211 and 5,744,312, which include exo $^-$ Vent (New England Biolabs), exo $^-$ Deep Vent (New England Biolabs), Bst (BioRad), exo $^-$ Pfu (Stratagene), Bca (Panvera), sequencing grade Taq (Promega), exo $^{-/-}$Klenow DNA polymerase, and thermostable DNA polymerases from thermoanaerobacter thermohydrosulfuricus.

[00544] One of the ribonuclease for use in the methods and compositions of the invention is capable of cleaving ribonucleotides in an RNA/DNA hybrid or heteroduplex. Preferably, the ribonuclease cleaves ribonucleotides in an RNA/DNA hybrid regardless of the identity and type of nucleotides adjacent to the ribonucleotide to be cleaved. It is preferred that the ribonuclease cleaves independent of sequence identity. Examples of suitable ribonucleases for the methods and compositions of the invention are well known in the art, including ribonuclease H (RNase H), e.g., Hybridase.

[00545] As is well known in the art, DNA-dependent DNA polymerase activity, RNA-dependent DNA polymerase activity, and the ability to cleave RNA from a RNA/DNA hybrid may be present in different enzymes, or two or more activities may be present in the same enzyme. Accordingly, in some embodiments, the same enzyme comprises RNA-dependent DNA polymerase activity and cleaves RNA from an RNA/DNA hybrid. In some embodiments, the same enzyme comprises DNA-dependent DNA polymerase activity and cleaves RNA from an RNA/DNA hybrid. In some embodiments, the same enzyme comprises DNA-dependent DNA polymerase activity, RNA-dependent DNA polymerase activity and cleaves RNA from an RNA/DNA hybrid. In some embodiments, different enzymes comprise RNA-dependent DNA polymerase activity and DNA-dependent DNA polymerase activity. In some embodiments, different enzymes comprise RNA -

IDT Exhibit 1020.0902

dependent DNA polymerase activity and cleave RNA from an RNA/DNA hybrid. In some embodiments, different enzymes comprise DNA-dependent DNA polymerase activity and cleave RNA from an RNA/DNA hybrid. In other embodiments the RNA targets are degraded by the use of other RNases such as RNase 1, for example, for removing the target RNA following extension of all DNA first primer along the RNA targets by transcription, or a combination of RNase H and other RNases.

Nucleic Acid Target

[00546] The DNA, RNA, or polynucleotide target is generally a polymeric nucleotide, which in the intact natural state can have about 30 to 5,000,000 or more nucleotides and in an isolated state can have about 20 to 50,000 or more nucleotides, usually about 100 to 20,000 nucleotides, more frequently 500 to 10,000 nucleotides. The polynucleotide target to be amplified includes nucleic acids from any source in purified or unpurified form, which can be DNA (dsDNA and ssDNA) or RNA, mitochondrial DNA and RNA, chloroplast DNA and RNA, DNA-RNA hybrids, or mixtures thereof, genes, chromosomes, plasmids, the genomes of biological material such as microorganisms, e.g., bacteria, yeasts, viruses, viroids, molds, fungi, plants, animals, humans, and fragments thereof. Exemplary RNAs include, but are not limited to, mRNAs, tRNAs, snRNAs, rRNAs, retroviruses, small non-coding RNAs, microRNAs, polysomal RNAs, pre-mRNAs, intronic RNA, viral RNA and fragments thereof. Exemplary DNAs include, but are not limited to, genomic DNA, plasmid DNA, phage DNA, nucleolar DNA, mitochondrial DNA, chloroplast DNA, cDNA, synthetic DNA, yeast artificial chromosomal DNA ("YAC"), bacterial artificial chromosome DNA ("BAC"), other extrachromosomal DNA, primer extension products and fragments thereof. Target polynucleotide includes DNA (e.g., genomic DNA, including human genomic DNA, and mammalian genomic DNA (such as mouse, rat,)) and RNA (e.g., mRNA, ribosomal RNA, and total RNA). It should be understood that template RNA includes coding and non-coding RNA. The sequences can be naturally occurring or recombinant nucleic acid targets, including cloned nucleic fragments of interest. The terms "polynucleotide target", "nucleic acid target", "target polynucleotide" and "polynucleotide target" are used interchangeably. The term "DNA target" is interchangeable with the term "target DNA", and the term "RNA target" is interchangeable with the term "target RNA". The target nucleic acid may be a mixture of DNA and RNA targets.

[00547] The target polynucleotide can be only a minor fraction of a complex mixture such as a biological sample and can be obtained from various biological material by procedures well known in the art. Polynucleotides can be obtained from sources containing very small quantities of nucleic acid, such a single cells, small numbers of cells, patient samples, forensic samples, and archeological samples. Obtaining and purifying nucleic acids use standard techniques in the art, including methods designed to isolate one or a very small number of cells, such a cell sorting or laser capture micro-dissection. The methods of the invention are particularly suited for use with genomic DNA (e.g., human and other mammalian genomic DNA), as well as RNA (e.g., total RNA or mRNA samples) or fragments thereof.

[00548] The target polynucleotide(s) can be known or unknown and may contain more than one desired specific nucleic acid sequence of interest, each of which may be the same or different from

IDT Exhibit 1020.0903

each other. If the target polynucleotide is double stranded (e.g., double stranded DNA or a double stranded DNA/RNA hybrid, such as is produced by first strand cDNA synthesis), the target may first be treated to render it single stranded (e.g., by denaturation or by cleavage of the RNA portion of a DNA/RNA hybrid). Denaturation may also be carried out to remove secondary structure present in a single stranded target molecule (e.g., RNA). In some cases, double stranded DNA target polynucleotide may be first cleaved by one or more restriction endonuclease enzymes.

[00549] When the target polynucleotide is DNA, the initial step of the amplification of a target nucleic acid sequence can be rendering the target single stranded. If the target nucleic acid is a double stranded (ds) DNA, the initial step can be target denaturation. The denaturation step may be thermal denaturation or any other method known in the art, such as alkali treatment. If the target nucleic acid is present in a DNA-RNA hybrid, the initial step can be denaturation of the hybrid to obtain a DNA, or removal of the RNA strand using other means known in the art, such as thermal treatment, digestion with an enzyme that cleaves RNA from an RNA/DNA hybrid (such as RNase H) or alkali treatment, to generate single stranded DNA. When the target is RNA, the initial step may be the synthesis of a single stranded cDNA. Techniques for the synthesis of cDNA from RNA are known in the art, and include reverse transcription of RNA strand using a primer that binds to a specific target, such as the poly-A tail of eukaryotic mRNAs or other specific or consensus sequences. In addition, reverse transcription can be primed by a population of degenerate or partially degenerate primers. First strand cDNA can be separated from the complex of RNA and first strand cDNA as described herein.

[00550] RNAs can be from any source in purified or unpurified form, which can be RNA such as total RNA, tRNA, mRNA, rRNA, mitochondrial RNA, chloroplast RNA, DNA-RNA hybrids, or mixtures thereof, from any source and/or species, including human, animals, plants, and microorganisms such as bacteria, yeasts, viruses, viroids, molds, fungi, plants, and fragments thereof. It is understood that the RNA can be coding or noncoding RNA (such as untranslated small RNAs). RNAs can be obtained and purified using standard techniques in the art. Use of a DNA target (including genomic DNA target) can involve initial transcription of the DNA target into RNA form, which can be achieved using methods disclosed in Kurn, U.S. Pat. No. 6,251,639 B1, and by other techniques (such as expression systems) known in the art. Thus, RNA template can be itself generated from a DNA source (such as genomic DNA), using methods known in the art, including Kurn, U.S. Pat. No. 6,251,639. RNA copies of genomic DNA would generally include untranscribed sequences generally not found in mRNA, such as introns, regulatory and control elements, etc. RNA targets may also be generated from cloned genomic DNA sequences that can be subjected to in vitro transcription. Use of a DNA-RNA hybrid can involve denaturation of the hybrid to obtain a single stranded RNA, denaturation followed by transcription of the DNA strand to obtain an RNA, or other methods known in the art such as digestion with an RNase H to generate single stranded DNA.

Reaction Conditions and Detection

[00551] Appropriate reaction media and conditions for carrying out the methods of the invention include those that permit nucleic acid extension, copying, and amplification according to the methods of the invention. Such media and conditions are known to persons of skill in the art, and are

IDT Exhibit 1020.0904

described in various publications, such as U.S. Pat. Nos. 5,554,516; 5,716,785; 5,130,238; 5,194,370; 6,090,591; 5,409,818; 5,554,517; 5,169,766; 5,480,784; 5,399,491; 5,679,512; and PCT Pub. No. WO99/42618. For example, a buffer may be Tris buffer, although other buffers can also be used as long as the buffer components are non-inhibitory to enzyme components of the methods of the invention. The pH is preferably from about 5 to about 11, more preferably from about 6 to about 10, even more preferably from about 7 to about 9, and most preferably from about 7.5 to about 8.5. The reaction medium can also include bivalent metal ions such as Mg.sup.2+, or Mn.sup.2+, at a final concentration of free ions that is within the range of from about 0.01 to about 15 mM, and most preferably from about 1 to 10 mM. The reaction medium can also include other salts, such as KCl or NaCl, that contribute to the total ionic strength of the medium. For example, the range of a salt such as KCl is preferably from about 0 to about 125 mM, more preferably from about 0 to about 100 mM, and most preferably from about 0 to about 75 mM. The reaction medium can further include additives that could affect performance of the amplification reactions, but that are not integral to the activity of the enzyme components of the methods. Such additives include proteins such as BSA, single stranded binding protein (for example, T4 gene 32 protein), and non-ionic detergents such as NP40 or Triton. Reagents, such as DTT, that are capable of maintaining enzyme activities can also be included. Such reagents are known in the art. Where appropriate, an RNase inhibitor (such as RNasin) that does not inhibit the activity of the RNase employed in the method can also be included. Any aspect of the methods of the invention can occur at the same or varying temperatures. In some embodiments, the amplification reactions (particularly, primer extension and transcription; and generally not the step of denaturing) are performed isothermally, which substantially avoids the thermocycling process. The isothermal amplification reaction is carried out at a temperature that permits hybridization of the oligonucleotides (primer) of the invention to the template polynucleotide and that does not substantially inhibit the activity of the enzymes employed. The temperature can be in the range of preferably about 25 °C. to about 85 °C., more preferably about 30 °C. to about 80 °C., and most preferably about 37 °C. to about 75 °C. The temperature for the transcription steps can be lower than the temperature(s) for the preceding steps. The temperature of the transcription steps can be in the range of preferably about 25 °C. to about 85 °C., more preferably about 30 °C. to about 75 °C., and most preferably about 37 °C. to about 70 °C.

[00552] Nucleotide and/or nucleotide analogs, such as deoxyribonucleoside triphosphates, that can be employed for synthesis of the primer extension products in the methods of the invention are provided in the amount of from preferably about 50 to about 2500 μM, more preferably about 100 to about 2000 μM, even more preferably about 200 to about 1700 μM, and most preferably about 250 to about 1500 μM. Nucleotides and/or analogs, such as ribonucleoside triphosphates, that can be employed for synthesis of the RNA transcripts in the methods of the invention are provided in the amount of from preferably about 0.25 to about 6 mM, more preferably about 0.5 to about 5 mM, even more preferably about 0.75 to about 4 mM, and most preferably about 1 to about 3 mM.

[00553] The oligonucleotide components of the reactions of the invention are generally in excess of the number of target nucleic acid sequence to be amplified. They can be provided at about or at least about any of the following: 10, 10^2, 10^4, 10^6, 10^8, 10^{10} times the amount of target nucleic acid.

IDT Exhibit 1020.0905

Primers can be provided at about or at least about any of the following concentrations: 50 nM, 100 nM, 500 nM, 1000 nM, 2500 nM, 5000 nM.

[00554] In the methods of the invention, the steps may be carried out in the order listed or, in some cases, may be carried out in a different order. In some methods a later step depends on the formation of a product from an earlier step, in which case such steps must be carried out in the order listed. One of ordinary skill in the art will understand which steps should be carried out in the order listed, and which steps can be carried out in a different order.

[00555] In some embodiments, the foregoing components are added simultaneously at the initiation of the isothermal amplification process. In another embodiment, components are added in any order prior to or after appropriate timepoints during the amplification process, as required and/or permitted by the amplification reaction. Such timepoints, some of which are noted below, can be readily identified by a person of skill in the art. The enzymes used for nucleic acid amplification according to the methods of the invention can be added to the reaction mixture either prior to a denaturation step, following the denaturation step, or following hybridization of a primer to a polynucleotide template, as determined by their thermal stability and/or other considerations known to the person of skill in the art. The first primer extension product and the second primer extension product synthesis reactions can be performed consecutively, followed by an amplification steps. In these embodiments, the reaction conditions and components may be varied between the different reactions.

[00556] In some embodiments, the amplification reactions can be stopped at various timepoints, and resumed at a later time. Said timepoints can be readily identified by a person of skill in the art. One timepoint is at the end of a first primer extension product synthesis. Another timepoint is at the end of a second primer extension product synthesis. Methods for stopping the reactions are known in the art, including, for example, cooling the reaction mixture to a temperature that inhibits enzyme activity or heating the reaction mixture to a temperature that destroys an enzyme. Methods for resuming the reactions are also known in the art, including, for example, raising the temperature of the reaction mixture to a temperature that permits enzyme activity or replenishing a destroyed (or depleted) enzyme. In some embodiments, one or more of the components of the reactions is replenished prior to, at, or following the resumption of the reactions. Alternatively, the reaction can be allowed to proceed (i.e., from start to finish) without interruption.

[00557] In some embodiments the reaction can be allowed to proceed without purification of intermediate complexes, for example, to remove primer. Products can be purified at various timepoints, which can be readily identified by a person of skill in the art. One timepoint is at the end of first primer extension product synthesis. Another timepoint is at the end of second primer extension synthesis. In some embodiments, the removal of primers and/or target at the end of a defined step by enzymes with appropriate nuclease activities are also useful, for example, cleavage of the RNA portion of free composite tailed primer prior to the isothermal amplification step by treatment with RNase 1.

[00558] The detection of the amplification product can be indicative of the presence of a target sequence. Quantitative analysis is also an aspect of the instant invention. Direct and indirect detection methods (including quantitation) are well known in the art. For example, by comparing

IDT Exhibit 1020.0906

the amount of product amplified from a test sample containing an unknown amount of a polynucleotide containing a target sequence to the product of amplification of a reference sample that has a known quantity of a polynucleotide that contains the target sequence, the amount of target sequence in the test sample can be determined. The amplification methods of the invention can also be extended to analysis of sequence alterations and sequencing of the target nucleic acid. Further, detection could be effected by, for example, examination of translation products from RNA amplification products. The global amplification by the methods of the invention and/or the amplification of selected targets, when present in the sample, are useful for various methods which enable highly parallel nucleic acid interrogations.

Characterization of Nucleic Acids

[00559] The methods of the invention are amenable to quantitative analysis, as in some embodiments, amplification can yield sufficient single stranded polynucleotide (generally, DNA and RNA) products which accurately reflect the representation of the various DNA or RNA sequences (e.g. genomic DNA or mRNA) in the starting material. The amplified products can be analyzed using, for example, probe hybridization techniques known in the art, such as Northern blotting, and hybridizing to probe arrays. In addition, the single stranded polynucleotide products may serve as starting material for other starting material for other analytical and/or quantification methods known in the art, such as real time PCR, quantitative TaqMan, quantitative PCR using molecular beacons, methods described in Kurn, U.S. Pat. No. 6,251,639, etc. Thus, the invention includes those further analytical and/or quantification methods as applied to any of the products of the methods herein.

[00560] In another embodiment, the amplification methods of the invention are utilized to generate multiple copies of single stranded polynucleotide products from RNA or DNA targets that are labeled by the incorporation of labeled nucleotides during DNA polymerization. For example, amplification according to the methods of the invention can be carried out with suitable labeled dNTPs or rNTPs. These labeled nucleotides can be directly attached to a label, or can comprise a moiety which could be attached to a label. The label may be attached covalently or non-covalently to the amplification products. Suitable labels are known in the art, and include, for example, a ligand which is a member of a specific binding pair which can be detected/quantified using a detectable second member of the binding pair. Thus, amplification of total RNA or mRNA according to the methods of the invention in the presence of, for example, Cy3-dUTP or Cy5-dUTP results in the incorporation of these nucleotides into the amplification products.

[00561] The labeled amplified products are suitable for analysis (for example, detection and/or quantification) by contacting them with, for example, microarrays (of any suitable surface, which includes glass, chips, plastic), beads, or particles, that comprise suitable probes such as cDNA and/or oligonucleotide probes. Thus, the invention provides methods to characterize (for example, detect and/or quantify) a DNA or RNA sequence of interest by generating labeled polynucleotide (generally, DNA) products using amplification methods of the invention, and analyzing the labeled products. Analysis of labeled products can be performed by, for example, hybridization of the labeled amplification products to, for example, probes immobilized at, for example, specific locations on a solid or semi-solid substrate, probes immobilized on defined particles, or probes

IDT Exhibit 1020.0907

immobilized on blots (such as a membrane), for example arrays. Other methods of analyzing labeled products are known in the art, such as, for example, by contacting them with a solution comprising probes, followed by extraction of complexes comprising the labeled amplification products and probes from solution. The identity of the probes provides characterization of the sequence identity of the amplified products, and thus by extrapolation the identity of the target DNA or target RNA present in a sample. Hybridization of the labeled products is detectable, and the amount of specific labels that are detected is proportional to the amount of the labeled amplification products of a specific DNA or RNA sequence of interest. This measurement is useful for, for example, measuring the relative amounts of the various RNA species in a sample, which are related to the relative levels of gene expression, as described herein or to detect the presence or absence of defined target DNA or RNA in a sample. The measurement is also useful for measuring the relative amounts of various DNA sequences corresponding, for example, to genetic regions in the sample. The amount of labeled products (as indicated by, for example, detectable signal associated with the label) hybridized at defined locations on an array can be indicative of the detection and/or quantification of the corresponding target DNA or target RNA species in the sample.

Sequencing of the Polynucleotide Products of the Invention

[00562] As described above, the methods can be used to obtain sequence information about a target RNA or target DNA of interest. The sequencing can be carried out on the primer extension products or amplification products produced by the methods herein. In some embodiments the sequencing is performed on the polynucleotides attached to solid surfaces as described herein. In one embodiment sequencing is performed on polynucleotides that are attached to the beads through oligonucleotides attached to the beads which capture amplified product, and are extended to produce a polynucleotide attached to the surface comprising a defined sequence at its 3' end.

[00563] The methods of the invention are useful, for example, for sequencing of a polynucleotide sequence of interest. The sequencing process is carried out as described for the methods described herein.

[00564] Known methods for sequencing include, for example, those described in:, Sanger, F. et al. , Proc. Natl. Acad. Sci. U. S. A. 75,5463-5467 (1977); Maxam, A. M. & Gilbert, W. Proc Natl Acad Sci USA 74,560-564 (1977); Ronaghi, M. et al. , Science 281, 363,365 (1998); Lysov, 1. et al., Dokl Akad Nauk SSSR 303, 1508-1511 (1988); Bains W. & Smith G. C. J. Theor Biol 135,303-307 (1988); Drnanac, R. et al. , Genomics 4,114-128 (1989); Khrapko, K. R. et al. , FEBS Lett 256.118-122 (1989); Pevzner P. A. J Biomol Struct Dyn 7, 63-73 (1989); and Southern, E. M. et al., Genomics 13, 1008-1017 (1992). Pyrophosphate-based sequencing reaction as described, e. g., in US patent 6,274, 320,6258, 568 and 6,210, 891), may also be used. In some cases, the methods above require that the nucleic acid attached to the solid surface be single stranded. In such cases, the unbound strand may be melted away using any number of commonly known methods such as addition of NaOH, application of low ionic (e. g., salt) strength, enzymatic degradation or displacement of the second strand, or heat processing. Where the solid surface comprises a plurality of beads, following this strand removal step, the beads can be pelleted and the supernatant discarded. The beads can then be resuspended in a buffer, and a sequencing primer or other non-amplification

IDT Exhibit 1020.0908

primer can be added. The primer is annealed to the single stranded amplification product. This can be accomplished by using an appropriate annealing buffer and temperature conditions, e. g., as according to standard procedures in the art.

[00565] The methods of the invention are useful, for example, for sequencing of an RNA sequence of interest. The sequencing process can be carried out by processing and amplifying a target RNA containing the sequence of interest by any of the methods described herein. Addition of nucleotides during primer extension can be analyzed using methods known in the art, for example, incorporation of a terminator nucleotide or sequencing by synthesis (e.g. pyrosequencing).

[00566] In embodiments wherein the end product is in the form of DNA primer extension products, in addition to the nucleotides, such as natural deoxyribonucleotide triphosphates (dNTPs), that are used in the amplification methods, appropriate nucleotide triphosphate analogs, which may be labeled or unlabeled, that upon incorporation into a primer extension product effect termination of primer extension, may be added to the reaction mixture. Preferably, the dNTP analogs are added after a sufficient amount of reaction time has elapsed since the initiation of the amplification reaction such that a desired amount of second primer extension product or fragment extension product has been generated. Said amount of the time can be determined empirically by one skilled in the art.

[00567] Suitable dNTP analogs include those commonly used in other sequencing methods and are well known in the art. Examples of dNTP analogs include dideoxyribonucleotides. Examples of rNTP analogs (such as RNA polymerase terminators) include 3'-dNTP. Sasaki et al., Biochemistry (1998) 95:3455-3460. These analogs may be labeled, for example, with fluorochromes or radioisotopes. The labels may also be labels which are suitable for mass spectroscopy. The label may also be a small molecule which is a member of a specific binding pair, and can be detected following binding of the other member of the specific binding pair, such as biotin and streptavidin, respectively, with the last member of the binding pair conjugated to an enzyme that catalyzes the generation of a detectable signal that could be detected by methods such as colorimetry, fluorometry or chemiluminescence. All of the above examples are well known in the art. These are incorporated into the primer extension product or RNA transcripts by the polymerase and serve to stop further extension along a template sequence. The resulting truncated polymerization products are labeled. The accumulated truncated products vary in length, according to the site of incorporation of each of the analogs, which represent the various sequence locations of a complementary nucleotide on the template sequence.

[00568] Analysis of the reaction products for elucidation of sequence information can be carried out using any of various methods known in the art. Such methods include gel electrophoresis and detection of the labeled bands using appropriate scanner, sequencing gel electrophoresis and detection of the radiolabeled band directly by phosphorescence, capillary electrophoresis adapted with a detector specific for the labels used in the reaction, and the like. The label can also be a ligand for a binding protein which is used for detection of the label in combination with an enzyme conjugated to the binding protein, such as biotin-labeled chain terminator and streptavidin conjugated to an enzyme. The label is detected by the enzymatic activity of the enzyme, which generates a detectable signal. As with other sequencing methods known in the art, the sequencing reactions for the various nucleotide types (A, C, G, T or U) are carried out either in a single reaction

IDT Exhibit 1020.0909

vessel, or in separate reaction vessels (each representing one of the various nucleotide types). The choice of method to be used is dependent on practical considerations readily apparent to one skilled in the art, such as the nucleotide tri phosphate analogs and/or label used. Thus, for example, when each of the analogs is differentially labeled, the sequencing reaction can be carried out in a single vessel. The considerations for choice of reagent and reaction conditions for optimal performance of sequencing analysis according to the methods of the invention are similar to those for other previously described sequencing methods. The reagent and reaction conditions should be as described above for the nucleic acid amplification methods of the invention.

The Solid Surface

[00569] In various exemplary embodiments, a solid surface may have a wide variety of forms, including membranes, slides, plates, micromachined chips, microparticles, beads and the like. Solid surfaces may comprise a wide variety of compositions including, but not limited to, glass, plastic, silicon, alkanethiolate derivatized gold, cellulose, low cross linked and high cross linked polystyrene, silica gel, polyamide, and the like, and can have various shapes and features (e.g., wells, indentations, channels, etc.). As used herein, the terms "solid surface" and "solid substrate" are used interchangeably. In some cases these will be referred to as the surface or the support. The surface can be hydrophilic or capable of being rendered hydrophilic and may comprise inorganic powders such as silica, magnesium sulfate, and alumina; natural polymeric materials, particularly cellulosic materials and materials derived from cellulose, such as fiber containing papers, e.g., filter paper, chromatographic paper, etc.; synthetic or modified naturally occurring polymers, such as nitrocellulose, cellulose acetate, poly (vinyl chloride), polyacrylamide, cross linked dextran, agarose, polyacrylate, polyethylene, polypropylene, poly(4-methylbutene), polystyrene, polymethacrylate, poly(ethylene terephthalate), nylon, poly(vinyl butyrate), etc.; either used by themselves or in conjunction with other materials; glass available as Bioglass, ceramics, metals, and the like. Natural or synthetic assemblies such as liposomes, phospholipid vesicles, and cells can also be employed. The surface can have any one of a number of shapes, such as strip, rod, particle, including bead, and the like.

[00570] In some embodiments, the solid surface comprises a bead or plurality of beads. The beads may be of any convenient size and fabricated from any number of known materials. Example of such materials include: inorganics, natural polymers, and synthetic polymers. Specific examples of these materials include: cellulose, cellulose derivatives, acrylic resins, glass, silica gels, polystyrene, gelatin, polyvinyl pyrrolidone, co- polymers of vinyl and acrylamide, polystyrene cross-linked with divinylbenzene or the like (as described, e. g, in Merrifield, Biochemistry 1964,3, 1385-1390), polyacrylamides, latex gels, polystyrene, dextran, rubber, silicon, plastics, nitrocellulose, natural sponges, silica gels, control pore glass, metals, cross-linked dextrans (e. g., Sephadex) agarose gel (SepharoseTM), and other solid phase supports known to those of skill in the art. The beads are generally about 2 to about 100 um in diameter, or about 5 to about 80 pm in diameter, in some cases, about 10 to about 40, um in diameter. In some embodiments the beads can be magnetic. Having magnetic beads can be useful for isolation and purification of the beads comprising nucleic acids described herein. Other methods to separate beads can also be used. For example, the capture beads

IDT Exhibit 1020.0910

may be labeled with a fluorescent moiety which would make the nucleic acid- bead complex fluorescent. The target capture bead complex may be separated, for example, by flow cytometry or fluorescence cell sorter.

Attachment of Oligonucleotides to the Solid Surfaces

[00571] One aspect of the invention involves attaching oligonucleotides to solid surfaces such that the oligonucleotides can hybridize with polynucleotides produced by the methods of the invention. Many methods of attaching oligonucleotides to surfaces are known. In some embodiments, the oligonucleotide is attached covalently to the solid surface. In some embodiments such attached oligonucleotides act to capture nucleic acids such as amplification products, and in some cases such attached oligonucleotides also act as primers. Methods of attaching oligonucleotides and primers to a surface are known in the art (see, e.g., Beier et al., 1999, Nucleic Acids Res. 27(9):1970-1977; Brison et al., 1982, Molecular and Cellular Biology 2:578 587; Cheung et al., 1999, Nat. Genet 21(1 Suppl):15-19; Chrisey et al., 1996, Nucleic Acids Res. 24(15):3031-3039; Cohen et. al., 1997, Nucleic Acids Res. 1997 Feb. 15; 25(4):911-912; Devivar et al., 1999, Bioorg. Med. Chem. Lett. 9(9):1239-1242; Heme et al., 1997. J. Am. Chem. Soc. 119:8916-8920; Kumar et al., 2000, Nucleic Acis Res. 28(14):e71; Lipshutz et al., 1999, Nat. Genet. 21(1 Suppl):20-24; Milner et al., 1997, Nat. Biotechnol. June;15(6):537-541; Morozov et al., 1999, Anal. Chem. 71(15):3110-3117; Proudnikov et al., 1998, Anal Biochem. 259(1):34-41; Rasmussen et al., 1991, Anal Biochem. 198(1):138-142; Rogers et al., 1999, Anal. Biochem. 266(1):23-30; Salo et al., 1999, Bioconjug Chem. 10(5):815-823; Singh-Gasson et al., 1999, Nat. Biotechnol. 17(10):974-978, and Pierce Chemical Company Catalog 1994, pp. 155-200), incorporated herein by reference).

Solid Surfaces in Microreactors or Emulsions

[00572] One aspect of the invention comprises methods and libraries relating to the attachment of nucleic acids on beads in an microreactors such as in an emulsion. As used herein, a microreactor is a small volume of fluid, generally in the volume range of microliters to nanoliters. The small volumes of the microreactors are isolated from one another allowing reactions to occur within each microreactor without significant contamination from or mixing with other microreactors unless desired. Emulsions or microemulsions comprise one approach to microreactors, in which the individual drops each can represent a microreactor. Microreactors such as emulsions can allow for the clonal amplification from a single molecule of a nucleic acid to produce a population of identical nucleic acids within the microreactor. Generally water-in-oil emulsions, also referred to as inverse emulsions are used. These emulsions have small water droplets dispersed in a hydrophobic medium. As used herein, the term oil is used broadly to refer to a hydrophobic fluid in which aqueous droplets can be dispersed. various exemplary embodiments a hydrophobic medium can be can include an oil (e.g., mineral oil, light mineral oil, silicon oil) or a hydrocarbon (e.g., hexane, heptane, octane, nonane, decane, etc.) and the like. Fluorinated hydrocarbons can also be used.

[00573] In some embodiments, solid surfaces can be included within each microreactor. For example, where a bead comprises the solid surface, an emulsion can be produced such that the solid surfaces are included within the emulsion. In some embodiments, an emulsion is produced such that

IDT Exhibit 1020.0911

a plurality of microreactors within the emulsion have one or fewer beads. Related methods are described, for example in WO/2004/069849, and WO/05010145A2. In some aspects of the invention, the nucleic acids are attached to the beads by binding of a ligand attached to the nucleic acid to a receptor bound to the solid surface.

[00574] The emulsion can be formed either in bulk, for example, by mixing, or can be formed by injection, for example using a microfluidic device. For use with the present invention, beads with or without attached nucleic acid template are suspended in a water-in-oil emulsion. Where the amplification reaction requires thermal cycling, such as a PCR reaction, a heat stable emulsion must be used. Where isothermal amplification is employed, the need for thermal stability may be reduced, expanding the types of emulsions that can be used. In some embodiments, a plurality of the microreactors include only one template nucleic acid species and one bead. There may be many droplets that do not contain a template nucleic acid or which do not contain a bead. Likewise there may be droplets that contain more than one copy of a template. The emulsion may be formed according to any suitable method known in the art. One method of creating emulsion is described below but any method for making an emulsion may be used. These methods are known in the art and include adjuvant methods, counter-flow methods, cross-current methods, rotating drum methods, and membrane methods. Furthermore, the size of the microcapsules may be adjusted by varying the flow rate and speed of the components. For example, in dropwise addition, the size of the drops and the total time of delivery may be varied. In some embodiments, the emulsion contains a density of about 3,000 beads encapsulated per microliter.

[00575] Various emulsions that are suitable for biologic reactions are referred to in Griffiths and Tawfik, EMBO, 22, pp. 24-35 (2003); Ghadessy et al. , Proc. Natl. Acad. Sci. USA 98, pp. 4552-4557 (2001); United States Patent No. 6,489, 103 and WO 02/22869, each fully incorporated herein by reference.

[00576] The microreactors should be sufficiently large to encompass sufficient amplification reagents for the degree of amplification required. However, the microreactors should be sufficiently small so that a population of microreactors, each containing a member of a DNA library, can be amplified by conventional laboratory equipment. The use of microreactors as described herein allows amplification of complex mixtures of templates (e. g. , genomic DNA samples or whole cell RNA) without intermixing of sequences, or domination by one or more templates (e. g. , PCR selection bias; see, Wagner et al. , 1994, Suzuki and Giovannoni, 1996; Chandler et al., 1997, Polz and Cavanaugh, 1998).

[00577] In some embodiments the micoreactors are produced such that a plurality of microreactors have one bead and one copy of a nucleic acid template. In some embodiments, limiting dilution can be used to isolate polynucleotides in a manner that is suitable for clonal amplification. In some embodiments the beads have the one copy of a nucleic acid bound to the bead. Producing a plurality of microreactors having only one bead and one nucleic acid template can be performed by diluting the bead and the nucleic acid template or the bead with the bound nucleic acid template to a dilution level at which, on average, each microreactor will contain one or fewer beads and/or one or fewer nucleic acid templates per microreactor. Determining the level of dilution at which this condition occurs can be done by calculation, and by experiment. Dilution to single bead and/or single

IDT Exhibit 1020.0912

molecule is described, for example in WO/2004/069849. For example, a sample comprising a plurality of beads, each comprising, on average one or fewer nucleic acids can be diluted to a concentration such that aliquots of the diluted sample that can be divided into individual microreactors (e.g., wells of a multi-well plate) can be predicted to comprise on average >0 and <1 nucleic acid molecule. Therefore, a percentage of reaction vessels can be predicted on a statistical basis (e.g., Poisson distribution) to comprise an isolated polynucleotide suitable for clonal amplification. Once isolated within the reaction vessels, polynucleotides can be amplified by various methods as described herein to yield clonal amplification.

[00578] Subsequent to amplification, the microreactors can be accessed in order to remove the contents, for example, the beads, for later analysis such as by sequencing. One method of accessing the microreactors comprises breaking the emulsion.

[00579] In one embodiment, following amplification of the nucleic acid template and the attachment of amplification copies to the bead, the emulsion is "broken" (also referred to as "demulsification" in the art). There are many methods of breaking an emulsion (see, e. g., U. S. Patent No. 5,989, 892 and references cited therein) and one of skill in the art would be able to select an appropriate method. In the present invention, one preferred method of breaking the emulsion uses additional oil to cause the emulsion to separate into two phases. The oil phase can then be removed, and a suitable organic solvent (e. g. , hexanes) is added. After mixing, the oil/organic solvent phase is removed. Subsequently, the aqueous layers above the beads are removed. The beads are then washed, for example, with a mixture of an organic solvent and buffer. Suitable organic solvents include alcohols such as methanol, ethanol, and the like. The beads bound to amplification products may then be resuspended in aqueous solution for use, for example, in a sequencing reaction according to known technologies.

[00580] In some embodiments, because the clonal amplicons that are produced are isolated as discrete populations, clonal amplification products can be carried analyzed, for example sequenced, in a parallel manner. Therefore, in some embodiments, at least at least 100, 500, 1000, 10000, 50000, 100000, 300000, 500000, or 1000000 populations of clonal amplification products can be analyzed in parallel. The skilled artisan will appreciate that various methods can be suitable for parallel analysis of clonal amplicons. Generally, such methods can produce a discrete detectable signal that can be associated or linked to individual populations of clonal amplicons.

Determination of Gene Expression Profile

[00581] The amplification methods of the invention can be used for use in determining the levels of expression of multiple genes in a sample since the methods described herein are capable of amplifying multiple target RNAs in the same sample. As described above, amplification products can be detected and quantified by various methods, as described herein and/or known in the art. Since RNA is a product of gene expression, the levels of the various RNA species, such as whole transcriptome or total RNAs, in a sample is indicative of the relative expression levels of the various genes (gene expression profile). Thus, determination of the amount of RNA sequences of interest present in a sample, as determined by quantifying amplification products of the sequences, provides for determination of the gene expression profile of the sample source.

IDT Exhibit 1020.0913

[00582] The methods of the present invention allow for the storage and subsequent analysis of samples, allowing for a sample to be bound to a solid substrate for archiving, then later to be analyzed by the methods described herein to determine a gene expression profile. In some embodiments, the sample can be analyzed multiple times, and stored between analyses.

[00583] Accordingly, the invention provides methods of determining gene expression profile in a sample, said method comprising: amplifying single stranded product from at least one RNA sequence of interest in the sample, using any of the methods described herein; and determining amount of amplification products of each RNA sequence of interest, wherein each said amount is indicative of amount of each RNA sequence of interest in the sample, whereby the expression profile in the sample is determined. Generally, labeled products are generated. In one embodiment, the target RNA is mRNA. It is understood that amount of amplification product may be determined using quantitative and/or qualitative methods. Determining the amount of amplification product includes determining whether amplification product is present or absent. Thus, an expression profile can includes information about presence or absence of one or more RNA sequence of interest. "Absent" or "absence" of product, and "lack of detection of product" as used herein includes insignificant, or de minimus levels. In some cases, the methods further provide for clonal amplification of the target RNA or a subset of the target RNA.

[00584] The methods of expression profiling are useful in a wide variety of molecular diagnostic, and especially in the study of gene expression in essentially any mammalian cell (including a single cell) or cell population. A cell or cell population (e.g. a tissue) may be from, for example, blood, brain, spleen, bone, heart, vascular, lung, kidney, pituitary, endocrine gland, embryonic cells, tumors, or the like. Expression profiling is also useful for comparing a control (normal) sample to a test sample, including test samples collected at different times, including before, after, and/or during development, a treatment, and the like.

Libraries

[00585] In another embodiment, the invention encompasses a library comprising a plurality of nucleic acid molecules, wherein each nucleic acid molecule is separately immobilized to a different bead. In another embodiment, the invention encompasses a library comprising a plurality of nucleic acid molecules, wherein each nucleic acid molecule is separately immobilized to a different bead and wherein each bead comprises over 100, 000 clonal amplification copies of each nucleic acid molecule, wherein the library is contained in a single vessel. As examples, the nucleic acid molecules may be genomic DNA, cDNA, episomal DNA, BAC DNA, or YAC DNA. The genomic DNA may be animal, plant, viral, bacterial, or fungal genomic DNA. Preferably, the genomic DNA is human genomic DNA or human cDNA.

Kits

[00586] One aspect of the invention comprises kits useful for carrying out the methods of the invention.

[00587] In one aspect, the kit comprises (a) a first primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the 3'-DNA portion comprises a random sequence or a specific sequence, and

IDT Exhibit 1020.0914

the 5' RNA portion further comprises sequence (A), (b) a second primer comprising a 5'-ligand. In some embodiments, the kit may further comprise (c), an RNA dependent DNA polymerase, (d) a DNA dependent DNA polymerase with strand displacement activity, (e) RNase H, (f) an amplification chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the sequence of the amplification primer is the substantial the same sequence as the (A) sequence, or a combination thereof. In some cases, the kit may further comprise instructions for the use of said kit.

[00588] In one aspect, the kit comprises (a) a first primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the 3'-DNA portion comprises a random sequence and the 5' RNA portion further comprises sequence (A), (b) a second primer comprising a 5'-ligand, and (c) a chimeric oligonucleotide comprising a 3'-DNA portion substantially comprising sequence (A) and a 5'-RNA sequence (C). In some cases, the kit may further comprise (d) RNase H, (e) an RNA dependent DNA polymerase, (f) a DNA dependent DNA polymerase with strand displacement activity, (g) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the chimeric amplification primer comprises a sequence which is substantially the same as sequence (C), or a combination thereof. In some cases, the kit may further comprise instructions for the use of said kit.

[00589] In some embodiments the second primer further comprises a sequence (B) at or near the 5'-end. In some embodiments the kit further comprises solid support with immobilized ligand binding component on it surface.

[00590] In some embodiments the kit further comprises solid surface with an oligonucleotide attached to the surface by the 5'-end and comprising a sequence (B). In other embodiment the oligonucleotide attached to the solid surface comprises a sequence hybridizable to sequence (A) and the oligonucleotide is attached by the 5'-end.

[00591] In some embodiments the kit comprises (a) a first primer that is a tailed DNA primer comprising a 5'-tail sequence (D), (b) a second primer that is a chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the 5'-end comprises a tail sequence (E), (c) a third primer which is a tailed primer comprising a 3'-sequence that comprises a sequence substantially the same as sequence (D), optionally a 5'-tail sequence (F), and 5'-ligand. In some cases, the kit may further comprise (d) an RNA dependent DNA polymerase, (e) a DNA dependent DNA polymerase with strand displacement activity, (f) RNase H, (g) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-.RNA portion wherein the chimeric amplification primer comprises a sequence which is substantially the same a sequence (E), or a combination thereof. In some cases, the kit may further comprise instructions for the use of said kit.

[00592] In some embodiments, the kit comprises (a) reagents for forming an emulsion and (b) a DNA polymerase with substantial strand displacement activity. Reagents for forming a suitable water in oil emulsion are known and commercially available for example in emPCR kits II and III (454/Roche LifeSciences). Said emulsion forming reagents may include for example decamethylcyclopentasiloxane, polyphenylmethylsiloxane, water and/or buffer. In some cases, the kit may further comprise (c) one or more RNA-DNA chimeric primers, (d) an all DNA primer, (e) a solid surface such as a bead or set of beads, a substantially planar array, a well or wells in a plate, or

IDT Exhibit 1020.0915

an isolated surface or set of isolated surfaces, (f) RNase H, (g) a chimeric oligonucleotide, or a combination thereof. In some cases, the kit may further comprise instructions for the use of said kit.

[00593] In some embodiments the kit comprises (a) a first primer that is a tailed DNA primer comprising a 5'-tail sequence (D), (b) a second primer that is a chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the 5'-end comprises a tail sequence (E), (c) a third primer which is a tailed primer comprising a 3'-sequence that comprises a sequence substantially the same as sequence (D), optionally a 5'-tail sequence (F), and 5'-ligand. In some cases, the kit may further comprise (d) an RNA dependent DNA polymerase, (e) a DNA dependent DNA polymerase with strand displacement activity, (f) RNase H, (g) a chimeric oligonucleotide comprising a 3'-DNA sequence (E) and a 5'-RNA sequence (G), (h) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the chimeric amplification primer comprises a sequence which is substantially the same as sequence (G), or a combination thereof. In some cases, the kit may further comprise instructions for the use of said kit.

[00594] The 3'-end sequence of the first and second primer may comprise a specific sequence or a random sequence. In some embodiments the kits useful for carrying out the methods of the invention may further comprise an inhibitor of the DNA dependent DNA polymerase, such as Actinomycin.

[00595] The components of the kits may comprise the same aspects and embodiments as described above for the components in the description of methods. For example, the ligands and receptors, the primers, the enzymes and the oligonucleotides can be those described herein to carry out the methods of the invention.

Example 1: Clonal Expansion of RNA

Step 1: Synthesis of first primer extension product

[00596] 100ng of an RNA template is provided. The provided RNA template is produced from a biological specimen using a commercially available kit (i.e. Qiagen RNeasy) according to the manufacturer's instructions. A first primer extension reaction mixture is assembled comprising a first primer consisting of a 3' annealing sequence, a portion of which is DNA, and a 5' tail sequence (A), a portion of which is RNA and the following reagents in a total volume of 10 µl:

100ng of RNA template
20 pmol of primer
0.5 µl dNTPs (25 mM)
0.1 µl RNasin
0.1 µl DTT
2 µl 5× AMV reverse transcriptase reaction buffer
DEPC treated water to 10 µl total volume

[00597] The reaction mixture is incubated for 2 min at 75° C., and then cooled to 37° C. 1 µl AMV reverse transcriptase (USB 70041Y, 15U/µl) is added to each reaction and the reaction mixture is further incubated at this temperature for 60 min. The resulting product is a first primer extension

IDT Exhibit 1020.0916

product of DNA with a 5' RNA tail sequence (A) annealed to an RNA template. The reaction mixture is heated to 90°C for 3 min. for enzyme inactivation.

Step 2: Synthesis of Second Primer Extension Product (Formation of (A)/(A') RNA/DNA heteroduplex)

[00598] The first primer extension product is mixed with 10 µl of the second primer extension mixture containing the following:

1 µl 10× Klenow reaction buffer

0.1 µl dNTPs (25 mM)

0.5 µl Klenow Exonuclease free (USB 70057Y 10U/µl)) DNA polymerase

8.4 µl water

[00599] The reaction mixture is incubated for 30 min at 37° C., followed by heating to 75° C. for 5 min to stop the reaction by inactivating the enzymes. The resulting primer extension products comprise a double stranded DNA product with a DNA/RNA heteroduplex at one end of sequence (A')/(A) and a partial duplex with a tailed B sequence at the other end.

Step 3: Cleavage of DNA/RNA Hybrid

[00600] To the second primer extension product reaction mixture is added 0.02 U Hybridase (RNase H) and the reaction mixture is incubated at 50° C. for 60 min and cooled to 4° C. The resulting RNase H digested primer extension product reaction mixture comprises a first primer extension product of DNA without a 5' RNA tail sequence (A) annealed to the second primer extension product of DNA with a 5' tail sequence (B) and a 3' annealing sequence (A').

Step 4: Annealing of Chimeric Oligonucleotide

[00601] Next, a chimeric oligonucleotide comprising a 3' annealing sequence (A), a portion of which is DNA, and a 5' tail sequence (C), a portion of which is RNA is added to the reaction mixture and annealed to the complementary sequence (A'). The annealing step is carried out at 50°C for 10 min. and the reaction is cooled to 4° C.

Step 5: Extension of the Second Primer Extension Product along the Chimeric Oligonucleotide

[00602] To the reaction mixture is added 2.5 units of Exonuclease Free Klenow polymerase and the reaction mixture is incubated for 30 min at 37° C., followed by heating to 75° C. for 5 min to stop the reaction by inactivating the enzymes. The resulting first primer extension product comprises a sequence (A), a 5' tail sequence (C), a portion of which is RNA, and a 3' tail sequence (B') of DNA. The product is a double stranded DNA with a DNA-RNA heteroduplex at one end (C-C').

Step 6: Attachment to a Solid Support

[00603] To the reaction mixture is added a solid support in the form of Polystyrene beads which are derivatized with a short DNA oligonucleotide comprised of sequence B. The reaction mixture is heated to 98°C for 2 minutes and cooled to room temperature to allow for annealing of the first primer extension product to the solid support. The beads are washed in 1x SSC to remove the

IDT Exhibit 1020.0917

second primer extension product. To the beads is added 1x Thermopole buffer (NEB), dNTPs, 1x NEB BSA, Bst DNA polymerase, and water in amounts known in the art to promote DNA polymerase activity. The reaction mixture is incubated in a thermocycler at 25°C for 5 minutes, 50°C for 30 min, 95°C for 5 min, and then cooled to 4°C. The resulting double stranded product, illustrated as the final product in **Figure 17,** is a SPIA substrate suitable for clonal expansion via SPIA amplification. Enzymes, buffers and salts are removed by washing the solid-support-bound SPIA substrate and resuspending in a suitable volume of water.

Step 7: Clonal Expansion

[00604] SPIA amplification is carried out using buffer and enzyme mixtures from NuGEN's WT-Ovation Pico RNA amplification system (NuGen Technologies Inc, San Carlos CA). In addition, the SPIA amplification uses a chimeric amplification primer that has a 5' RNA sequence and a short 3' DNA sequence (approximately 7 base pairs) and is complementary to the sequence (C') of the SPIA substrate produced in step 6. The SPIA amplification is set up as follows:

3µl Amplification primer (50µM stock)
10µl Solid-support-bound SPIA substrate
17µl water
40µl amplification buffer (WT-Ovation Pico System)
20µl Amplification Enzyme Mix (WT-Ovation Pico System)

[00605] The reactions are incubated in a thermocycler at 50°C for 60 min, followed by 95°C for 5 min, and cooled down. The amplification products are generated in close proximity to the remaining unhybridized B-sequence oligonucleotides on the solid support on which the parent SPIA substrate is bound and are therefore also immobilized on the same solid support as they are generated. This process ultimately provides a plurality of solid supports (e.g. beads), each bead with a clonally expanded sequence hybridized therein, such that the hybridized sequence has a known sequence (B') and (C) at each end. The product of this example is useful for such methods as archiving of nucleic acid sequences, reducing the complexity of nucleic acid samples, and providing a set of clonally expanded sequences on beads suitable for use in next generation sequencing platforms such as the SOLiD system by Applied Biosystems.

Example 2: Diagnosis and Prognosis of Cancer

[00606] A suggested course of treatment can be determined by RNA expression analysis of a tumor biopsy. A needle biopsy is performed on a subject to obtain tissue from the suspicious mass for further analysis. The biopsied tissue recovered from the subject is processed to extract and purify total RNA using a commercially available Qiagen RNeasy kit according to the manufacturer's instructions.

[00607] 500pg of total RNA representing at least a portion of the transcriptome of the biopsied material is amplified by the methods of the present invention as described briefly herein. To the RNA in a reaction mixture is added: 100 pmol of a first primer comprising random first primer and a poly dT first primer, a 5' segment and a 3' segment, a portion of the 5' segment comprising RNA,

IDT Exhibit 1020.0918

and a portion of the 3' segment comprising DNA. The 3' DNA segment of the random first primer further comprises an annealing sequence that comprises random hexamers. The 5' RNA segment of the random first primer comprises a tag sequence (A).

[00608] To the reaction mixture 10pmol of a poly-T first primer is also added comprising a 5' segment and a 3' segment, a portion of the 5' segment comprising RNA, and a portion of the 3' segment comprising DNA in the reaction mixture. The 3' DNA segment of this poly-T first primer further comprises an annealing sequence that hybridizes to and is complementary to a portion of the poly A tail of mRNA transcripts present in the total RNA of the reaction mixture, and the 5' RNA segment comprises a tag sequence (A).

[00609] The volume of the reaction mixture is adjusted to 10µl with DEPC treated water, and then the reaction mixture is heated to 75°C for 2 minutes and cooled. First strand synthesis is carried out using the buffer and enzyme reagents provided with the WT-Ovation Pico RNA Amplification kit, and incubation conditions are as described in the User Guide and Quick Protocol: (http://www.nugeninc.com/tasks/sites/nugen/assets/File/user_guides/userguide_wt_ov_pico.pdf and http://www.nugeninc.com/tasks/sites/nugen/assets/File/quick_protocols/qp_wt_ov_pico.pdf.).

[00610] The first primer extension product is mixed with 10µl of the second primer extension mixture containing the following:
second strand cDNA synthesis enzyme mixture and second strand buffer mixture from the NuGEN's WT-Ovation Pico amplification system (as above)
20pmol of a second primer comprising a 5' segment and a 3' segment, a portion of the 5' segment comprising a tag sequence (B), and a portion of the 5' segment comprising an annealing sequence.

[00611] The reaction mixture is incubated under the conditions described in the use Guide for the WT-Ovation Pico RNA Amplification system, followed by heating to 75°C for 5 minutes to stop the reaction by inactivating the enzymes. The resulting primer extension products comprise a partial double stranded DNA product with a DNA/RNA heteroduplex at one end of a sequence (A')/(A). The first primer extension product in the reaction mixture is then digested using exonuclease 1 (0.5 µl at 37°C for 30 min.) followed by inactivation of the enzymes (80°C for 20 minutes). The first and second primer extension products are then purified using Agencourt magnetic beads as per the manufacturer's instructions (User Guide as above) .

[00612] Isothermal linear amplification (SPIA) is then carried out in a reaction mixture containing the above purified reaction products (10 µl), 2 µl chimeric amplification primer (100 µM stock solution), 18 µl water, 40 µl amplification buffer and 20 µl amplification enzyme mixture as provided in the WT-Ovation Pico RNA Amplification System (NuGen Technologies). The amplification is carried out according to the instructions provided for the WT-Ovation Pico RNA Amplification System.

[00613] The amplified product is then analyzed and quantitated by Real Time qPCR with SYBR Green, using an MJ Opticon thermocycler. Amplification reactions are diluted 1:100 in Tris-EDTA and 2 µl of the diluted DNA are analyzed using primer pairs specific for abl, ras, and her2 in three separate reactions.

IDT Exhibit 1020.0919

[00614] The results of the qPCR are analyzed to determine that abl is overexpressed in the cells of the suspicious mass. The results are combined with immunohistochemical and cytological analysis to determine a suggested course of therapy.

Altenatively, a gene expression profile of the sample can be obtained using microarrays such as GeneChip (Affymetrix). The amplified reaction products are subjected to fragmentation and labeling with Biotin using NuGEN's Ovation-F/L reagents and protocol. The fragmented and labeled products are used for hybridization to GeneChip according to the manufacturer protocol. The resulting hybridization data provides a gene expression profile of the sample.

Example 3: Personal Genomics

[00615] An individual is tested by a personal genomics business using the methods of the present invention for single nucleotide polymorphisms (SNPs) within the BRCA1, BRCA2, p53, MPO, NAT1, NAT2, and ras coding regions that are related to increased risks for specific types of cancer.

[00616] The individual supplies a small sample of tissue (i.e. a cheek swab) to the personal genomics business. Genomic DNA from the sample of tissue is isolated using a commercially available kit (i.e. Promega's Wizard ® Genomic DNA Purification Kit), according to the manufacturer's protocol.

[00617] 1 to 10ng of purified genomic DNA is used to clonally amplify the sequences corresponding to the genomic regions with known, cancer related, SNPs of the BRCA1, BRCA2, p53, MPO, NAT1, NAT2, and ras genes on a solid support (i.e. a bead). The target sequences are clonally amplified by isothermal linear amplification using the steps shown in Figure 13 steps I(b), II(b), and III(b); Figure 17 steps IV to X; and Figure 18 of the present application.

[00618] The set of beads comprising the clonally amplified sequences are loaded onto a SOLiD™ Analyzer and the sequences of the regions of interest are determined using the manufacturer's protocol using a primer complementary to the sequence (A) on the 5' end of the amplified product.

[00619] The resulting sequences are used to determine the presence or absence of SNPs related to cancer in the genes of interest. A report is generated that includes the SNPs identified, the impact of the SNPs on lifetime risk of developing specific diseases or conditions, and suggestions for prophylactic or therapeutic interventions.

Example 4: Analysis of Fetal DNA

[00620] A fetal sample is obtained by amniocentesis. DNA is extracted and purified from the sample using a commercially available kit (i.e. Promega's Wizard ® Genomic DNA Purification Kit). The DNA is attached to a solid support according to the method outlined in Figure 19 and clonally amplified according to the method outlined in Figure 18. The set of beads comprising the clonally amplified sequences are loaded onto a Genome Sequencer FLX Titanium Series from Roche/454 Life Sciences and the sequences of the regions of interest are determined using the manufacturer's protocol using a primer complementary to the sequence (A) on the 5' end of the amplified product.

IDT Exhibit 1020.0920

[00621] The resulting sequences are used to determine the presence or absence of SNPs related to prenatal diseases or conditions. A report is generated that includes the SNPs identified, the impact of the SNPs on lifetime risk of developing specific diseases or conditions, and suggestions for prophylactic or therapeutic interventions.

[00622] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

IDT Exhibit 1020.0921

CLAIMS

WHAT IS CLAIMED IS:

1. A method for amplifying a target nucleic acid sequence or its complement on a solid support to form a plurality of amplified products comprising clonally amplifying said target sequence or its complement by:

(i). linear amplification;

(ii). amplification using a single primer;

(iii). amplification from a double-stranded nucleic acid having a 3' overhang;

(iv). amplification using a DNA-RNA chimeric primer; or

(v). (i), (ii), (iii), and (iv);

wherein said target nucleic acid sequence is coupled to said solid support.

2. The method of claim 1 wherein said amplification is isothermal.

3. The method of claim 1 wherein said solid support is a bead or an array.

4. The method of claim 1 wherein said amplification results in at least 50, 100, 500, 1,000, 5,000 copies of the target nucleic acid sequence.

5. The method of claim 1, wherein said solid surface comprises a plurality of primers of substantially identical sequence.

6. The method of claim 1, wherein said target nucleic acid sequence is from a linear template.

7. The method of claim 1, wherein said target nucleic acid sequence is greater than 200 base pairs.

8. A method for clonally amplifying a target nucleic acid sequence or its complement comprising:

delivering the target nucleic acid sequence or its complement into an emulsion; and

performing linear amplification of said target nucleic acid sequence inside said emulsion.

9. The method of claim 8 wherein said amplification results in a plurality of non-multimerized amplified products.

10. The method of claim 1 or 8 wherein the amplified products have fewer than 2 errors for every 100,000 nucleotides.

11. A kit comprising a first primer, and an RNA-DNA chimeric oligonucleotide; the kit optionally further comprising a second primer, a polymerase having substantial strand displacement activity, RNase H, and instructions for use of said reagents.

12. A kit comprising a container containing reagents for forming a water in oil emulsion and a DNA polymerase with substantial strand-displacement activity; the kit optionally further comprising one or more RNA-DNA chimeric primers, an RNA-DNA chimeric oligonucleotide, RNase H, and instructions for use of said reagents.

13. A method comprising:

(a) extending a first primer or a set of first primers with a DNA polymerase comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P) is complementary to

IDT Exhibit 1020.0922

a target nucleic acid and a portion of the 5' RNA segment, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product or set of first primer extension products hybridized to the target nucleic acid;

(b) separating or removing the first primer extension product or the set of first primer extension products from the target nucleic acid;

(c) extending a second primer with a DNA polymerase comprising a 3' DNA segment and a 5' DNA segment, wherein a portion of the 3' DNA segment is complementary to the first primer extension product or set of first primer extension products and a 5' portion, sequence (B), of the of the second primer is not complementary to the first primer extension product or the set of first primer extension products, to produce a double-stranded DNA product or a set of double-stranded DNA products, each product comprising the first primer extension product hybridized to a second primer extension product, whereby the second primer extension product has a DNA sequence (A') that is complementary and hybridized to the RNA sequence (A) of the first primer extension product at its 3' end, thereby forming the double stranded nucleic acid product or products with an RNA/DNA heteroduplex at one end;

(d) cleaving the RNA from the RNA/DNA heteroduplex;

(e) annealing a chimeric oligonucleotide comprising a 5' end wherein the 5' end comprises sequence (C), and the 5' further end comprises RNA, and a 3' end wherein the 3' end comprises the DNA sequence (A), to the double stranded nucleic acid product or products; and

(f) extending the double stranded nucleic acid product or products with a DNA polymerase to produce a sequence (C)-(C') RNA-DNA heteroduplex.

14. The method of claim 13 wherein the second primer further comprises a ligand at a 5' end of the 5' DNA segment.

15. The method of claim 13 further comprising binding the ligand to a solid surface, whereby the first and second primer extension product or products are bound to the solid surface.

16. The method of claims 13 or 15 further comprising an amplification with the steps of:

(g) cleaving the RNA portion of the chimeric oligonucleotide in the DNA-RNA heteroduplex, whereby sequence (C') of step (f) is single stranded.

(h) annealing an amplification primer to the single stranded portion of the second primer extension product complementary to sequence (C), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(i) extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product hybridized to the second primer extension product;

(j) cleaving the RNA from the amplified product hybridized to the second primer extension product in the RNA-DNA heteroduplex; and

IDT Exhibit 1020.0923

(k) repeating steps (h) to (j) to produce multiple copies of amplified product comprising a 5' end and a 3' end, wherein the 5' end comprises sequence (A) and the 3' end comprises sequence (B').

17. The method of claim 13 wherein the 3' portion of the first primer that is complementary to the target nucleic acid comprises a random nucleotide sequence.

18. The method of claim 13 using a set of first primers wherein the members of the set of first primers each comprises a distinct 3' DNA annealing sequence, each specific for a target or region of template nucleic acid.

19. The method of claim 13 wherein the 3' portion of the first primer that is complementary to the target nucleic acid comprises a degenerate or a random sequence for annealing to multiple target sequences.

20. The method of claim 13 wherein the target nucleic acid is RNA.

21. The method of claim 13 wherein the target nucleic acid is DNA.

22. The method of claim 13 wherein, in step (b), the target RNA is cleaved by heat, enzyme treatment, or chemical treatment or enzymes.

23. The method of claim 20 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

24. The method of claim 20 wherein the RNA target is contained within a sample that also comprises DNA, and actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

25. The method of claim 13 wherein the 3' segment of the second primer complementary to a portion of the first primer extension product comprises one or more sequences selected from the group consisting of a random nucleotide sequence, a specific sequence complementary to a specific sequence of the first primer extension product, and a sequence common to multiple first primer extension products.

26. The method of claim 16 wherein the amplification is a clonal amplification.

27. The method of claim 16 wherein the solid surface is a bead or an isolated area on a surface.

28. The method of claim 27 wherein the bead or isolated area is the only bead or isolated area within isolated liquid volume such that the amplified product is contained within such liquid volume.

29. The method of claim 28 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

30. The method of claim 29 wherein the droplet is part of a microemulsion.

31. The method of claim 28 wherein the liquid volume is a well in a plate.

32. The method of claim 28 wherein the bead or isolated area comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B), at their 3' ends, whereby upon the amplification of step (k) multiple copies of amplified product comprising sequence (B') at their 5' end are hybridized to the bead or isolated area.

33. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 32 further comprising extending the oligonucleotide comprising sequence (B) to produce a multiple polynucleotides covalently attached to the bead or isolated area that

IDT Exhibit 1020.0924

are substantially complementary to the amplified product, wherein the multiple polynucleotides comprise sequence (A') near their 5' ends and sequence (B) near their 3' ends.

34. A sequencing method comprising performing the method of claim33, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a sequencing primer, a portion of which is complementary to a portion of sequence (C') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

35. The method of claim 34 wherein the sequencing method comprises cleavable labeled terminators.

36. The method of claim 34 wherein the sequencing method comprises pyrophosphate detection.

37. The method of claim 34 wherein the sequencing method is an isothermal sequencing method.

38. The method of claim 34 wherein the sequencing method comprises cycle sequencing.

39. The method of claim 34 wherein the sequencing method comprises sequencing by ligation.

40. A method for generating an amplified nucleic acid of comprising:

(a) providing a first nucleic acid or a set of first nucleic acids comprising DNA corresponding to a region of a target nucleic acid and further comprising a 5' end and a 3' end, wherein said 5' end of the first nucleic or the set of first nucleic acids comprises a sequence (A) and a sequence (C), wherein said sequence (A) is 3' of sequence (C), and wherein sequence (C) comprises RNA, and wherein said 3' end of the first nucleic acid comprises sequence (B');

(b) hybridizing said first nucleic acid or set of first nucleic acids to a solid support comprising an oligonucleotide comprising a sequence (B) complementary to the sequence (B') of the first nucleic acid;

(c) extending the oligonucleotide to produce a double stranded product comprising a second nucleic acid hybridized to the first nucleic acid, wherein the second nucleic acid comprises a 3' segment complementary to a portion of the first nucleic acid and a 5' segment complementary to a portion of the first nucleic acid, whereby a portion of the 3' segment comprises a sequence (A') and a portion of the 3' segment comprises a sequence (C'), and wherein the sequence (A') is 5' of the sequence (C'), and whereby a portion of the 5' segment comprises a sequence (B).

41. The method of claim 40 wherein the solid support comprises a substantially planar array.

42. The method of claim 40 wherein the solid support comprises a bead.

43. The method of claim 42 wherein the bead comprises a magnetic bead.

44. The method of claim 42 wherein the bead comprises only one copy of the first and second nucleic acid.

45. The method of claim 40, wherein the target nucleic acid is RNA.

46. The method of claim 40, wherein the target nucleic acid is DNA.

47. The method of claim 40, wherein a plurality of first and second nucleic acids are provided corresponding to different sequences in the target nucleic acid.

IDT Exhibit 1020.0925

48. The method of claim 47 wherein the plurality of first and second nucleic acids are bound to one or a plurality of beads or a plurality of isolated areas on a surface.

49. The method of claim 48 wherein the plurality of first and second nucleic acids are bound to a plurality of beads under conditions such that generally, one or fewer first and second nucleic acid is bound to one bead or one isolated area on a surface.

50. The method of claim 49 further comprising amplification by treating the beads or isolated areas with reagents to produce multiple copies of an amplification product complementary to all or a portion of the second nucleic acid products.

51. The method of claim 50 wherein the plurality of beads or isolated areas are contained within a plurality of isolated volumes such that generally one or fewer beads or isolated area is associated with each isolated volume, and whereby the production of multiple copies of amplification product results in multiple copies of substantially one amplification product in each volume.

52. The method of claim 51 wherein the amplification is carried out with a reaction mixture comprising RNase H, an amplification primer comprising a DNA portion and a 5' RNA portion, and a DNA polymerase with strand displacement activity.

53. The method of claim 49 further comprising the step of storing the beads or isolated areas comprising generally one or fewer primer extension products per bead or isolated area.

54. The method of claim 50 further comprising clonally amplifying the primer extension product bound to the beads or isolated areas after storing them, wherein the amplification is carried out with a reaction mixture comprising RNase H, an amplification primer comprising a DNA portion and a 5' RNA portion, and a DNA polymerase with strand displacement activity.

55. A method for attaching a polynucleotide sequence that is representative of a sequence within a nucleic acid target molecule to a solid surface comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P), is complementary to a target nucleic acid and a 5' portion of the of the primer, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid;

(b) separating or removing the first primer extension product from the target nucleic acid;

(c) extending a second primer to produce a double stranded product comprising a second primer extension product hybridized to the first primer extension product, wherein the second primer comprises a 3' segment complementary to a portion of the first primer extension product and a ligand, whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer; and

(d) binding the ligand to a receptor bound to a solid surface whereby the second primer extension product is attached to the solid surface.

56. A method comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid;

IDT Exhibit 1020.0926

(b) cleaving the target RNA from the RNA/DNA hybrid; and

(c) extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer.

57. The method of claim 56 wherein the 3' portion of the primer that is complementary to the target RNA comprises a random nucleotide sequence.

58. The method of claim 56 wherein the 3' portion of the primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

59. The method of claim 56 wherein the 3' portion of the primer that is complementary to the target RNA comprises a specific sequence that is complementary to a multiplicity of targets.

60. The method of claim 56 wherein, in step (b), the target RNA is cleaved by heat, enzyme treatment, or chemical treatment.

61. The method of claim 56 wherein the RNA target is in a sample that also comprises DNA, and wherein actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

62. The method of claim 56 wherein the 3' segment of the second primer complementary to a portion of the first primer extension product comprises a random nucleotide sequence, a specific sequence complementary to a specific sequence of the first primer extension product, or a sequence common to multiple first primer extension products.

63. The method of claim 56 wherein the second primer further comprises a nucleotide sequence (B) that is not complementary to the first primer extension product sequence.

64. The method of claim 56 further comprising step (d) of binding the ligand to a solid surface, whereby the second primer extension product becomes bound to the solid surface.

65. The method of claim 64 wherein the binding of the ligand results in the double stranded product being bound to the solid surface.

66. The method of claim 64 wherein the binding of the ligand results in the single stranded second primer extension product being bound to the solid surface.

67. The method of claims 64 further comprising treating the solid surface with reagents to produce multiple copies of an amplification product that are substantially complementary the second primer extension product.

68. The method of claim 64 wherein the solid surface comprises a substantially planar array.

69. The method of claim 64 wherein the solid surface comprises a bead.

70. The method of claim 69 wherein the bead comprises a magnetic bead.

71. The method of claim 69 wherein the bead comprises only one copy of second primer extension product.

72. The method of claim 71 wherein the second primer extension product is single-stranded.

73. The method of claim 71 wherein the second primer extension product is double-stranded.

IDT Exhibit 1020.0927

74. The method of claim 57 wherein a plurality of second primer extension products are produced corresponding to different sequences in the target RNA.

75. The method of claim 74 wherein the plurality of second primer extension products is bound to one or a plurality of beads or a plurality of isolated areas on a surface.

76. The method of claim 75 wherein the plurality of second primer extension products is bound to a plurality of beads under conditions such that generally, one or fewer second primer extension products is bound to one bead or one isolated area on a surface.

77. The method of claim 75 further comprising treating the beads or isolated areas with reagents to produce multiple copies of an amplification product complementary to all or a portion of the second primer extension products.

78. The method of claim 77 wherein the plurality of beads or isolated areas are contained within a plurality of isolated volumes such that generally one or fewer beads or isolated area is associated with each isolated volume, and whereby the production of multiple copies of amplification product results in multiple copies of substantially one amplification product in each volume.

79. The method of claim 78 wherein the amplification is carried out with a reaction mixture comprising RNase H, an amplification primer comprising a DNA portion and a 5' RNA portion, and a DNA polymerase with strand displacement activity.

80. The method of claim 76 further comprising the step of storing the beads or isolated areas comprising generally one or fewer primer extension products per bead or isolated area.

81. The method of claim 80 further comprising clonally amplifying the primer extension product bound to the beads or isolated areas after storing them.

82. A method for amplifying a nucleic acid representative of a target RNA comprising the method of claim 64 and further comprising the steps of:

(e) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded;

(f) annealing an amplification primer to the single stranded portion of the second primer extension product complementary to sequence (A), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(g) extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product hybridized to the second primer extension product;

(h) cleaving the RNA from the amplified product hybridized to the second primer extension product; and

(i) repeating steps (f) to (h) to produce multiple copies of amplified product.

83. The method of claim 82 wherein the second primer further comprises a sequence (B that is not complementary to the first primer extension product sequence, whereby the amplification product comprises a sequence at its 3' end which is complementary to (B).

84. The method of claim 56 wherein the target RNA comprises messenger RNA.

85. A method comprising:

(a) denaturing a double-stranded target DNA;

IDT Exhibit 1020.0928

(b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end; and

(c) extending a second primer comprising a ligand and a 3' DNA region that comprises a random sequence, wherein the primer is optionally a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product whereby the second primer extension product comprises a ligand.

86. The method of claim 85 wherein step (b) comprises a first incubation at a temperature below about 30 °C, and a second incubation at a temperature above about 40 °C.

87. The method of claim 86 wherein a DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer.

88. The method of claim 85 further comprising step (d) of binding the ligand to a solid surface, whereby the plurality of second primer extension products become bound to the solid surface.

89. The method of claim 88 wherein the binding of the ligand results in the double stranded product being bound to the solid surface.

90. The method of claim 88 wherein the binding of the ligand results in the single stranded second primer extension product being bound to the solid surface.

91. The method of claims 88 further comprising treating the solid surface with reagents to produce multiple copies of amplification products that are substantially complementary to the plurality of second primer extension products.

92. The method of claim 88 wherein the solid surface comprises a substantially planar array.

93. The method of claim 88 wherein the solid surface comprises a plurality of beads or a plurality of isolated areas on a surface.

94. The method of claim 93 wherein the plurality of second primer extension products is bound to a plurality of beads or isolated areas under conditions such that generally, one or fewer copies of a single second primer extension product is bound to one bead or one isolated area.

95. The method of claim 94 further comprising treating the beads or isolated areas with reagents to produce multiple copies of an amplification product substantially complementary to the second primer extension products.

96. The method of claim 95 wherein the plurality of beads or isolated areas are contained within a plurality of isolated volumes such that generally one or fewer beads or isolated area is associated with each isolated volume, and whereby the production of multiple copies of amplification product results in multiple copies of substantially one amplification product in each volume.

97. The method of claim 96 wherein the amplification is carried out with a reaction mixture comprising RNase H, an amplification primer with a DNA portion and a 5' RNA portion, and a DNA polymerase with strand displacement activity.

98. The method of claim 94 further comprising the step of storing the beads or isolated areas comprising generally one or fewer primer extension products.

IDT Exhibit 1020.0929

99. The method of claim 98 further comprising clonally amplifying the primer extension product bound to the beads or isolated areas after storing them.

100. A method for amplifying a nucleic acid representative of a target DNA comprising the method of claim 88, further comprising the steps of:

(e) cleaving the RNA from the first primer extension products such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded;

(f) annealing an amplification primer to the single stranded portion of the second primer extension products complementary to sequence (A), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(g) extending the amplification primer with a DNA polymerase having strand displacement activity to produce an amplified product hybridized to the second primer extension product;

(h) cleaving the RNA from the amplified product hybridized to the second primer extension product; and

(i) repeating steps (f) to (h) to produce multiple copies of amplified product

101. The method of claim 100 wherein the second primer comprises a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, whereby the amplified product comprises a portion complementary to sequence (B) at or near its 3' end.

102. The method of claim 85 wherein the target DNA is genomic DNA

103. .The method of claim 85 wherein the target DNA comprises multiple genomes.

104. A method for archiving and/or clonal expansion of a nucleotide sequence comprising the steps of:

(a) obtaining a plurality of partially double-stranded DNA products comprising a first polynucleotide and a second polynucleotide, wherein the second polynucleotide comprises a sequence (A') at its 3' end and a ligand, wherein the sequence (A') portion of the second polynucleotide is single-stranded, wherein optionally the second polynucleotide comprises a sequence (B) at or near its 5' end;

(b) attaching the partially double stranded DNA products to a plurality of beads or a plurality of isolated areas on a surface by binding the ligands to the bead or isolated area;

(c) annealing an amplification primer to the single stranded portion of the second polynucleotide complementary to sequence (A'), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(d) extending the amplification primer with an enzyme having strand displacement activity to produce a plurality of amplified products hybridized to the second polynucleotide products;

(e) cleaving the RNA from the amplified product hybridized to the second polynucleotide products using RNase H; and

(f) repeating steps (c) to (e) to produce multiple copies of amplified products.

105. The method of claim 104 wherein the ligands of the DNA products are attached to beads or isolated areas, and on average, one DNA product is attached to one or fewer beads or isolated areas.

106. The method of claim 105 further comprising storing the plurality of beads or isolated areas for later analysis.

107. The method of claim 106 wherein the beads or isolated areas are stored after step (b), then later amplified with steps (c) through (f).

IDT Exhibit 1020.0930

108. The method of claim 107 wherein the amplification is a clonal amplification carried out in multiple isolated volumes wherein on average, one isolated volume has one or fewer beads or isolated areas.

109. The method of claim 108 wherein the multiple isolated volumes are droplets in a non-aqueous phase.

110. A method comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid;

(b) cleaving the target RNA from the RNA/DNA hybrid;

(c) extending a second primer, comprising a ligand and a 3' segment complementary to a portion of the first primer extension product, to produce a double stranded product with a DNA/RNA heteroduplex at one end; wherein the double stranded product comprises a second primer extension product hybridized to the first primer extension product, and wherein a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer;

(d) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded;

(e) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C);

(f) optionally extending the oligonucleotide to form an oligonucleotide extension product hybridized to the second primer extension product;

(g) extending the second primer extension product to create a heteroduplex such that the second primer comprises a DNA sequence (C') that is complementary to sequence (C); and

(h) cleaving the RNA from the heteroduplex created in step (g) to produce a single-stranded portion of the second primer extension product corresponding to sequence (C').

111. The method of claim 110 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a random nucleotide sequence.

112. The method of claim 110 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

113. The method of claim 110 wherein the 3' portion of the primer that is complementary to the target RNA comprises a specific sequence that is complementary to a multiplicity of targets.

114. The method of claim 110 wherein the RNA target is contained within a sample that also comprises DNA, and actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

115. The method of claim 110 wherein, in step (b), the target RNA is cleaved by chemical heat, or enzyme treatment.

116. The method of claim 110 wherein the 3' segment of the second primer complementary to a portion of the first primer extension product comprises a random nucleotide sequence, a specific sequence complementary

IDT Exhibit 1020.0931

to a specific sequence of the first primer extension product, or a sequence common to multiple first primer extension products.

117. The method of claim 110 further comprising: (i) binding the ligand on the second primer extension product to a solid surface.

118. The method of claim 117 wherein step (i) of binding the ligand to the solid surface is performed before step (h).

119. The method of claim 117 wherein step (i) of binding the ligand to the solid surface is performed after step (h).

120. A method of amplifying a sequence representative of an sequence within an RNA target molecule comprising the method of claim 117, and further comprising the steps of:

(j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension product complementary to sequence (C');

(k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product;

(l) cleaving the RNA from the amplified product; and

(m) repeating steps (j) to (l) to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence complementary to sequence (A').

121. The method of claim 120 wherein the second primer further comprises a segment (B) that is not complementary to the first primer extension product sequence, whereby the amplified product comprises a sequence (B') at or near its 3' end that is substantially complementary to sequence (B), and a sequence (A) near its 5' end that is complementary to sequence (A').

122. The method of claim 121 wherein the amplification is a clonal amplification.

123. The method of claim 122 wherein the solid surface is a bead or an isolated area on a surface.

124. The method of claim 123 wherein the bead or isolated area is the only bead or isolated area associated with a isolated liquid volume such that the amplified product is contained within such liquid volume.

125. The method of claim 124 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

126. The method of claim 125 wherein the solid surface is a bead and the droplet is part of a microemulsion.

127. The method of claim 124 wherein the liquid volume is a well in a plate.

128. The method of claim 117 wherein the solid surface is a substantially planar substrate.

129. The method of claim 124 wherein the bead or isolated area comprises covalently attached thereto multiple oligonucleotides comprising the sequence (B) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (B') at their 5' end are hybridized to the bead or isolated area.

130. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 129, and further comprising extending the (B) sequences to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and that comprise sequence (A') near their 3' ends.

IDT Exhibit 1020.0932

131. A sequencing method comprising performing the method of claim 130, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (A') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

132. The method of claim 131 wherein the sequencing method comprises cleavable labeled terminators.

133. The method of claim 131 wherein the sequencing method comprises pyrophosphate detection.

134. The method of claim 131 wherein the sequencing method is an isothermal sequencing method.

135. The method of claim 131 wherein the sequencing method comprises cycle sequencing.

136. A method of performing bridge PCR comprising performing the method of claim 121 further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the A and B' sequences on the amplified product in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

137. A method of performing rolling circle amplification comprising performing the method of claim 121 further comprising the steps of:

(n) hybridizing the amplified product to a nucleic acid sequence comprising regions complementary to A and B' sequences in close proximity;

(o) optionally extending the gap with a DNA polymerase enzyme;

(p) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

138. The method of claim 137 wherein the primer is complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product.

139. The method of claim 137 wherein the primer is an oligonucleotide attached to a solid surface.

140. A method of PCR amplification comprising performing the method of claim 121 further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

141. A method of strand displacement amplification (SDA) comprising performing the method of claim 121 wherein sequences (A) and (B') in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

142. A method comprising:

(a) denaturing a double-stranded target DNA;

(b) annealing to the target DNA and extending with and enzyme comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a plurality of first primer extension products, each with sequence (A) at its 5' end;

(c) extending a second primer comprising a ligand and a 3' DNA region that comprises a random sequence, wherein the primer is optionally a tailed primer comprising a nucleic acid sequence (B) that is 5' of the random sequence, to form a plurality of double-stranded products each comprising a first primer extension product and a second primer extension product whereby the second primer extension product comprises a ligand;

IDT Exhibit 1020.0933

(d) cleaving the RNA from the first primer extension products such that a portion of the second primer extension products that are complementary to sequence (A) are single stranded;

(e) annealing to the second primer extension product an oligonucleotide comprising a 3'-DNA segment that is complementary to sequence (A') and a 5'-RNA segment comprising sequence (C);

(f) optionally extending the oligonucleotide to form a plurality of oligonucleotide extension products hybridized to the second primer extension products;

(g) extending the second primer extensions product to create a heteroduplex such that the second primer extension products comprise a DNA sequence (C') that is complementary to sequence (C); and

(h) cleaving the RNA from the heteroduplex created in step (g).

143. The method of claim 142 further comprising: (i) binding the ligand on the second primer extension products to a solid surface.

144. The method of claim 143 wherein step (i) of binding the ligand to the solid surface is performed before step (h).

145. The method of claim 143 wherein step (i) of binding the ligand to the solid surface is performed after step (h).

146. A method of amplifying a sequence representative of an sequence within a DNA target molecule comprising the method of claim 143, and further comprising the steps of:

(j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the second primer extension products complementary to sequence (C');

(k) extending the amplification primer with an enzyme having strand displacement activity to produce a amplified products;

(l) cleaving the RNA from the amplified products; and

(m) repeating steps (j) to (l) to produce multiple copies of amplified products wherein the 5' portion of the amplified product has a sequence complementary to sequence (A').

147. The method of claim 146 wherein the second primer comprises the segment (B) that is not complementary to the first primer extension product sequence, whereby the amplified products comprise a sequence (B') at or near their 3' ends that is substantially complementary to sequence (B), and a sequence (A) near their 5' ends that is complementary to sequence (A').

148. The method of claim 147 wherein the amplification is a clonal amplification.

149. The method of claim 148 wherein the solid surface is a bead or isolated area on a surface.

150. The method of claim 149 wherein the bead or isolated area is the only bead or isolated area within isolated liquid volume such that the amplified product is contained within such liquid volume.

151. The method of claim 150 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

152. The method of claim 151 wherein solid surface is a bead and the droplet is part of a microemulsion.

153. The method of claim 150 wherein the liquid volume is a well in a plate.

154. The method of claim 143 wherein the solid surface is a substantially planar substrate.

IDT Exhibit 1020.0934

155. The method of claim 150 wherein the bead or isolated area comprises covalently attached multiple oligonucleotides comprising the sequence (B) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified products comprising sequence (B') at their 5' end are hybridized to the bead or isolated area.

156. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 155 further comprising extending the (B) sequences to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and that comprise sequence (A') near their 5' ends.

157. A sequencing method comprising performing the method of claim 155, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (A') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

158. The method of claim 157 wherein the sequencing method comprises cleavable labeled terminators.

159. The method of claim 157 wherein the sequencing method comprises pyrophosphate detection.

160. The method of claim 157 wherein the sequencing method is an isothermal sequencing method.

161. The method of claim 157 wherein the sequencing method comprises cycle sequencing.

162. A method of performing bridge PCR comprising performing the method of claim 147 further comprising the steps of exposing the amplified products to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the A and B' sequences on the amplified products in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

163. A method of performing rolling circle amplification comprising performing the method of claim 147 further comprising the steps of:

(n) hybridizing the amplified products to a target nucleic acid comprising regions complementary to A and B' sequences in close proximity;

(o) optionally extending the gap with a polymerase enzyme;

(p) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

164. The method of claim 163 wherein the primer is complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product.

165. The method of claim 163 wherein the primer is an oligonucleotide attached to a solid surface.

166. A method of PCR amplification comprising performing the method of claim 147 further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

167. A method of strand displacement amplification (SDA) comprising performing the method of claim 147 wherein sequences (A) and (B') in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

168. A method comprising:

(a) extending a first primer comprising a 3' portion complementary to a target RNA and a 5' portion, sequence (D), not complementary to the target RNA, to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid;

(b) cleaving the target RNA from the RNA/DNA hybrid;

IDT Exhibit 1020.0935

(c) extending a second primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to the first primer extension product and a 5' portion, sequence (E), of the of the second primer is not complementary to the first primer extension product, to produce a double-stranded DNA product comprising the first primer extension product hybridized to a second primer extension product, whereby the second primer extension product has a sequence (D') that is complementary to sequence (D) at its 3' end;

(d) denaturing the double-stranded DNA product;

(e) annealing to the second primer extension product and extending a third primer comprising, from its 5' end, a ligand, optionally a sequence (F), and a sequence (D), wherein sequence (D) is complementary to sequence (D') on the second primer extension product to produce a double-stranded DNA product comprising the second primer extension product hybridized to a third primer extension product, whereby the third primer extension product comprises a sequence (E') at its 3' end complementary to sequence (E).

169. The method of claim 168 further comprising binding the ligand to a solid surface, whereby the third primer extension product is bound to the solid surface.

170. The method of claim 168 further comprising an amplification with the steps of:

(f) cleaving the RNA portion of the second primer extension product in the DNA-RNA heteroduplex, whereby sequence (E') of the third primer extension product is single stranded.

(g) annealing an oligonucleotide comprising a 3' DNA segment (E) that is complementary to sequence (E') and a 5' RNA segment comprising sequence (G);

(h) extending the third primer extension product to produce a sequence (G') at its 3' end complementary to sequence (G);

(i) cleaving the RNA from the heteroduplex created in step (h) to produce a single-stranded portion of the third primer extension product corresponding to sequence (G').

171. The method of claim 170 further comprising binding the ligand to a solid surface, whereby the third primer extension product comprising sequence (G') is bound to the solid surface.

172. The method of claim 170 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a random nucleotide sequence.

173. The method of claim 170 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

174. The method of claim 170 wherein the RNA target is contained within a sample that also comprises DNA, and actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

175. The method of claim 170 wherein, in step (b), the target RNA is cleaved by heat, enzyme treatment, or chemical treatment or enzymes.

176. The method of claim 170 wherein the 3' segment of the second primer complementary to a portion of the first primer extension product comprises a random nucleotide sequence, a specific sequence complementary to a specific sequence of the first primer extension product, or a sequence common to multiple first primer extension products.

177. A method of amplifying a sequence representative of an sequence within an RNA target molecule comprising the method of claim 171 further comprising the steps of:

IDT Exhibit 1020.0936

(j) annealing an amplification primer, wherein the amplification primer has a DNA portion and a 5' RNA portion, to the single stranded portion of the third primer extension product complementary to sequence (G');

(k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product;

(l) cleaving the RNA from the amplified product; and

(m) repeating steps (j) to (l) to produce multiple copies of amplified product wherein the 5' portion of the amplified product has a sequence (E) complementary to sequence (E') and the 3' end of the amplified product has sequence (D') complementary to sequence (D) and optionally sequence (F') complementary to sequence (F).

178. The method of claim 177 wherein the amplification is a clonal amplification.

179. The method of claim 177 wherein the solid surface is a bead or an isolated area on a surface.

180. The method of claim 179 wherein the bead or isolted area is the only bead or isolated area within isolated liquid volume such that the amplified product is contained within such liquid volume.

181. The method of claim 180 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

182. The method of claim 181 wherein the droplet is part of a microemulsion.

183. The method of claim 180 wherein the liquid volume is a well in a plate.

184. The method of claim 178 wherein the solid surface is a substantially planar substrate.

185. The method of claim 180 wherein the bead or isolated area comprises covalently attached thereto multiple oligonucleotides comprising the sequence (D), and/or sequence (F) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (D') (and/or F') at their 5' end are hybridized to the bead or isolated area.

186. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 185 further comprising extending the oligonucleotide at the (D), and/or (F) sequences to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product comprising sequence (E') near their 5' ends.

187. A sequencing method comprising performing the method of claim 186, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (E') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

188. The method of claim 187 wherein the sequencing method comprises cleavable labeled terminators.

189. The method of claim 187 wherein the sequencing method comprises pyrophosphate detection.

190. The method of claim 187 wherein the sequencing method is an isothermal sequencing method.

191. The method of claim 187 wherein the sequencing method comprises cycle sequencing.

192. A method comprising:

(a) denaturing a double-stranded target DNA;

(b) annealing to the target DNA and extending a first primer comprising a 3' portion comprising a random sequence and a 5' portion, sequence (D), which is not complementary to the target DNA, to form a plurality of first primer extension products, each comprising sequence (D) at its 3' end;

IDT Exhibit 1020.0937

(c) extending a second primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion comprises a random sequence, and a 5' portion, sequence (E), of the of the second primer is not complementary to the first primer extension products, to produce a plurality of double-stranded DNA products comprising a first primer extension product hybridized to a second primer extension product, whereby the second primer extension products have a sequence (D') that is complementary to sequence (D) at their 3' ends;

(d) denaturing the double-stranded DNA products;

(e) annealing to the second primer extension products and extending a third primer comprising, from its 5' end, a ligand, optionally a sequence (F), and a sequence (D), wherein sequence (D) is complementary to sequence (D') on the second primer extension products to produce double-stranded DNA products comprising second primer extension products hybridized to third primer extension products, whereby the third primer extension products comprise a sequence (E') at its 3' end complementary to sequence (E) in a DNA-RNA heteroduplex.

193. The method of claim 192 further comprising binding the ligand to a solid surface, whereby the third primer extension products are bound to the solid surface.

194. The method of claim 192 further comprising the steps of:

(f) cleaving the RNA portion of the second primer extension products in the DNA-RNA heteroduplex, whereby sequence (E') of the third primer extension products is single stranded.

(g) annealing an oligonucleotide comprising a 3' DNA segment (E) that is complementary to sequence (E') and a 5' RNA segment comprising sequence (G);

(h) extending the third primer extension products to produce a sequence (G') at their 3' ends complementary to sequence (G);

(i) cleaving the RNA from the heteroduplex created in step (h) to produce single-stranded portions of the third primer extension products corresponding to sequence (G').

195. The method of claim 194 further comprising binding the ligand to a solid surface, whereby the third primer extension products comprising sequence (G') are bound to the solid surface.

196. A method of amplifying a sequence representative of a sequence within a DNA target molecule comprising following the steps of claim 195 further comprising the steps of:

(j) annealing an amplification primer to the single stranded portion of the third primer extension products complementary to sequence (G'); wherein the amplification primer has a DNA portion and a 5' RNA portion,

(k) extending the amplification primer with an enzyme having strand displacement activity to produce a amplified products;

(l) cleaving the RNA from the amplified products; and

(m) repeating steps (j) to (l) to produce multiple copies of amplified products wherein the 5' portion of the amplified products have a sequence (E) complementary to sequence (E') and the 3' end of the amplified products have sequence (D') complementary to sequence (D) and optionally sequence (F') complementary to sequence (F).

197. The method of claim 196 wherein the amplification is a clonal amplification.

198. The method of claim 196 wherein the solid surface is a plurality of beads or a plurality of isolated areas on a surface.

IDT Exhibit 1020.0938

199. The method of claim 198 wherein each bead or isolated area in the plurality of beads or isolated areas is the only bead or isolated area within isolated liquid volume such that the amplified product corresponding to the sequence on that bead or isolated area is associated with such liquid volume.

200. The method of claim 199 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

201. The method of claim 199 wherein the droplet is part of a microemulsion.

202. The method of claim 199 wherein the liquid volume is a well in a plate.

203. The method of claim 196 wherein the solid surface is a substantially planar substrate.

204. The method of claim 196 wherein the bead or isolated area comprises covalently attached thereto multiple oligonucleotides comprising the sequence (D), and/or sequence (F) at their 3' ends, whereby upon the amplification of step (m) multiple copies of amplified product comprising sequence (D') (and/or sequence (F') at their 5' end are hybridized to the bead or isolated area.

205. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 204 further comprising extending the (D), and/or (F) sequences to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product comprising sequence (E') near their 5' ends.

206. A sequencing method comprising performing the method of claim 205, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (E') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

207. The method of claim 206 wherein the sequencing method comprises cleavable labeled terminators.

208. The method of claim 206 wherein the sequencing method comprises pyrophosphate detection.

209. The method of claim 206 wherein the sequencing method is an isothermal sequencing method.

210. The method of claim 206 wherein the sequencing method comprises cycle sequencing.

211. A library comprising a plurality of double-stranded oligonucleotides each of the oligonucleotides comprising:

(a) a first strand comprising DNA which has, proceeding from its 5' end (i) a ligand, (ii) a specific sequence (B, D, or DF), (iii) a sequence corresponding to or complementary to a sequence within a nucleic acid target, (iv) a specific sequence (A' or E'); and a specific sequence (C' or G'); and

(b) a second strand having from its 5' end, (i) a specific RNA sequence (C or G) complementary to specific sequence C' or G', (ii) a specific sequence (A or E) complementary to sequence (A' or E'), (iii) a sequence complementary to or corresponding to a sequence within a target nucleic acid.

212. The library of claim 211 wherein the ligands are bound to a solid surface.

213. The library of claim 212 wherein the solid surface comprises a plurality of beads.

214. The library of claim 213 wherein each of the plurality of beads comprises a single molecule of double-stranded oligonucleotide.

215. The library of claim 212 wherein the solid surface comprises a plurality of isolated areas on a surface.

216. A kit comprising:

(a) a first primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the 5' RNA portion further comprises sequence (A);

IDT Exhibit 1020.0939

(b) a second primer comprising a 5'-ligand;

(c) an RNA dependent DNA polymerase;

(d) a DNA dependent DNA polymerase with strand displacement activity;

(e) RNase H; and

(f) an amplification chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the sequence of the amplification primer is the substantial the same sequence as the (A) sequence.

217. The kit of claim 216 wherein the 3'-DNA portion of the first primer comprises a random sequence.

218. The kit of claim 216 wherein the second primer is a DNA primer that comprises a random sequence at its 3' end.

219. The kit of claim 216 wherein the 3'-DNA portion of the first primer comprises a random sequence, and the second primer is a DNA primer that comprises a random sequence at its 3' end.

220. A kit comprising:

(a) a first primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the 5' RNA portion further comprises sequence (A);

(b) a second primer comprising a 5'-ligand;

(c) an RNA dependent DNA polymerase;

(d) a DNA dependent DNA polymerase with strand displacement activity;

(e) RNase H;

(f) a chimeric oligonucleotide comprising a 3'-DNA portion substantially comprising sequence (A) and a 5'-RNA sequence (C); and

(g) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-RNA portion, wherein the chimeric amplification primer comprises a sequence which is substantially the same as sequence (C).

221. The kit of claim 220 wherein the 3'-DNA portion of the first primer comprises a random sequence.

222. The kit of claim 220 wherein the second primer is a DNA primer that comprises a random sequence at its 3' end.

223. The kit of claim 220 wherein the 3'-DNA portion of the first primer comprises a random sequence, and the second primer is a DNA primer that comprises a random sequence at its 3' end.

224. The kit of claim 220 wherein the second primer further comprises a sequence (B) at or near the 5'-end.

225. The kit of claim 220 wherein the kit further comprises solid support with immobilized receptor to the ligand on it surface.

226. The kit of claim 220 wherein the kit further comprises solid surface with an oligonucleotide attached to the surface by the 5'-end and comprising a sequence substantially the same as sequence (B).

227. The kit of claim 220 wherein the kit further comprises solid surface with an oligonucleotide attached to the surface by the 5'-end hybridizable to sequence (A).

228. The kit of claim 216 or 220 wherein the kit further comprises an inhibitor of the DNA dependent DNA polymerase,

IDT Exhibit 1020.0940

229. The kit of claim 228 wherein the inhibitor of the DNA dependent DNA polymerase is Actinomycin.

230. A kit comprising:

(a) a first primer that is a tailed DNA primer comprising a 5'-tail sequence (D);

(b) a second primer that is a chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the 5'-end comprises a tail sequence (E);

(c) a third primer which is a tailed primer comprising a 3'-sequence that comprises a sequence substantially the same as sequence (D), optionally a 5'-tail sequence (F), and 5'-ligand;

(d) an RNA dependent DNA polymerase;

(e) a DNA dependent DNA polymerase with strand displacement activity;

(f) RNase H; and

(g) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the chimeric amplification primer comprises a sequence which is substantially the same a sequence (E).

231. The kit of claim 230 wherein the 3'-DNA portion of the first primer comprises a random sequence.

232. The kit of claim 230 wherein the second primer is a DNA primer that comprises a random sequence at its 3' end.

233. The kit of claim 230 wherein the 3'-DNA portion of the first primer comprises a random sequence, and the second primer is a DNA primer that comprises a random sequence at its 3' end.

234. A kit comprising:

(a) a first primer that is a tailed DNA primer comprising a 5'-tail sequence (D);

(b) a second primer that is a chimeric primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the 5'-end comprises a tail sequence (E);

(c) a third primer which is a tailed primer comprising a 3'-sequence that comprises a sequence substantially the same as sequence (D), optionally a 5'-tail sequence (F), and 5'-ligand;

(d) an RNA dependent DNA polymerase;

(e) a DNA dependent DNA polymerase with strand displacement activity;

(f) RNase H;

(g) a chimeric oligonucleotide comprising a 3'-DNA sequence (E) and a 5'-RNA sequence G; and

(h) a chimeric amplification primer comprising a 3'-DNA portion and a 5'-RNA portion wherein the chimeric amplification primer comprises a sequence which is substantially the same as sequence (G).

235. The kit of claim 234 wherein the 3'-DNA portion of the first primer comprises a random sequence.

236. The kit of claim 234 wherein the second primer is a DNA primer that comprises a random sequence at its 3' end.

237. The kit of claim 234 wherein the 3'-DNA portion of the first primer comprises a random sequence, and the second primer is a DNA primer that comprises a random sequence at its 3' end.

238. The kit of claim 230 or 234 wherein the kit further comprises an inhibitor of the DNA dependent DNA polymerase,

IDT Exhibit 1020.0941

239. The kit of claim 238 wherein the inhibitor of the DNA dependent DNA polymerase is Actinomycin.

240. A method for attaching a polynucleotide sequence that is representative of a sequence within a nucleic acid target molecule to a solid surface comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer, sequence (P), is complementary to a target nucleic acid and a 5' portion of the of the primer, sequence (A), is not complementary to the target nucleic acid, to form a first primer extension product hybridized to the target nucleic acid;

(b) separating or removing the first primer extension product from the target nucleic acid;

(c) extending a second primer to produce a double-stranded product comprising a second primer extension product hybridized to the first primer extension product, wherein the second primer comprises a 3' segment complementary to a portion of the first primer extension product and 5' segment non-complementary sequence (B) to the first primer extension product, whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer and a portion of the 5' end of the second primer extension product comprises non-complementary sequence (B);

(d) adding an exonuclease to the double-stranded DNA/RNA hybrid, whereby single stranded 3' nucleotides are removed from the first primer extension product;

(e) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product;

(f) denaturing the first and second primer extension products;

(g) binding the sequence (B') of the first primer extension product to a third primer comprising sequence (B) bound to a solid surface, whereby the first primer extension product is attached to the solid surface; and

(h) extending the sequence (B) of the third primer to produce a double-stranded product comprising a third primer extension product hybridized to the first primer extension product, wherein the 5'end of the third primer comprises a sequence (B) complementary to the sequence (B') of the first primer extension product, whereby a portion of the 3' end of the third primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer.

241. A method comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid;

(b) removing the target RNA from the RNA/DNA hybrid;

(c) extending a second primer, comprising a 3' segment complementary to a portion of the first primer extension product and a 5' segment non-complementary to the first primer extension product comprising sequence (B), to produce a double-stranded DNA product with a DNA/RNA heteroduplex at one end, wherein the double-stranded product comprises a second primer extension product hybridized to the first primer extension product and wherein a portion of the 3' end of the second

IDT Exhibit 1020.0942

primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer;

(d) adding an exonuclease to the double-stranded DNA product, whereby single stranded 3' nucleotides are removed from the 3' region of the first primer extension product that is not hybridized to the second primer extension product;

(e) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product;

(f) denaturing the double-stranded DNA product;

(g) attaching the single-stranded first primer extention product to a solid support by annealing sequence (B') to the solid support comprising an oligonucleotide attached thereto, comprising a sequence (B); and

(h) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the first extension product, wherein the third primer extension product comprises a 3' sequence (A'), whereby a DNA/RNA heteroduplex at one end is generated.

242. The method of claim 241 further comprising:

(i) cleaving the RNA region from the first polynucleotide product hybridized to the third primer extension product using RNase H;

(j) annealing an amplification primer to sequence (A') on the single-stranded portion of the third primer extension product, wherein the amplification primer has a DNA portion and a 5' RNA portion;

(k) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the solid support;

(l) repeating steps (i) to (k) to produce multiple copies of an amplified product wherein the amplified product comprises sequence (B') at its 3' end; and

(m) capturing the amplified product on the solid support wherein the solid support comprises sequence (B).

243. The method of claim 241 wherein the 3' portion of the primer that is complementary to the target RNA comprises a random nucleotide sequence.

244. The method of claim 241 wherein the 3' portion of the primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

245. The method of claim 241 wherein the 3' portion of the primer that is complementary to the target RNA comprises a specific sequence that is complementary to a multiplicity of targets.

246. The method of claim 241 wherein, in step (b), the target RNA is cleaved by heat, enzyme treatment, or chemical treatment.

247. The method of claim 241 wherein the RNA target is in a sample that also comprises DNA, and wherein actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

248. The method of claim 241 wherein the 3' portion of the second primer complementary to a portion of the first primer extension product comprises a random nucleotide sequence, a specific sequence complementary to a specific sequence of the first primer extension product, or a sequence common to multiple first primer extension products.

IDT Exhibit 1020.0943

249. The method of claim 241 wherein the solid support comprises a bead.

250. The method of claim 241 wherein the solid support comprises an isolated area.

251. The method of claim 241 wherein the solid support comprises a plurality of beads or a plurality of isolated areas on a surface.

252. The method of claim 241 wherein the solid support comprises a substantially planar array.

253. The method of claim 249 wherein the bead comprises a magnetic bead.

254. The method of claim 249 wherein the bead comprises only one copy of the first primer extension product.

255. The method of claim 254 wherein the first primer extension product is single-stranded.

256. The method of claim 243 wherein a plurality of first primer extension products are produced corresponding to different sequences in the target RNA.

257. The method of claim 256 wherein the plurality of first primer extension products is bound to a solid support comprising either one or a plurality of beads or a plurality of isolated areas on a surface.

258. The method of claim 257 wherein the plurality of first primer extension products is bound to a plurality of beads under conditions such that generally, one or fewer first primer extension products is bound to one bead or one isolated area on a surface.

259. The method of claim 257 wherein the plurality of beads or isolated areas are contained within a plurality of isolated volumes such that generally one or fewer beads or isolated area is associated with each isolated volume, and whereby the production of multiple copies of amplification product results in multiple copies of substantially one amplification product in each volume.

260. The method of claim 257 further comprising the step of storing the beads or isolated areas comprising generally one or fewer primer extension products per bead or isolated area.

261. The method of claim 260 further comprising clonally amplifying the primer extension product bound to the beads or isolated areas after storing them.

262. The method of claim 241 wherein the target RNA comprises messenger RNA.

263. A method comprising:

(a) denaturing a double-stranded target DNA;

(b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the primer comprises sequence (A), which is not complementary to the target DNA; to form a a plurality of first primer extension product hybridized to the target DNA and comprising sequence (A) at its 5' end;

(c) separating the first primer extension product from the target DNA;

(d) annealing to the first primer extension product and extending a second primer comprising a 3' complementary DNA region that comprises a random sequence, wherein the second primer is a tailed primer comprising a 5' sequence (B), to form a double-stranded product comprising a first primer extension product and a second primer extension product, whereby a double-stranded product with a DNA/RNA heteroduplex at one end is generated;

(e) adding an exonuclease to the double-stranded DNA product, whereby single stranded 3' nucleotides are removed from the 3' region of the first primer extension product that is not hybridized to the second primer extension product;

IDT Exhibit 1020.0944

(f) extending the first primer extension product to produce a sequence (B'), complementary to sequence (B) on the second primer extension product;

(g) denaturing the double-stranded DNA product;

(h) attaching the single-stranded first primer extention product to a solid support by annealing sequence (B') to the solid support comprising an oligonucleotide attached thereto, comprising a sequence (B), whereby a plurality of first primer extension products become bound to the solid surface; and

(i) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the first primer extension product, comprising a 3' sequence (A'), whereby a DNA/RNA heteroduplex at one end is generated.

264. The method of claim 263 further comprising:

(j) cleaving the RNA from the first polynucleotide product hybridized to the amplified product using RNase H;

(k) annealing an amplification primer to the single-stranded portion of the amplified product complementary to sequence (A'), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(l) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the bead or isolated area;

(m) repeating steps (j) to (l) to produce multiple copies of an amplified product wherein the amplified product comprises sequence (B') at its 3' end; and

(n) capturing the amplified product on the solid support comprising sequence (B).

265. The method of claim 263 wherein step (b) comprises a first incubation at a temperature below about 30 °C, and a second incubation at a temperature above about 40 °C.

266. The method of claim 263 wherein a DNA polymerase which is active at temperatures above about 45 °C is used to extend the first primer.

267. The method of claim 263 wherein the solid support comprises a bead.

268. The method of claim 263 wherein the solid support comprises an isolated area.

269. The method of claim 263 wherein the solid support comprises a plurality of beads or a plurality of isolated areas on a surface.

270. The method of claim 263 wherein the solid support comprises a substantially planar array.

271. The method of claim 267 wherein the bead comprises a magnetic bead.

272. The method of claim 269 wherein the plurality of first primer extension products is bound to a plurality of beads or isolated areas under conditions such that generally, one or fewer copies of a single first primer extension product is bound to one bead or one isolated area.

273. The method of claim 269 wherein the plurality of beads or isolated areas are contained within a plurality of isolated volumes such that generally one or fewer beads or isolated area is associated with each isolated volume, and whereby the production of multiple copies of amplification product results in multiple copies of substantially one amplification product in each volume.

IDT Exhibit 1020.0945

274. The method of claim 272 further comprising the step of storing the beads or isolated areas comprising generally one or fewer primer extension products.

275. The method of claim 274 further comprising clonally amplifying the primer extension product bound to the beads or isolated areas after storing them.

276. The method of claim 263 wherein the target DNA is genomic DNA.

277. The method of claim 263 wherein the target DNA comprises multiple genomes.

278. A method comprising:

(a) extending a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer is complementary to a target RNA and a 5' portion, sequence (A), of the of the primer is not complementary to the target RNA; to form a first primer extension product hybridized to the target RNA, forming an RNA/DNA hybrid;

(b) removing the target RNA from the RNA/DNA hybrid;

(c) extending a second primer, comprising a 3' segment complementary to a portion of the first primer extension product and a 5' segment non-complementary to the first primer extension product comprising sequence (B), to produce a double-stranded product with a DNA/RNA heteroduplex at one end; wherein the double-stranded product comprises a second primer extension product hybridized to the first primer extension product, and whereby a portion of the 3' end of the second primer extension product comprises a sequence (A') that is complementary to the sequence (A) of the of the first primer;

(d) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single-stranded;

(e) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA sequence (A) that is complementary to sequence (A') and a 5'-RNA segment comprising sequence (C) that is non-complementary to the second primer extension product;

(f) extending the oligonucleotide at the 3' segment to form an oligonucleotide extension product hybridized to the second primer extension product;

(g) denaturing the double-stranded DNA product;

(h) attaching the single-stranded first primer extension product to a solid support by annealing sequence (B') to the solid support comprising a sequence (B); and

(i) extending sequence (B) on the solid support to produce a third primer extension product, comprising a 3' sequence (A') and (C'), whereby a DNA/RNA heteroduplex at one end is generated.

279. The method of claim 278 comprising:

(j) cleaving the RNA from the heteroduplex polynucleotide product hybridized to the amplified product using RNase H to produce a single-stranded portion of the third primer extension product corresponding to sequence (C');

(k) annealing an amplification primer to the single-stranded portion of the third primer extension product complementary to sequence (C'), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(l) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the third primer extension product on the solid support;

IDT Exhibit 1020.0946

(m) repeating steps (j) to (l) to produce multiple copies of the amplified product comprising sequences (A) and (B'); and

(n) capturing the amplified product on the solid support wherein the solid support comprises sequence (B).

280. The method of claim 278 further comprising extending sequence (B), whereby multiple copies are bound to solid support through sequence (B) having sequence (A') at 3' end.

281. The method of claim 278 further comprising sequencing by synthesis using a sequence (A) complementary to (A') as the priming sequence.

282. The method of claim 278 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a random nucleotide sequence.

283. The method of claim 278 wherein the 3' portion of the first primer that is complementary to the target RNA comprises a sequence that is complementary to polyadenosine (poly-A).

284. The method of claim 278 wherein the 3' portion of the primer that is complementary to the target RNA comprises a specific sequence that is complementary to a multiplicity of targets.

285. The method of claim 278 wherein the RNA target is contained within a sample that also comprises DNA, and actinomycin is added prior to step (a) to selectively inhibit the production of extension product complementary to the DNA during step (a).

286. The method of claim 278 wherein, in step (b), the target RNA is cleaved by chemical heat, or enzyme treatment.

287. The method of claim 278 wherein the 3' segment of the second primer complementary to a portion of the first primer extension product comprises a random nucleotide sequence, a specific sequence complementary to a specific sequence of the first primer extension product, or a sequence common to multiple first primer extension products.

288. The method of claim 278 wherein the amplification is a clonal amplification.

289. The method of claim 278 wherein the solid support is a bead.

290. The method of claim 278 wherein the solid support is an isolated area on a surface.

291. The method of claim 278 wherein the bead or isolated area is the only bead or isolated area associated with an isolated liquid volume such that the amplified product is contained within such liquid volume.

292. The method of claim 291 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

293. The method of claim 291 wherein the solid surface is a bead and the droplet is part of a microemulsion.

IDT Exhibit 1020.0947

294. The method of claim 291 wherein the liquid volume is a well in a plate.

295. The method of claim 278 wherein the solid support is a substantially planar substrate.

296. A method of producing a solid support with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 278, and further comprising extending the (B) sequences to produce multiple polynucleotides covalently attached to the solid support that are substantially complementary to the amplified product and that comprise sequence (A') near their 3' ends.

297. A sequencing method comprising performing the method of claim 296, further comprising the steps of removing the third primer extension product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (A) to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

298. The method of claim 297 wherein the sequencing method comprises cleavable labeled terminators.

299. The method of claim 297 wherein the sequencing method comprises pyrophosphate detection.

300. The method of claim 297 wherein the sequencing method is an isothermal sequencing method.

301. The method of claim 297 wherein the sequencing method comprises cycle sequencing.

302. A method of performing bridge PCR comprising performing the method of claim 278 further comprising the steps of exposing the amplified product to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the A and B' sequences on the amplified product in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

303. A method of performing rolling circle amplification comprising performing the method of claim 278 further comprising the steps of:

(o) hybridizing the amplified product to a nucleic acid sequence comprising regions complementary to A and B' sequences in close proximity;

(p) optionally extending the gap with a DNA polymerase enzyme;

(q) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

304. The method of claim 303 wherein the primer is complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product.

305. The method of claim 303 wherein the primer is an oligonucleotide attached to a solid surface.

306. A method of PCR amplification comprising performing the method of claim 278 further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

IDT Exhibit 1020.0948

307. A method of strand displacement amplification (SDA) comprising performing the method of claim 278 wherein sequences (A) and (B') in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

308. A method comprising:

(a) denaturing a double-stranded target DNA;

(b) annealing to the target DNA and extending with a DNA polymerase comprising strand displacement activity, a first primer comprising a DNA segment and a 5' RNA segment, wherein a 3' portion of the primer comprises a random sequence, and a 5' portion of the of the primer comprises sequence (A), which is not complementary to the target DNA; to form a first primer extension product hybridized to the target DNA and comprising sequence (A) at its 5' end;

(c) separating the first primer extension product from the target DNA;

(d) annealing to the first primer extension product and extending a second primer comprising a 3' complementary DNA region that comprises a random sequence, wherein the second primer is a tailed primer comprising a 5' sequence (B), to form a double-stranded product comprising a first primer extension product and a second primer extension product, whereby a double-stranded product with a DNA/RNA heteroduplex at one end is generated.

(e) cleaving the RNA in the heteroduplex from the first primer extension product such that a portion of the second primer extension product that is complementary to sequence (A) is single stranded;

(f) annealing to the second primer extension product an oligonucleotide comprising a 3'- DNA segment that is complementary to sequence (A') and a 5' RNA segment comprising sequence (C);

(g) extending the oligonucleotide along the second primer extension product to form an oligonucleotide extension product comprising a sequence (B'), complementary to sequence (B) on the second primer extension product;

(h) denaturing the double-stranded DNA product;

(i) attaching the single-stranded first primer extension product to solid support by annealing sequence (B') to the bead or isolated area comprising a sequence (B); and

(j) extending sequence (B) on the solid support to produce a third primer extension product, hybridized to the oligonucleotide extension product, comprising a 3' sequence (A') and (C'), whereby a DNA/RNA heteroduplex at one end is generated.

309. The method of claim 308 further comprising:

(k) cleaving the RNA from the heteroduplex polynucleotide product hybridized to the amplified product using RNase H to produce a single-stranded portion of the second primer extension product corresponding to sequence (C');

(l) annealing an amplification primer to the single-stranded portion of the amplified product complementary to sequence (C'), wherein the amplification primer has a DNA portion and a 5' RNA portion;

(m) extending the amplification primer with an enzyme having strand displacement activity to produce an amplified product hybridized to the amplified product on the bead or isolated area; and

(n) repeating steps (k) to (m) to produce multiple copies of the second polynucleotide product comprising sequences (A) and (B').

IDT Exhibit 1020.0949

310. The method of claim 308 wherein the amplification is a clonal amplification.

311. The method of claim 308 wherein the solid support is a bead or isolated area on a surface.

312. The method of claim 311 wherein the bead or isolated area is the only bead or isolated area within isolated liquid volume such that the amplified product is contained within such liquid volume.

313. The method of claim 312 wherein the liquid volume is an aqueous droplet within a non-aqueous fluid.

314. The method of claim 313 wherein solid surface is a bead and the droplet is part of a microemulsion.

315. The method of claim 312 wherein the liquid volume is a well in a plate.

316. The method of claim 308 wherein the solid support is a substantially planar substrate.

317. The method of claim 309 wherein the bead or isolated area comprises covalently attached multiple oligonucleotides comprising the sequence (B) at their 3' ends, whereby upon the amplification of step (m), multiple copies of amplified products comprising sequence (B') at their 5' end are hybridized to the bead or isolated area.

318. A method of producing a bead or isolated area with multiple copies of a nucleotide sequence covalently attached thereto by performing the method of claim 317 further comprising extending the (B) sequences to produce a multiple polynucleotides covalently attached to the bead or isolated area that are substantially complementary to the amplified product and that comprise sequence (A') near their 5' ends.

319. A sequencing method comprising performing the method of claim 317, further comprising the steps of removing the amplified product to render the covalently attached polynucleotides single-stranded, and extending a primer to sequence (A') to produce detectable oligonucleotide fragments characteristic of the sequence of the polynucleotide bound to the bead or isolated area.

320. The method of claim 319 wherein the sequencing method comprises cleavable labeled terminators.

321. The method of claim 319 wherein the sequencing method comprises pyrophosphate detection.

322. The method of claim 319 wherein the sequencing method is an isothermal sequencing method.

323. The method of claim 319 wherein the sequencing method comprises cycle sequencing.

324. A method of performing bridge PCR comprising performing the method of claim 308 further comprising the steps of exposing the amplified products to a solid substrate comprising oligonucleotide sequences attached thereto complementary to the A and B' sequences on the amplified products in the presence of components necessary for polymerase chain reaction, and thermal cycling the system to perform bridge PCR amplification.

325. A method of performing rolling circle amplification comprising performing the method of claim 308 further comprising the steps of:

(o) hybridizing the amplified products to a target nucleic acid comprising regions complementary to A and B' sequences in close proximity;

IDT Exhibit 1020.0950

(p) optionally extending the gap with a polymerase enzyme;

(q) ligating to form a circular nucleic acid comprising the amplified product, and performing rolling circle amplification by extending a primer that is complementary to a sequence in the circular nucleic acid.

326. The method of claim 325 wherein the primer is complementary to sequence (A), sequence (B'), or a sequence that was between sequences (A) and (B') in the amplified product.

327. The method of claim 325 wherein the primer is an oligonucleotide attached to a solid surface.

328. A method of PCR amplification comprising performing the method of claim 308 further comprising the steps of amplifying the amplified product using primers complementary to sequences (A) and (B), or using primers complementary to sequences (A') and (B').

329. A method of strand displacement amplification (SDA) comprising performing the method of claim 308 wherein sequences (A) and (B') in the amplified product are designed to be cleaved by a restriction enzyme, and performing strand displacement amplification on the amplified product.

IDT Exhibit 1020.0951

Figure 1

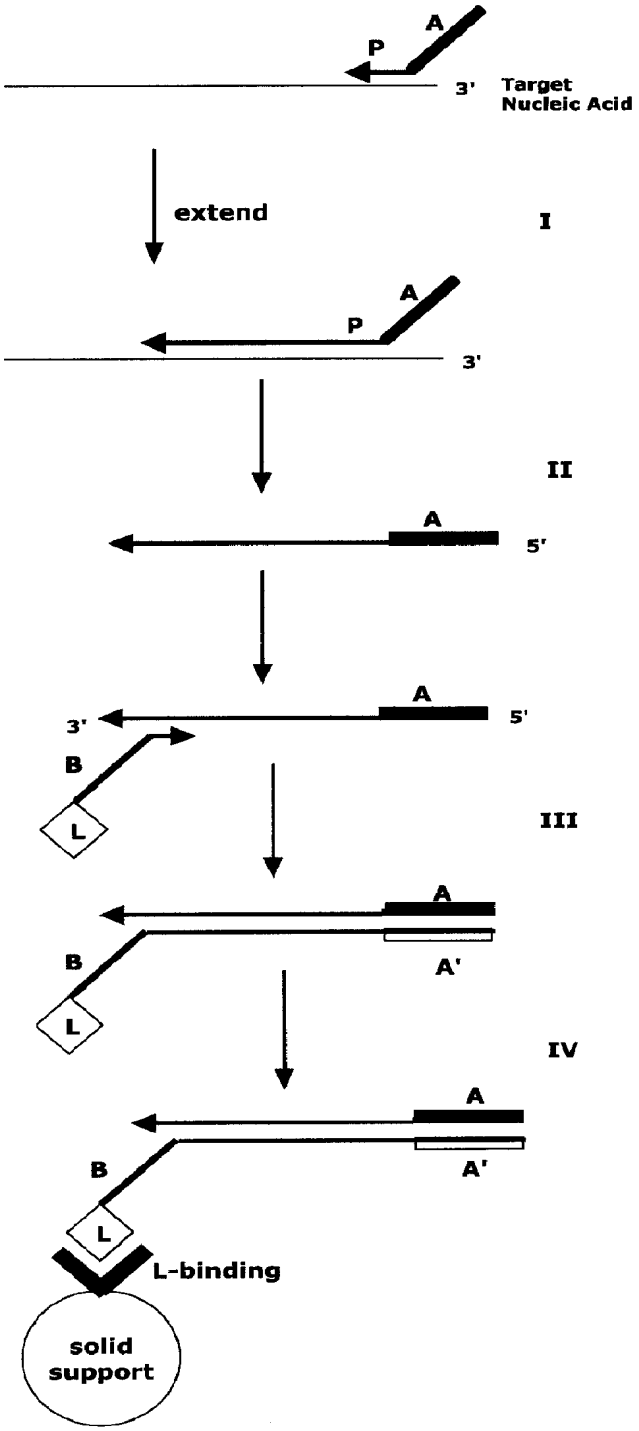

A
P
3'
Target
Nucleic Acid
extend
I
A
P
3'
II
A
5'
A
3'
5'
B
L
III
A
B
A'
L
IV
A
B
A'
L
L-binding
solid
support

Figure 2

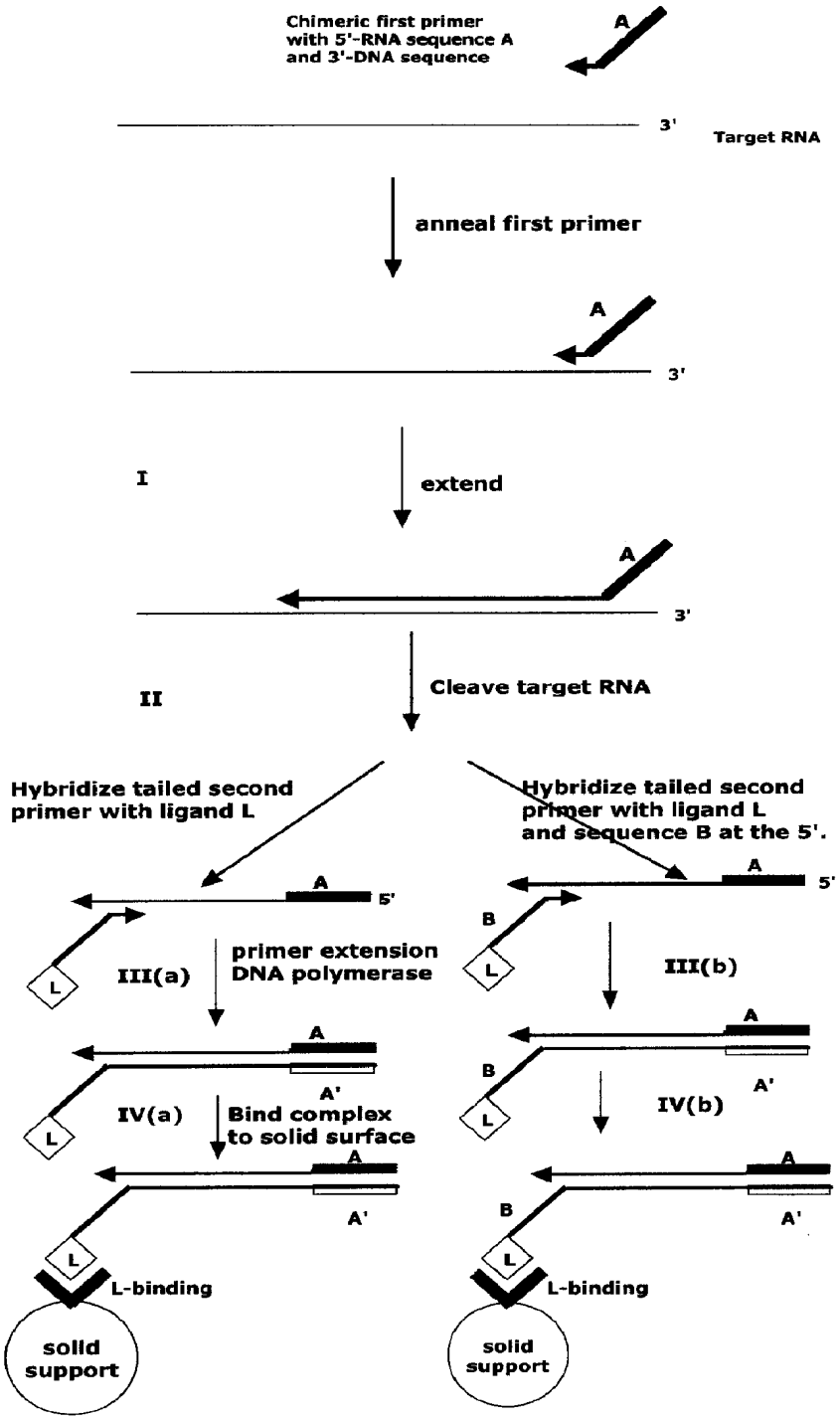

Chimeric first primer with 5'-RNA sequence A and 3'-DNA sequence
A
3'
Target RNA
anneal first primer
A
3'
I
extend
A
3'
II
Cleave target RNA
Hybridize tailed second primer with ligand L
Hybridize tailed second primer with ligand L and sequence B at the 5'.
A
5'
B
L
III(a)
primer extension DNA polymerase
III(b)
A
A'
IV(a)
Bind complex to solid surface
IV(b)
L-binding
solid support

Figure 3

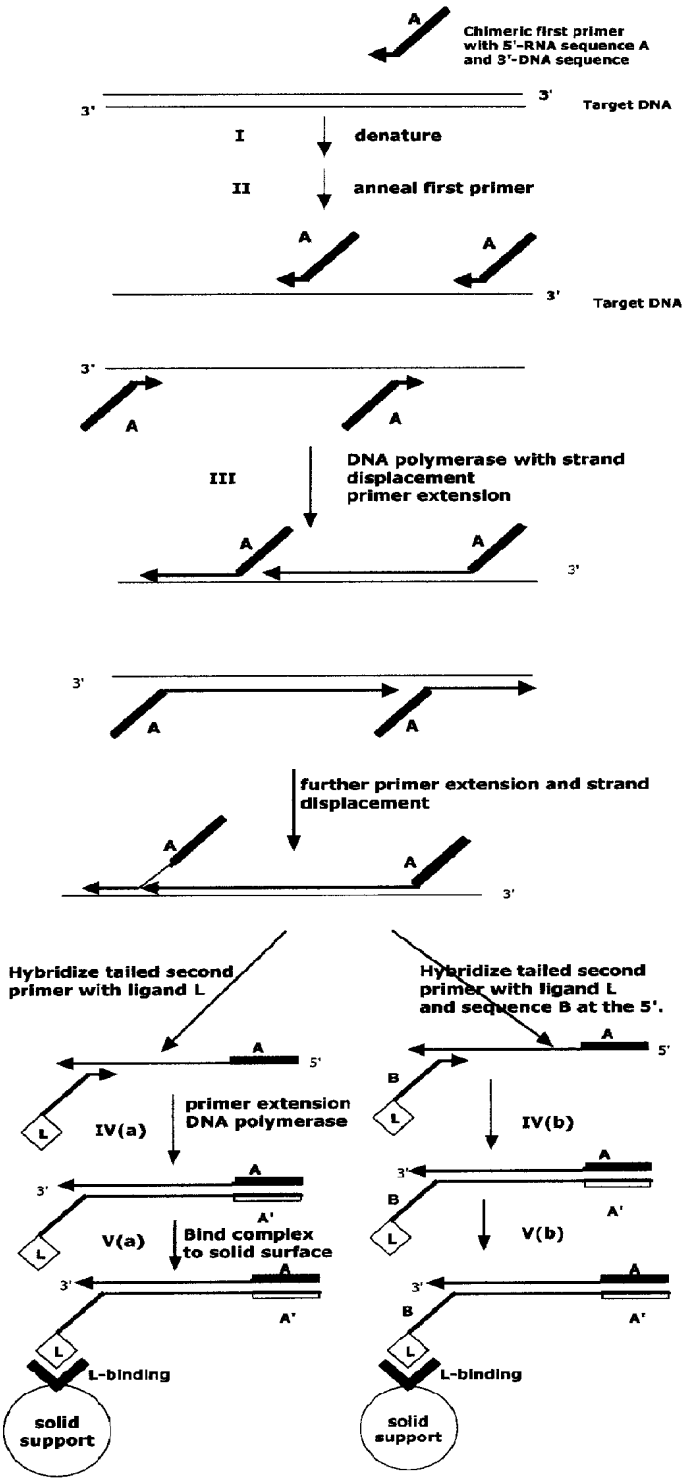

A
Chimeric first primer with 5'-RNA sequence A and 3'-DNA sequence
3'
Target DNA
I
denature
II
anneal first primer
A
A
3'
Target DNA
3'
A
A
III
DNA polymerase with strand displacement primer extension
A
A
3'
3'
A
A
further primer extension and strand displacement
A
A
3'
Hybridize tailed second primer with ligand L
A
5'
L
IV(a)
primer extension DNA polymerase
3'
A
A'
L
V(a)
Bind complex to solid surface
3'
A
A'
L
L-binding
solid support
Hybridize tailed second primer with ligand L and sequence B at the 5'.
A
5'
B
L
IV(b)
3'
A
A'
B
L
V(b)
3'
A
A'
B
L
L-binding
solid support

Figure 4

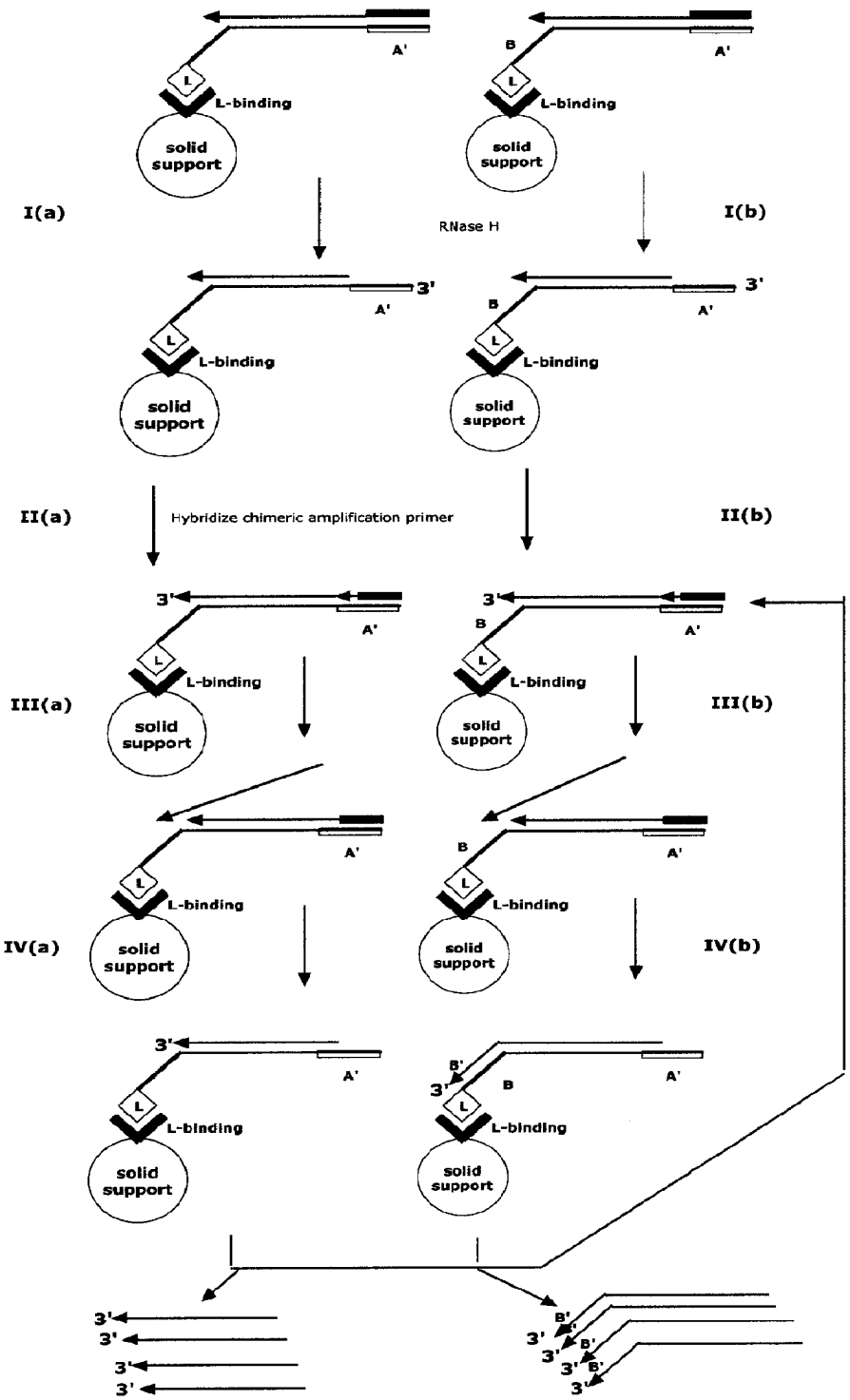

I(a)
I(b)
RNase H
II(a)
Hybridize chimeric amplification primer
II(b)
III(a)
III(b)
IV(a)
IV(b)
L-binding
solid support
A'
B
B'
3'
L

Figure 5A

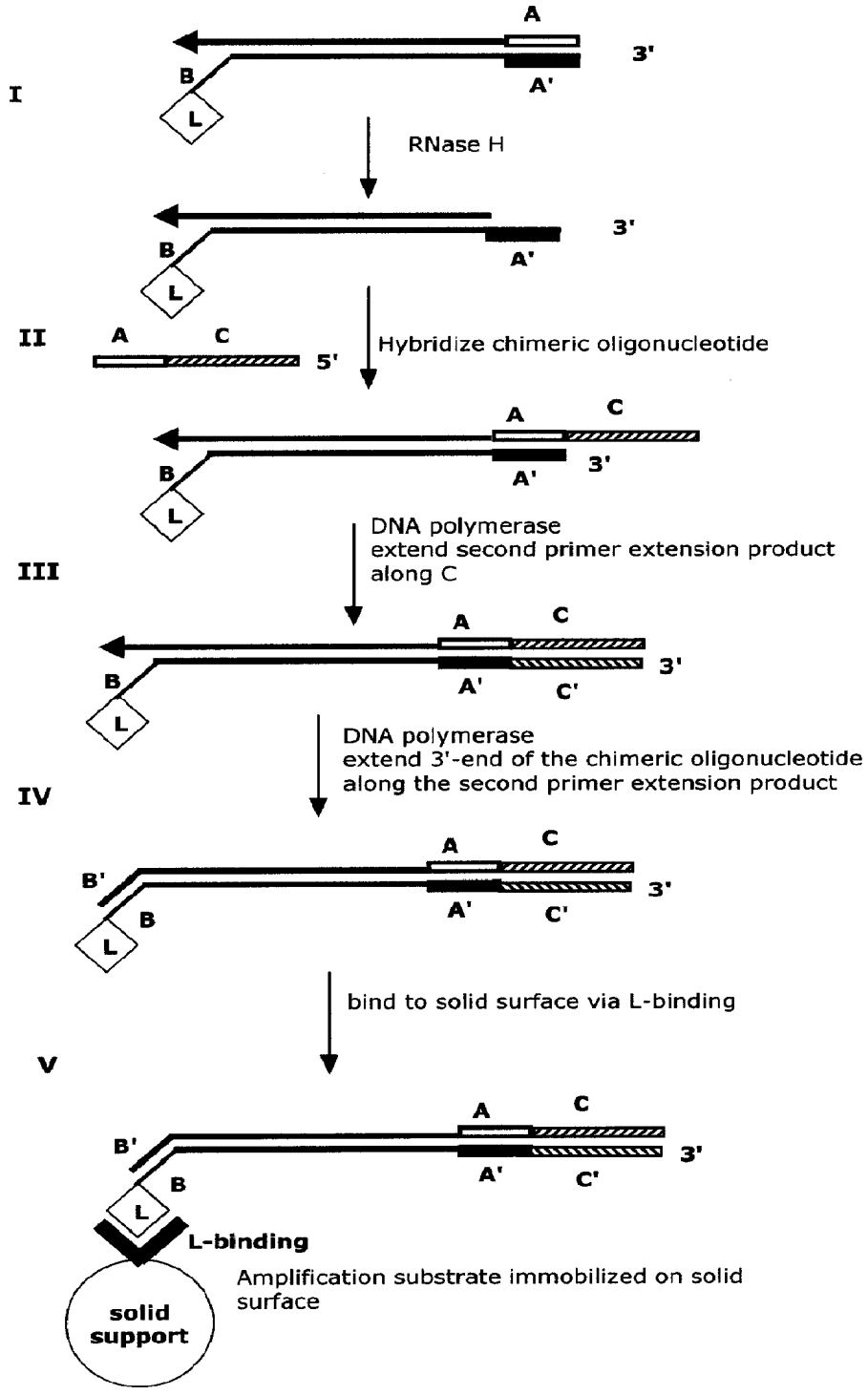

I
A
3'
B
A'
L
RNase H
II
A
C
5'
Hybridize chimeric oligonucleotide
III
DNA polymerase
extend second primer extension product
along C
C'
IV
DNA polymerase
extend 3'-end of the chimeric oligonucleotide
along the second primer extension product
B'
V
bind to solid surface via L-binding
L-binding
Amplification substrate immobilized on solid
surface
solid
support

Figure 5B

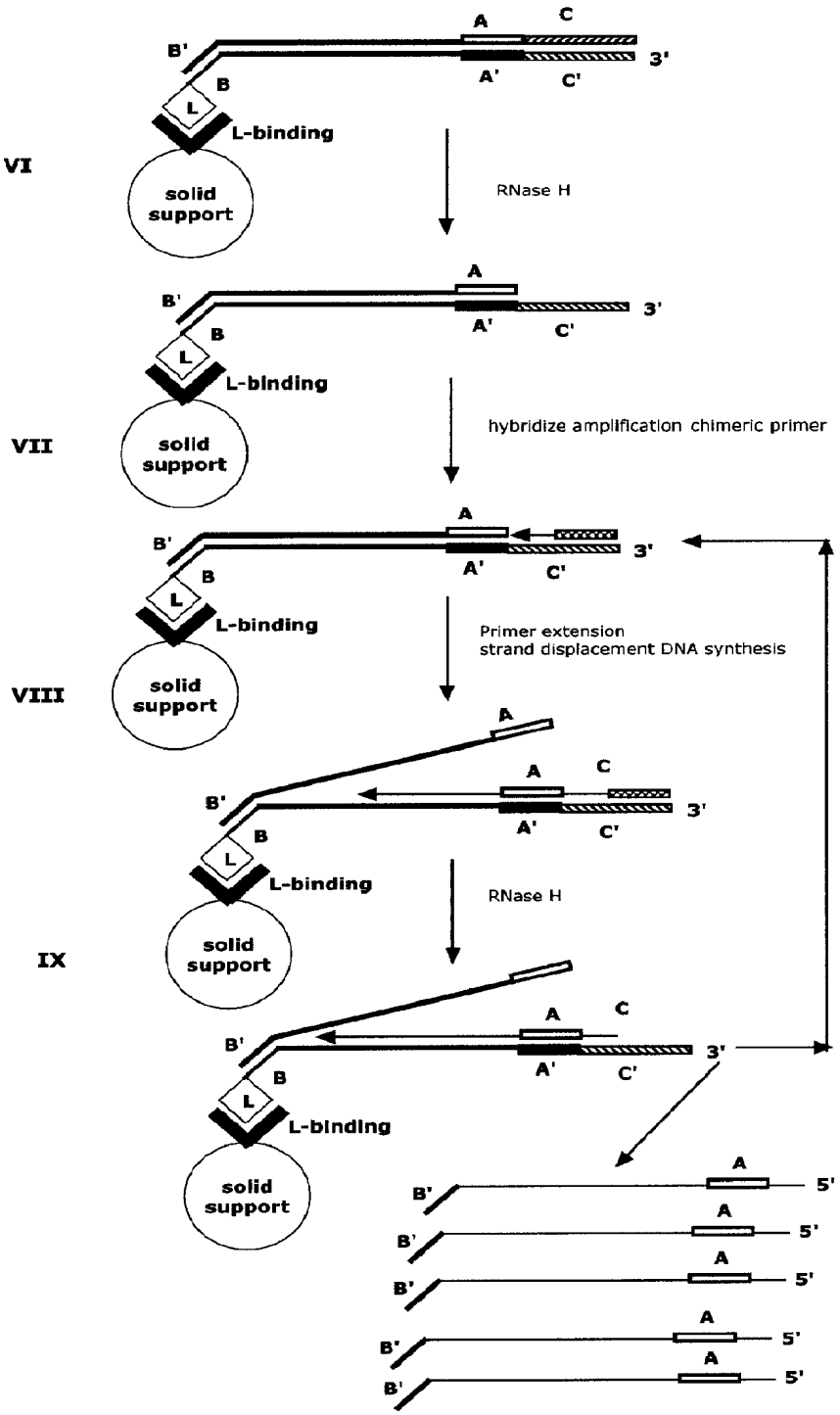

VI
VII
VIII
IX
A
C
B'
3'
A'
C'
B
L
L-binding
solid support
RNase H
hybridize amplification chimeric primer
Primer extension
strand displacement DNA synthesis
5'

Figure 6

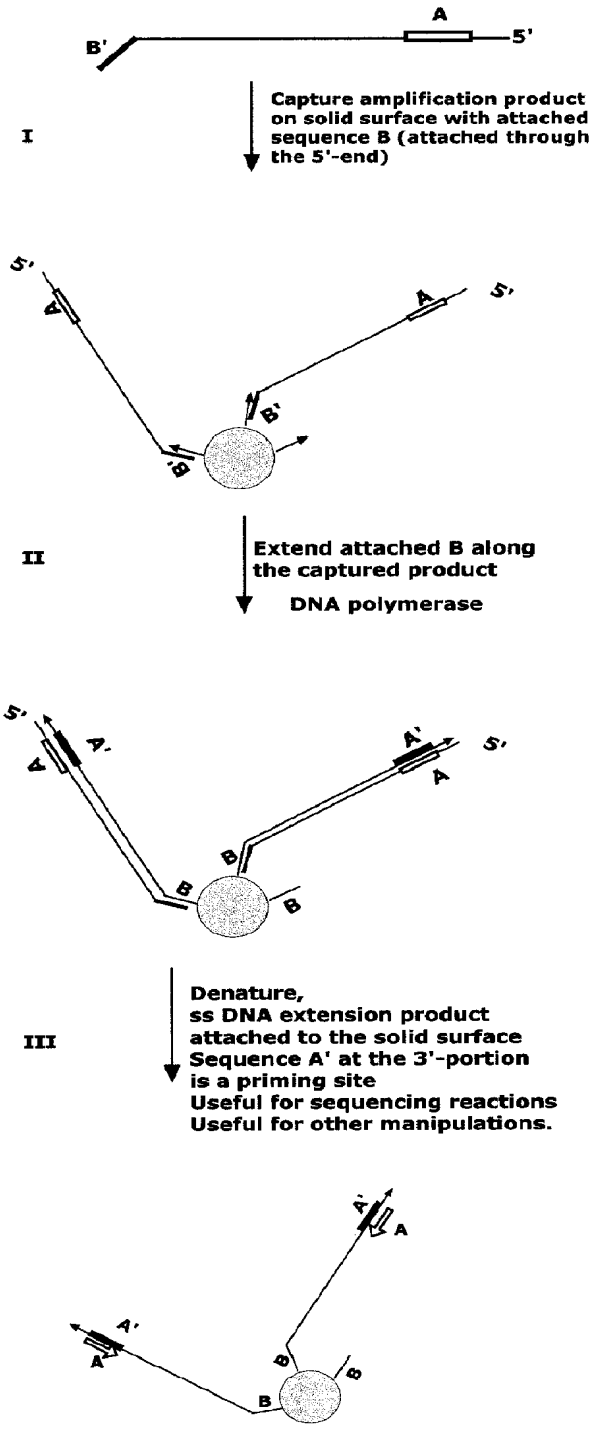

A
B'
5'
I
Capture amplification product on solid surface with attached sequence B (attached through the 5'-end)
5'
A
5'
A
B'
B'
II
Extend attached B along the captured product
DNA polymerase
5'
A
A'
5'
A
B
B
B
III
Denature,
ss DNA extension product attached to the solid surface
Sequence A' at the 3'-portion is a priming site
Useful for sequencing reactions
Useful for other manipulations.
A'
A
A'
A
B
B
B

Figure 7

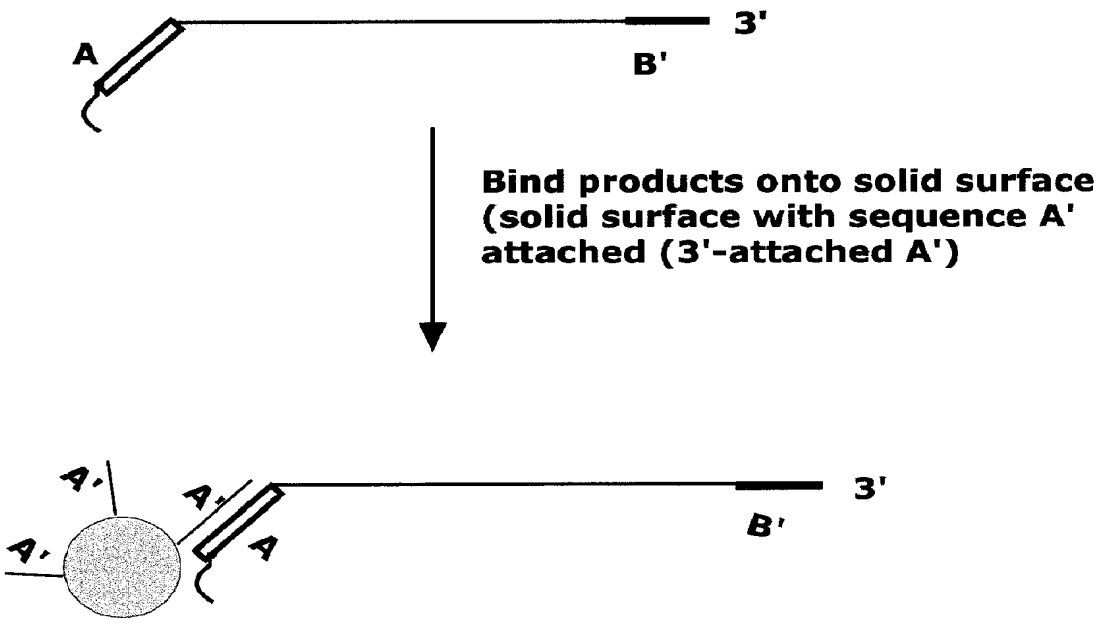

A
3'
B'
Bind products onto solid surface
(solid surface with sequence A'
attached (3'-attached A')
A'
A'
A'
A
3'
B'

Figure 8

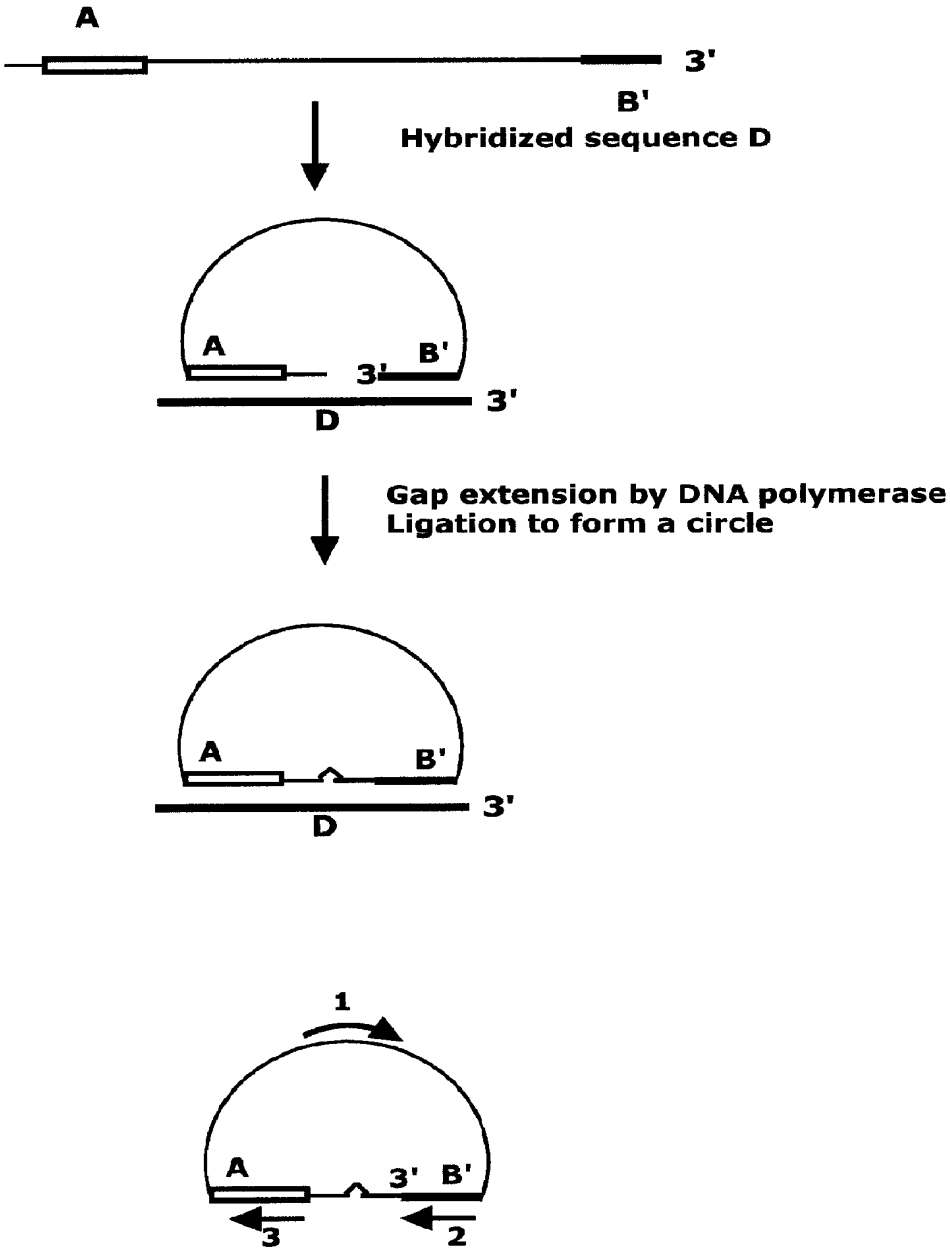

A
3'
B'
Hybridized sequence D
A
B'
3'
3'
D
Gap extension by DNA polymerase
Ligation to form a circle
A
B'
3'
D
1
A
3'
B'
3
2

Figure 9

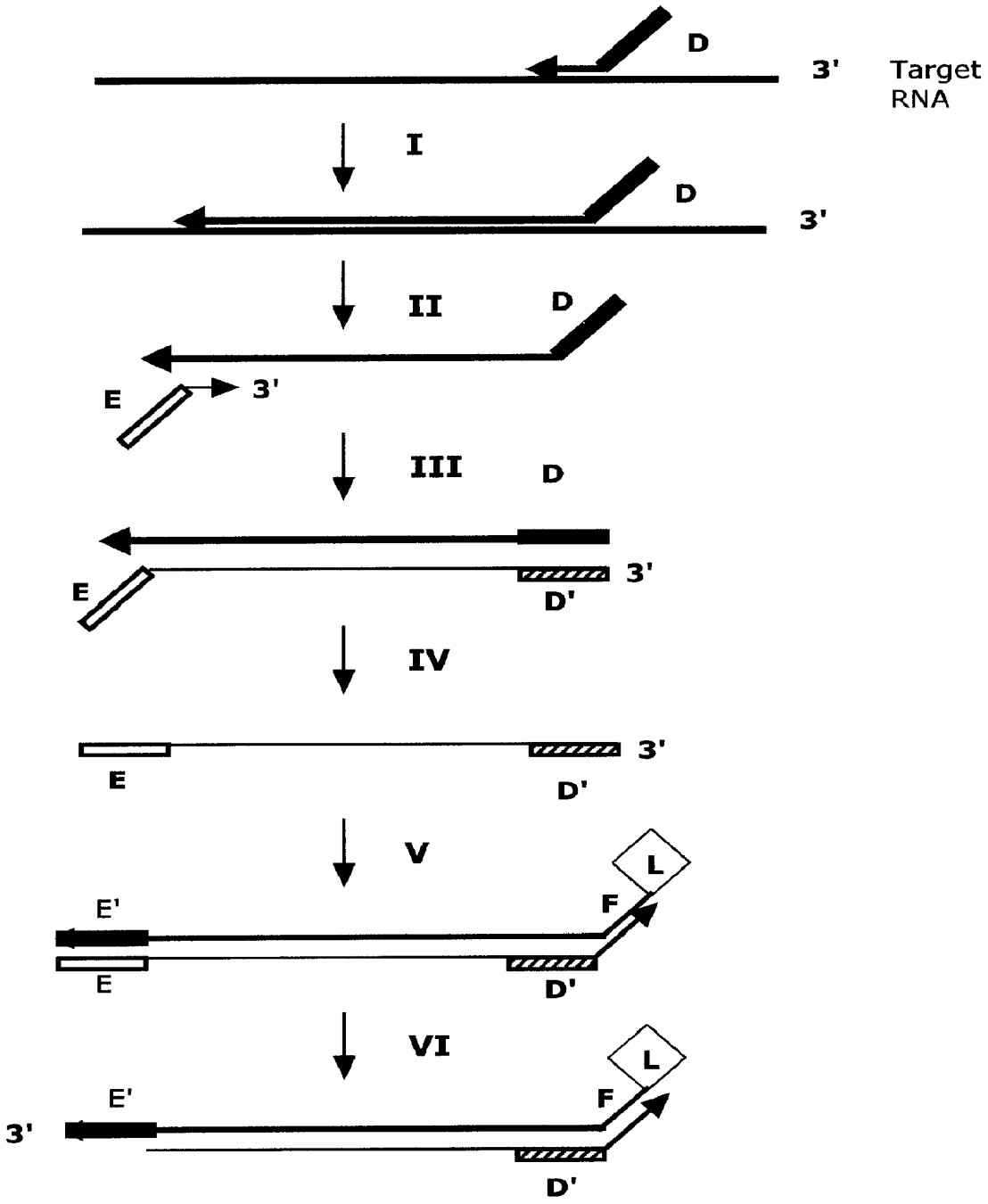


Figure 10

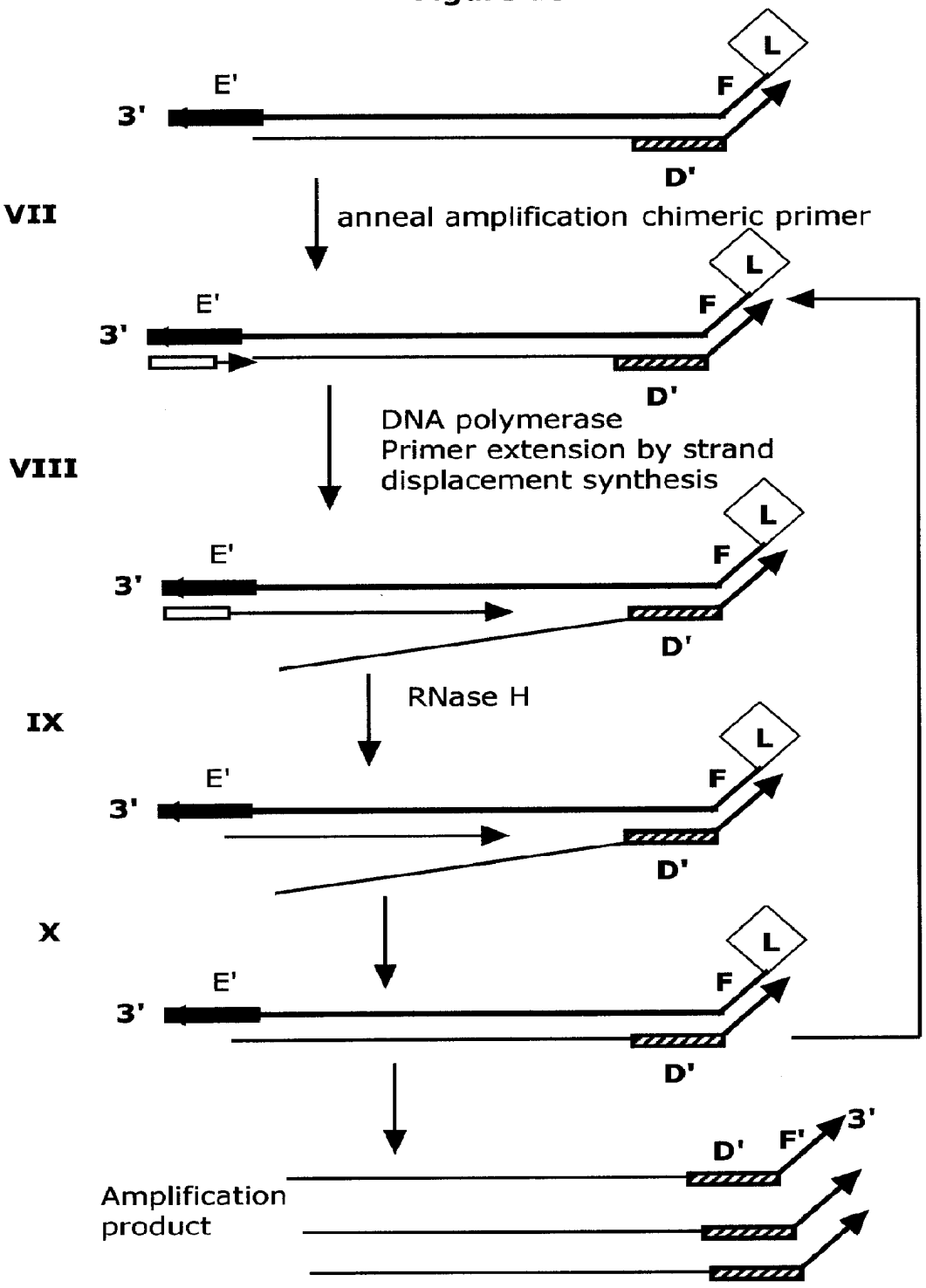

L
E'
F
3'
D'
VII
anneal amplification chimeric primer
DNA polymerase
Primer extension by strand
displacement synthesis
VIII
RNase H
IX
X
F'
Amplification
product

Figure 11

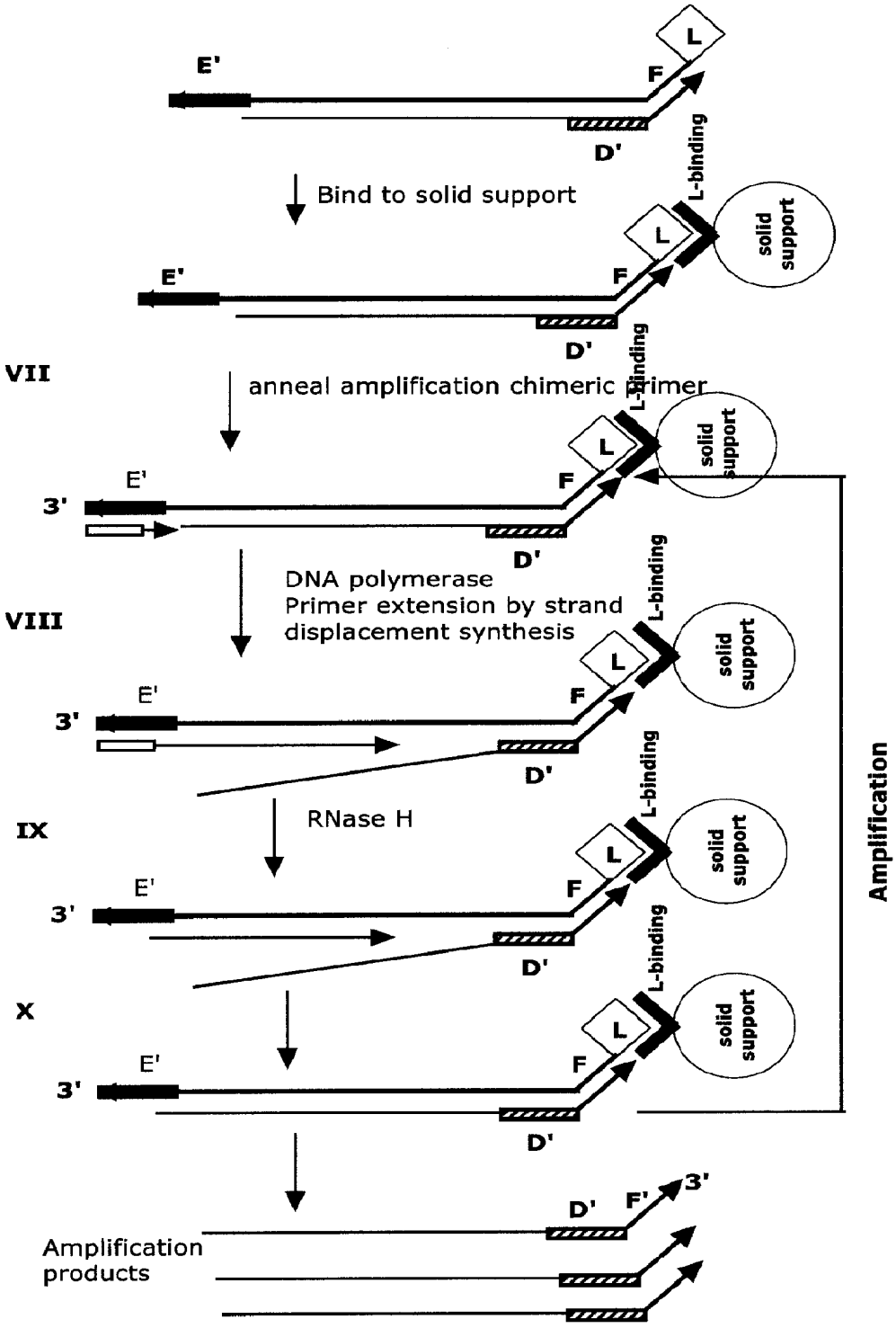

E'
L
F
D'
Bind to solid support
L-binding
solid support
VII
anneal amplification chimeric primer
3'
DNA polymerase
Primer extension by strand displacement synthesis
VIII
RNase H
IX
X
Amplification
Amplification products
F'

Figure 12

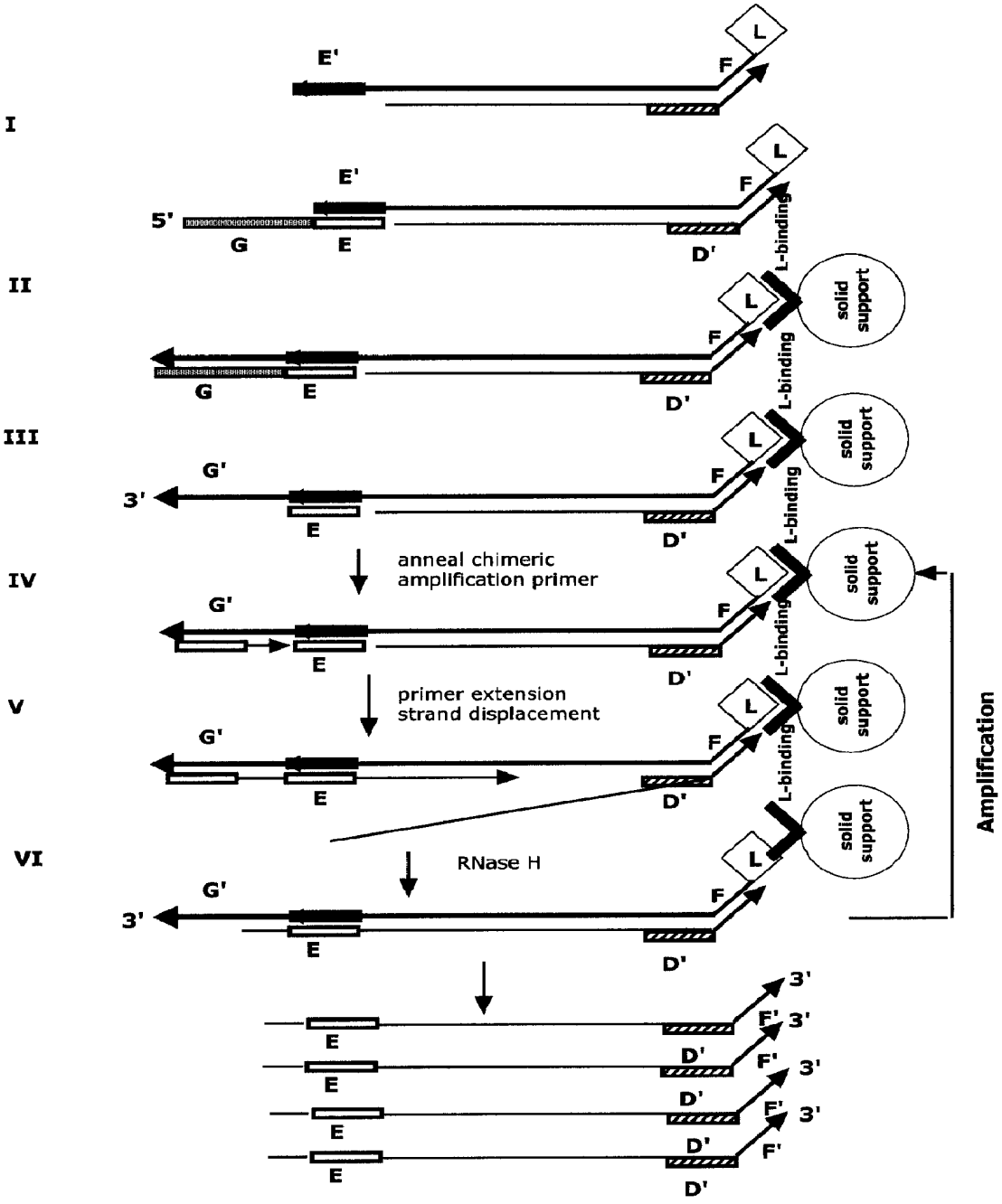

I
II
III
IV
V
VI
E'
F
L
5'
G
E
D'
L-binding
solid support
G'
3'
anneal chimeric amplification primer
primer extension strand displacement
RNase H
F'
Amplification

Figure 13

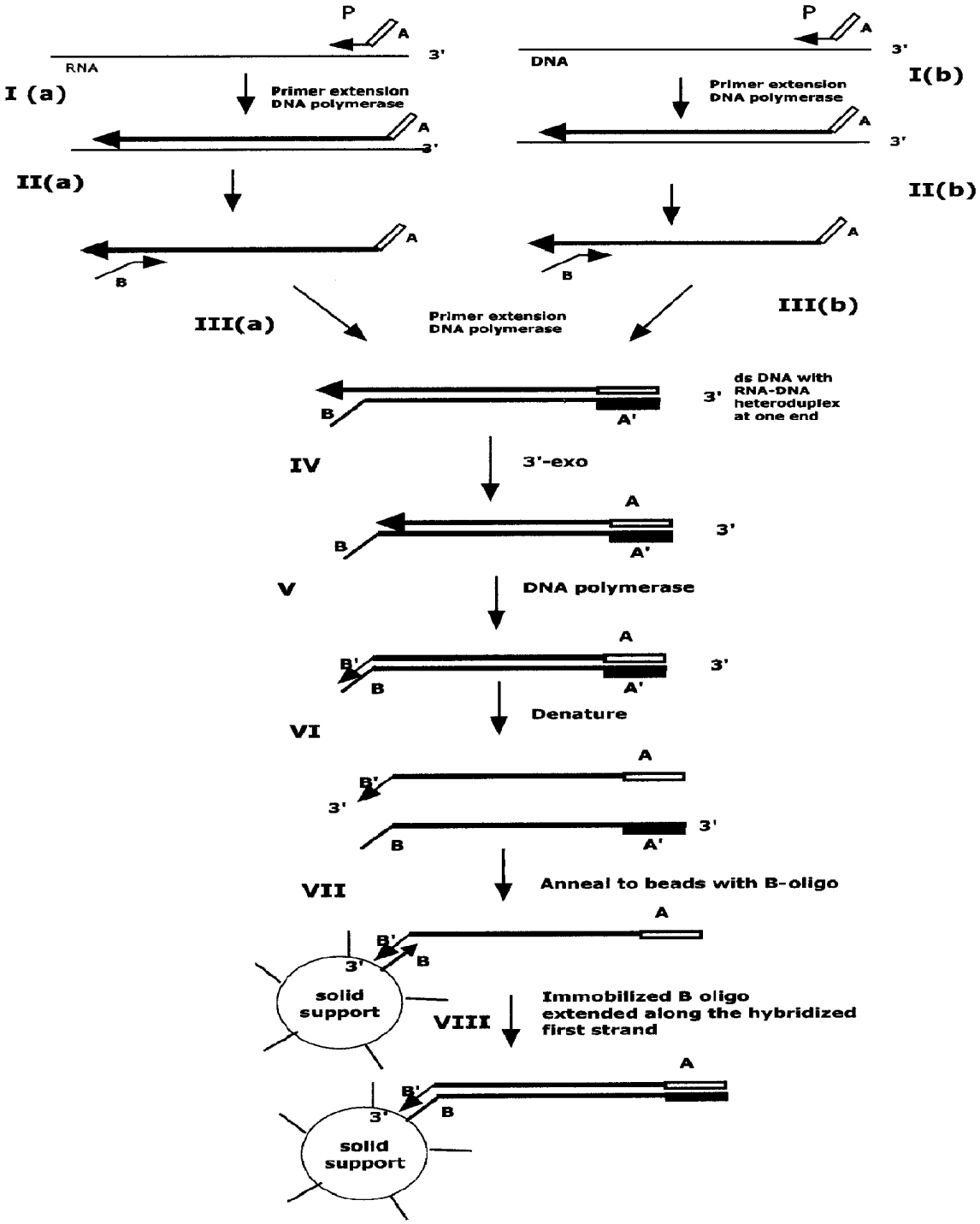

P
A
3'
RNA
I (a)
Primer extension
DNA polymerase
A
3'
II(a)
A
B
III(a)
DNA
I(b)
II(b)
III(b)
Primer extension
DNA polymerase
B
A'
ds DNA with
RNA-DNA
heteroduplex
at one end
IV
3'-exo
V
DNA polymerase
B'
VI
Denature
VII
Anneal to beads with B-oligo
solid
support
VIII
Immobilized B oligo
extended along the hybridized
first strand

Figure 14

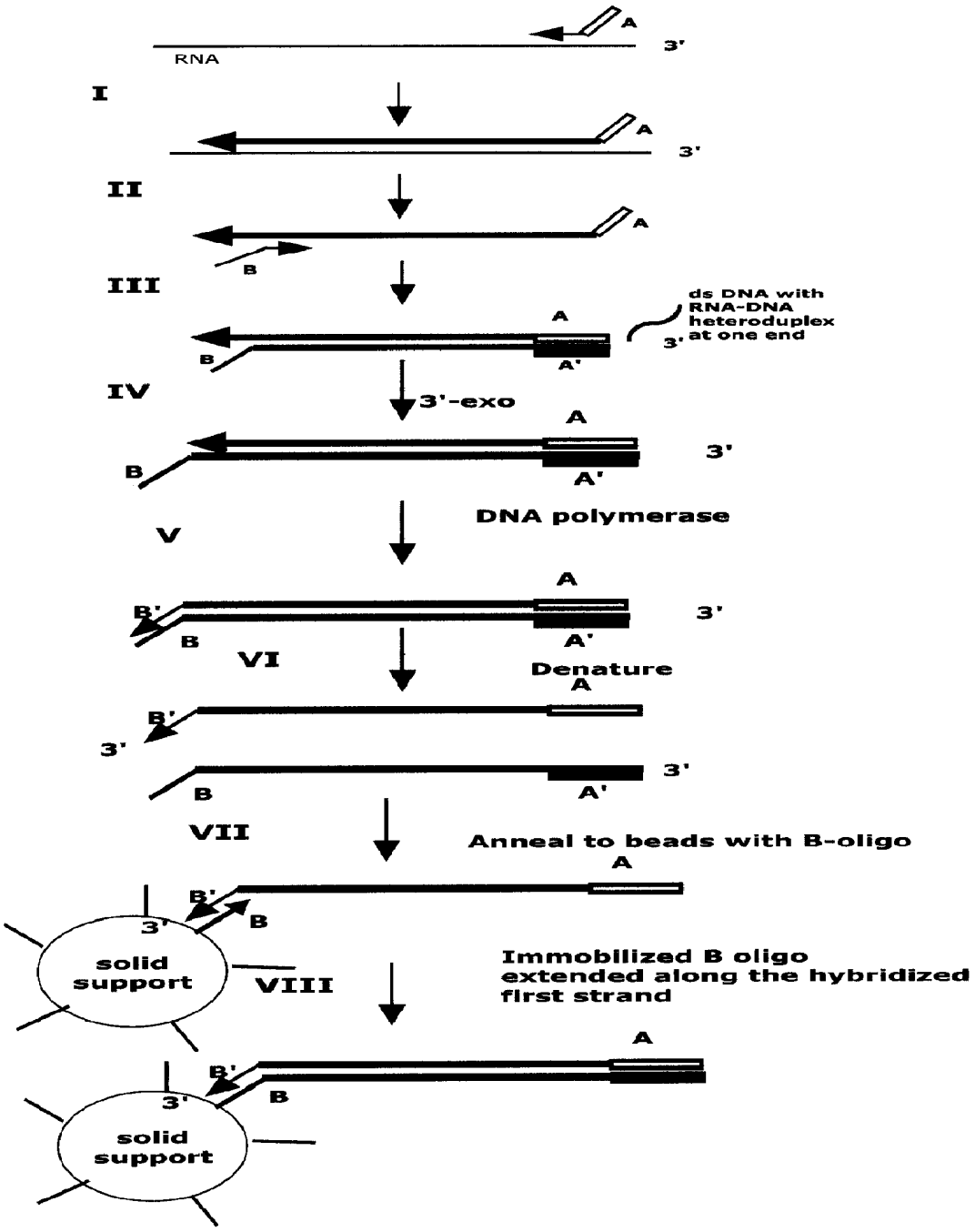

RNA
A
3'
I
II
III
B
ds DNA with
RNA-DNA
heteroduplex
at one end
A'
IV
3'-exo
V
DNA polymerase
B'
VI
Denature
VII
Anneal to beads with B-oligo
solid
support
VIII
Immobilized B oligo
extended along the hybridized
first strand

Figure 15

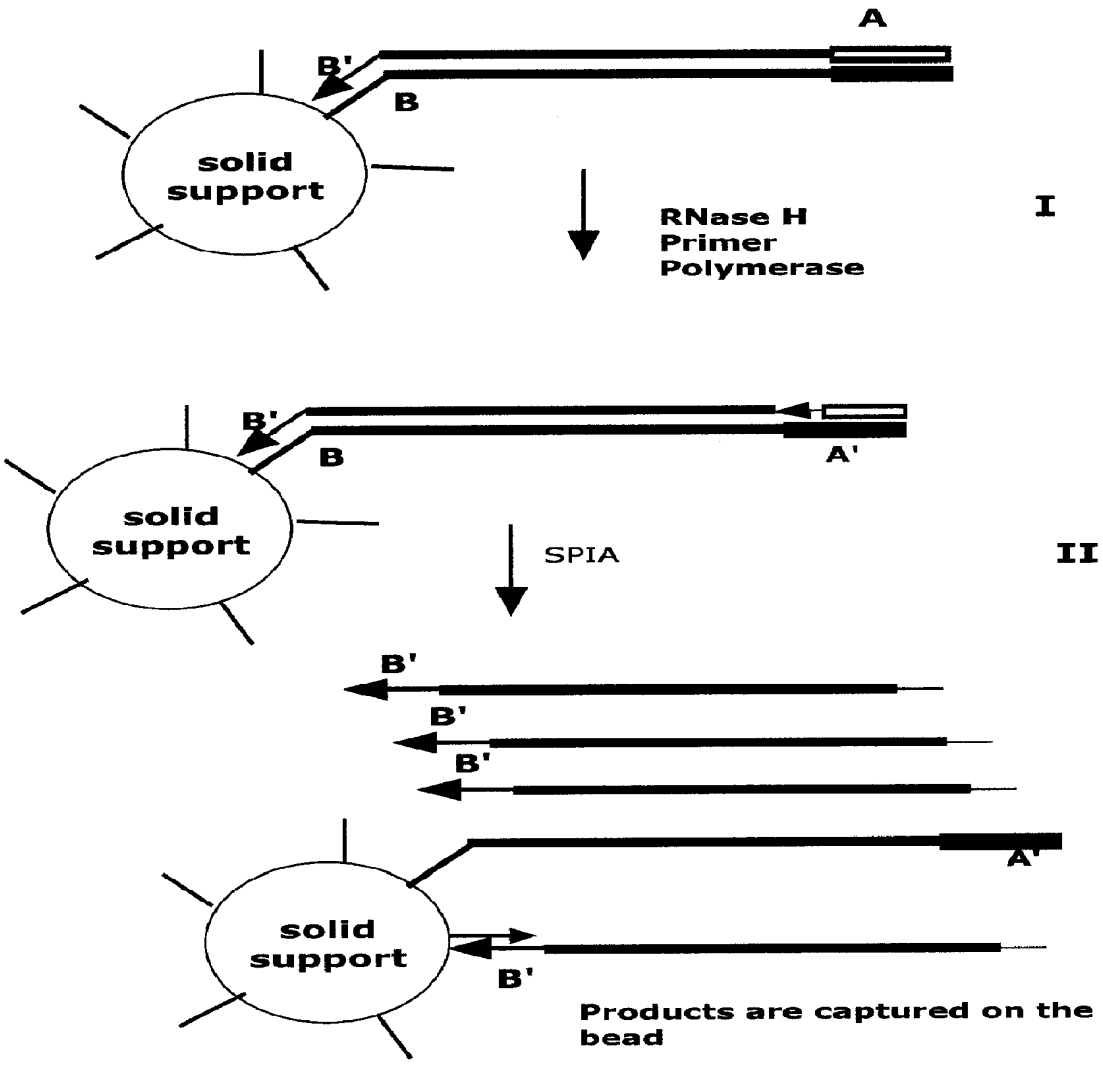

A
B'
B
solid support
RNase H
Primer
Polymerase
I
B'
B
A'
solid support
SPIA
II
B'
B'
B'
A'
solid support
B'
Products are captured on the bead

Figure 16

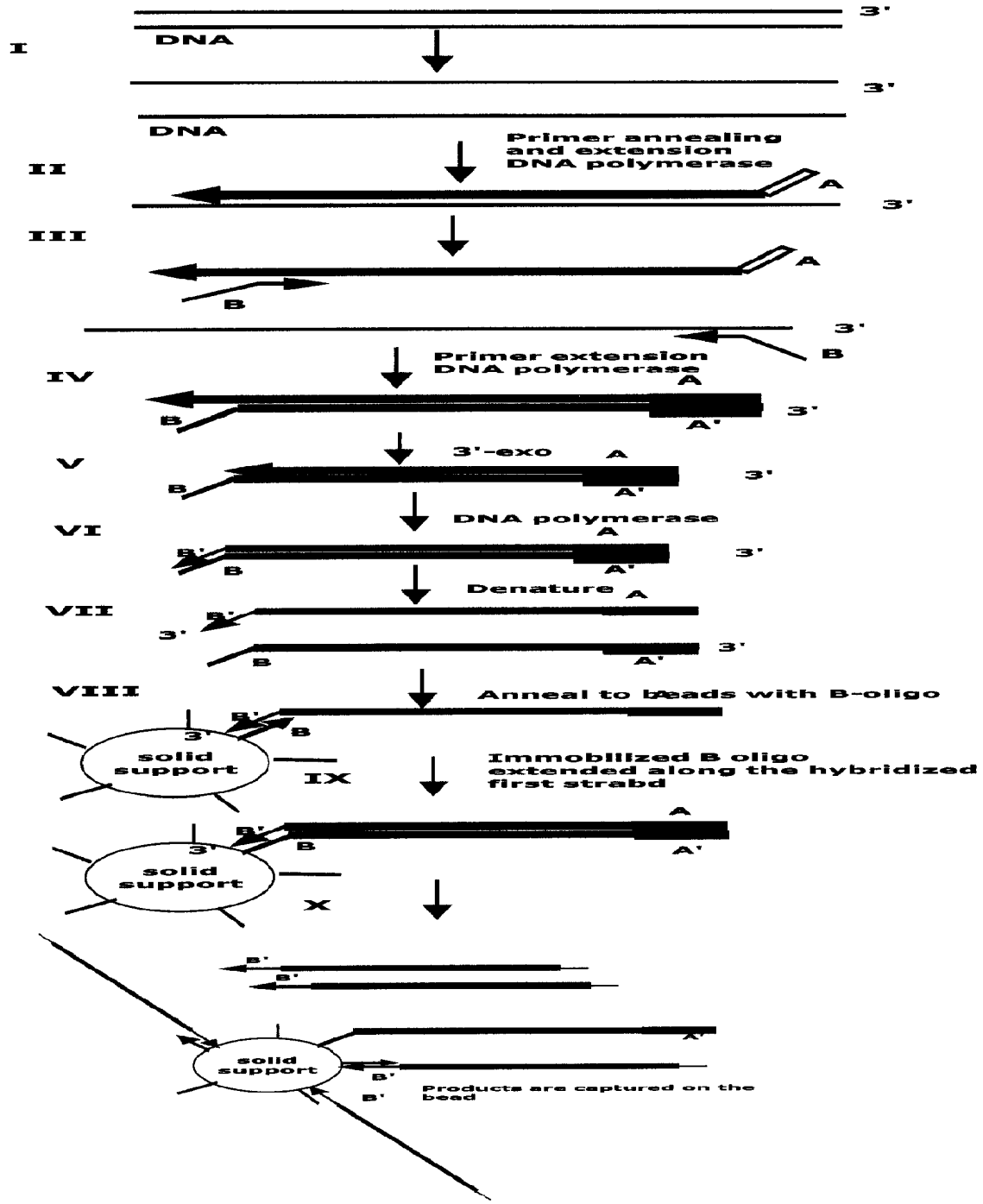

I
DNA
3'
3'
DNA
II
Primer annealing
and extension
DNA polymerase
A
3'
III
A
B
3'
B
IV
Primer extension
DNA polymerase
A
B
A'
3'
V
3'-exo
A
B
A'
3'
VI
DNA polymerase
A
B'
B
A'
3'
VII
Denature
A
B'
3'
B
A'
3'
VIII
Anneal to beads with B-oligo
B'
3'
B
solid
support
IX
Immobilized B oligo
extended along the hybridized
first strabd
A
B'
3'
B
A'
solid
support
X
B'
B'
A'
solid
support
B'
B'
Products are captured on the
bead

Figure 17

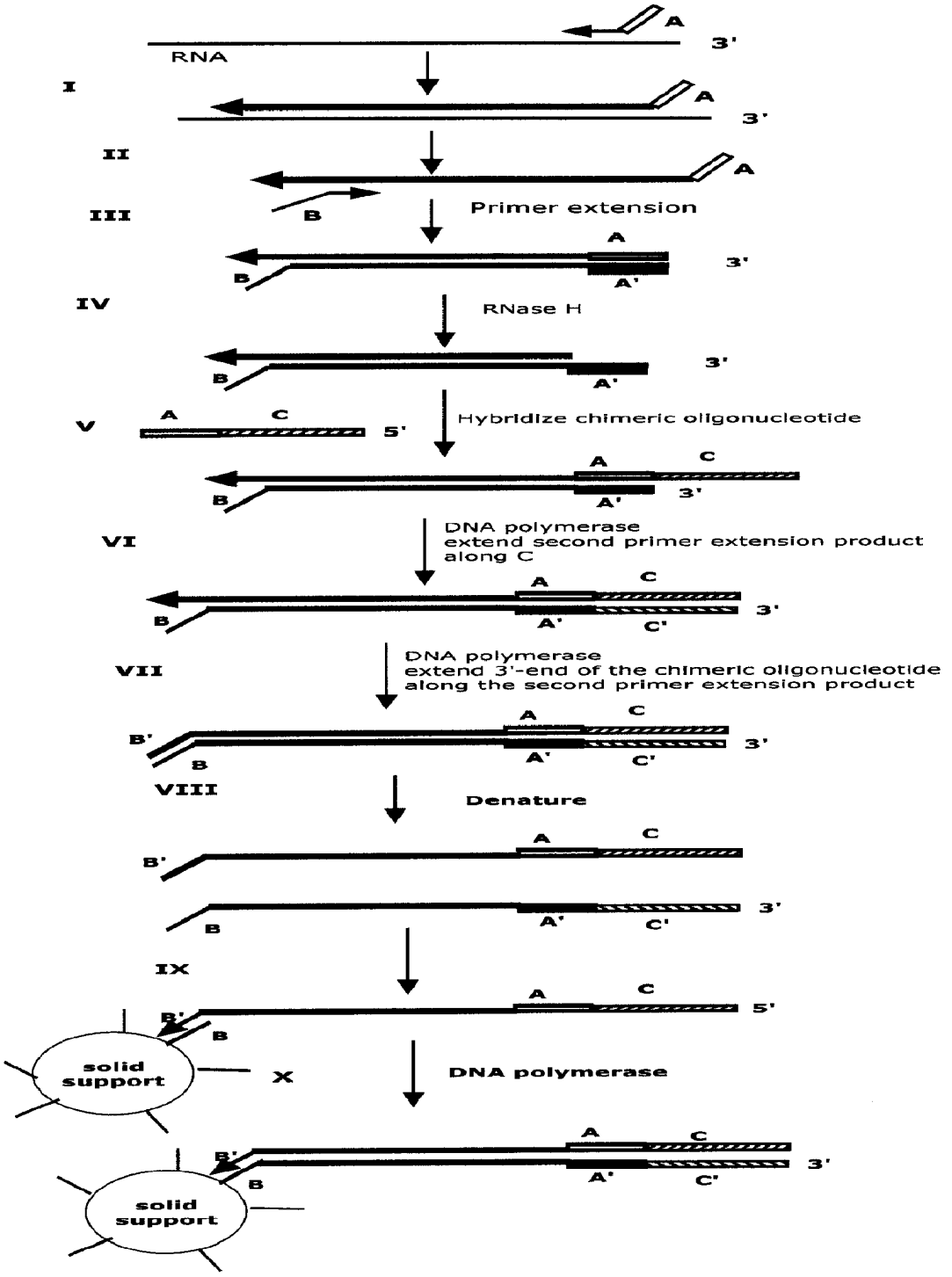

RNA
A
3'
I
II
III
Primer extension
B
IV
RNase H
A'
V
C
5'
Hybridize chimeric oligonucleotide
VI
DNA polymerase
extend second primer extension product
along C
C'
VII
DNA polymerase
extend 3'-end of the chimeric oligonucleotide
along the second primer extension product
B'
VIII
Denature
IX
solid
support
X
DNA polymerase

Figure 18

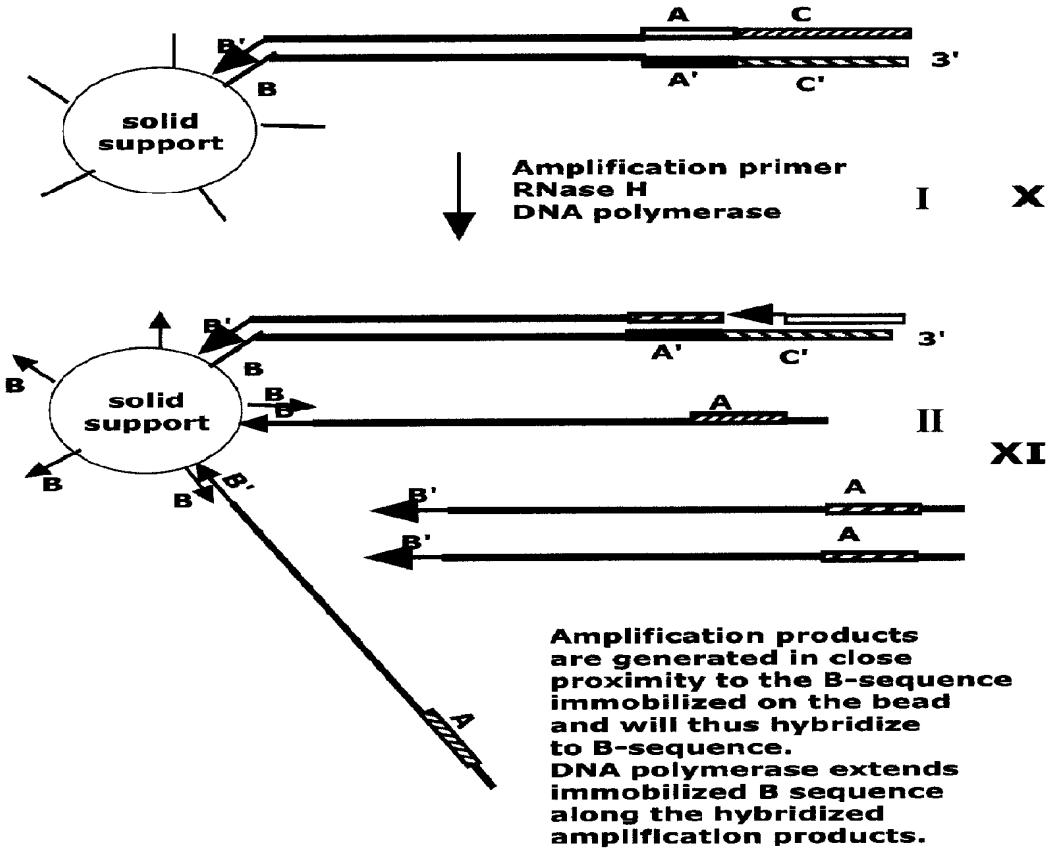
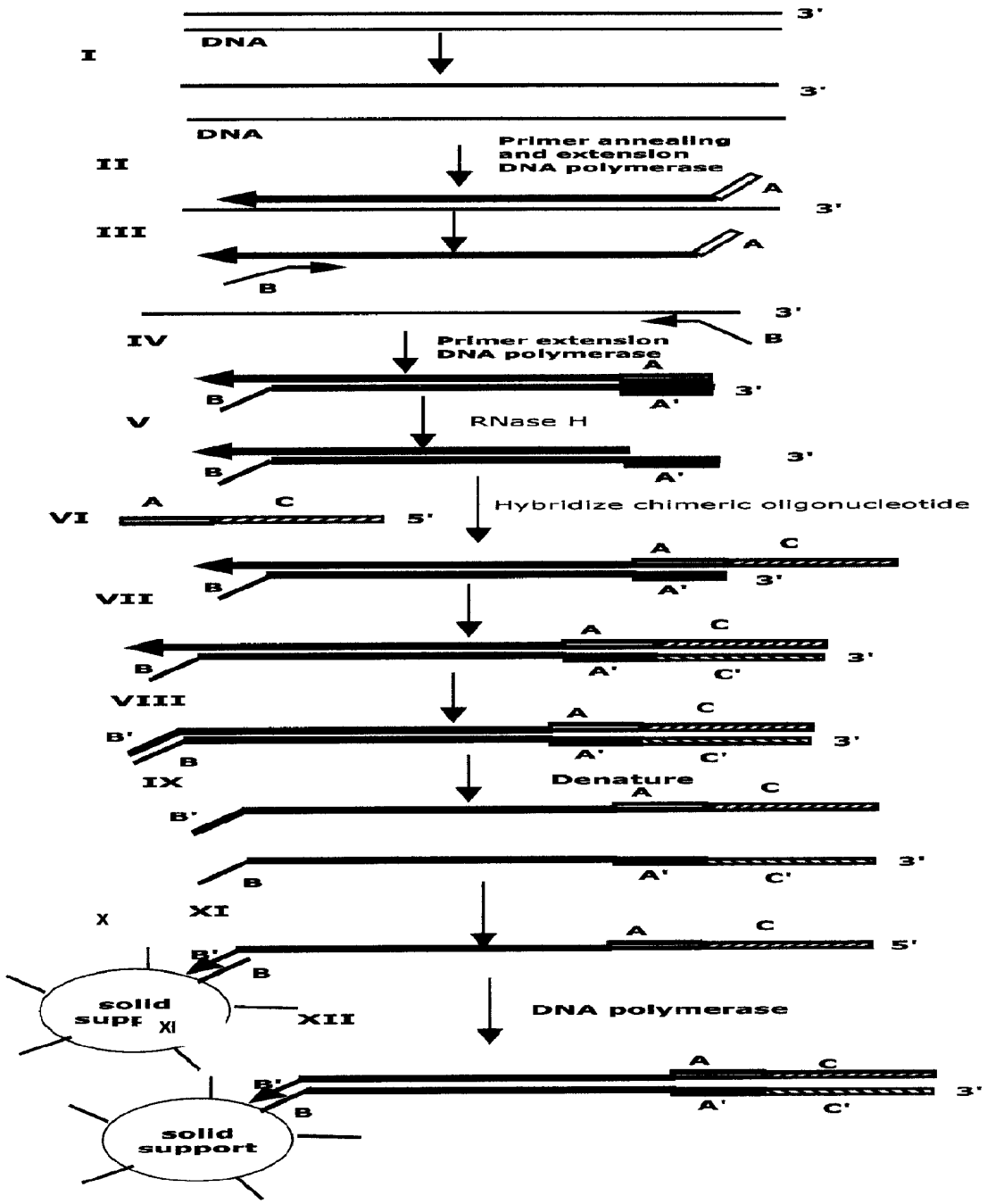

A
C
B'
3'
B
A'
C'
solid support
Amplification primer
RNase H
DNA polymerase
I
X
B'
3'
A'
C'
B
B
solid support
B
A
II
XI
B
B
B'
A
B'
A
A
Amplification products are generated in close proximity to the B-sequence immobilized on the bead and will thus hybridize to B-sequence.
DNA polymerase extends immobilized B sequence along the hybridized amplification products.

Figure 19


I
DNA
3'
3'
II
DNA
Primer annealing
and extension
DNA polymerase
A
3'
III
A
B
IV
3'
B
Primer extension
DNA polymerase
A
B
A'
3'
V
RNase H
3'
B
A'
VI
A
C
5'
Hybridize chimeric oligonucleotide
VII
A
C
B
A'
3'
VIII
A
C
B
A'
C'
3'
IX
B'
B
A
C
A'
C'
3'
Denature
B'
A
C
B
A'
C'
3'
X
XI
B'
B
A
C
5'
solid
supp
XI
XII
DNA polymerase
B'
B
A
C
A'
C'
3'
solid
support

Espacenet

Bibliographic data: WO9207951 (A1) — 1992-05-14

QUANTITATION OF ALDEHYDE-CONTAINING LESIONS IN NUCLEIC ACIDS

Inventor(s): KOW YOKE WAH [US] ± (KOW, YOKE, WAH)

Applicant(s): KOW YOKE WAH [US] ± (KOW, YOKE, WAH)

Classification: **- international:** ***C07K16/44; C12Q1/68; G01N33/53;***
(IPC1-7): C07H15/12; C12N15/00; C12Q1/68
- cooperative: **C07K16/44**; **C12Q1/68**; **G01N33/5308**

Application number: WO1991US07813 19911017 Global Dossier

Priority number (s): US19900600491 19901019

Abstract of WO9207951 (A1)

Methods of quantifying aldehyde-containing lesions, such as abasic sites, in nucleic acid, and reagents specific for aldehyde groups of aldehyde-containing lesions are described. A nucleic acid is reacted with a reagent which specifically complexes with the aldehyde group of the lesion. The amount of complexed reagent is then detected as an indication of the quantity of aldehyde-containing lesions in the nucleic acid.

IDT Exhibit 1020.0972

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:
C12Q 1/68, C07H 15/12, C12N 15/00	A1	WO 92/07951
		(43) International Publication Date: 14 May 1992 (14.05.92)

(21) International Application Number: PCT/US91/07813

(22) International Filing Date: 17 October 1991 (17.10.91)

(30) Priority data:
600,491 19 October 1990 (19.10.90) US

(71)(72) Applicant and Inventor: KOW, Yoke, Wah [US/US]; 84 Pinecrest Drive, Essex, VT 05451 (US).

(74) Agents: DeCONTI, Giulio, A., Jr. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: QUANTITATION OF ALDEHYDE-CONTAINING LESIONS IN NUCLEIC ACIDS

(57) Abstract

Methods of quantifying aldehyde-containing lesions, such as abasic sites, in nucleic acid, and reagents specific for aldehyde groups of aldehyde-containing lesions are described. A nucleic acid is reacted with a reagent which specifically complexes with the aldehyde group of the lesion. The amount of complexed reagent is then detected as an indication of the quantity of aldehyde-containing lesions in the nucleic acid.

IDT Exhibit 1020.0973

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU+**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE*	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

+ Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.

IDT Exhibit 1020.0974

QUANTITATION OF ALDEHYDE-CONTAINING LESIONS IN NUCLEIC ACIDS

Background of the Invention

Apyrimidinic/apurinic (AP), or abasic, sites are the most common lesions produced by DNA. AP sites are formed as a result of cleavage of the N-glycosylic bond between the base and deoxyribose moiety. Abasic sites can be generated spontaneously under physiological conditions by hydrolysis of the N-glycosylic bond. In humans, it has been estimated that about 10,000 AP sites are generated spontaneously per cell per day. AP sites can also be produced following chemical modification of DNA bases by DNA damaging agents such as bleomycin, by alkylating agents, and by ionizing radiation.

AP sites are also intermediates in the base excision repair pathway initiated by the action of DNA N-glycosylases. Thus, the cellular steady state level of AP sites would be elevated as a consequence of base modifications and their subsequent repair. If left unrepaired, AP sites can lead to cell death and/or mutation induction. AP sites are strong blocks to DNA synthesis *in vitro*, and are cytotoxic lesions when present in phage transfecting DNA. Although AP sites are bypassed inefficiently in bacterial cells, when bypass occurs, mutation can result since AP sites are non-instructive lesions.

There are several methods currently available to quantitate AP sites in DNA. These procedures usually involve the estimation of single strand breaks formed as a result of alkali treatment of AP site-containing DNA, including alkali elution and DNA unwinding. Recently, Weinfeld et al., (1990) Biochemistry 29:1737-1743, using a ^{32}[P]-postlabeling assay, was able to quantitate AP sites in DNA. AP sites can also be measured by the use of ^{14}C-labeled methoxyamine. Talpaert-Borle, M. and Liuzzi, M. (1983) Biochem. Biophys. Acta 740:410-416; Liuzzi, M. and Talpaert-Borle, M. (1988) Int. J. Radiat. Biol. 54:709-722. Because ^{14}C-methoxyamine is only available at low specific activity, this method is relatively insensitive unless large quantities of DNA are available.

All these known procedures detect AP sites in DNA at the femtomole level. However, due to the chemical instability of AP sites, the lengthy procedures involved in the preparation of DNA samples can lead to strand breaks and possible oxidation of the aldehyde group, potentially reducing the number of measurable AP sites in the DNA.

Summary of the Invention

This invention pertains to a method of quantifying aldehyde-containing lesions such as abasic sites in nucleic acids, such as DNA, and to reagents specific for aldehyde groups of an aldehyde-containing lesion for use in the method. The method comprises reacting the nucleic acid with a reagent that specifically reacts with the aldehyde group of the lesion (an "aldehyde reactive probe" or

ARP) to form a detectable complex and then detecting the complexed reagent. The amount of complexed reagent is proportional to the quantity of aldehyde-containing lesions in the nucleic acid.

The ARP is chemically specfic for aldehyde groups in nucleic acid. The ARP can be an O-alkyl hydroxylamine, preferably O-4-nitrobenzyl hydroxylamine and N-biotinyl N'-(aminooxy)acetyl hydrazide. The complexed ARP can be detected in a variety of different ways, including colorimetrically and immunologically. Colorimetric detection can be achieved by labelling the ARP with an enzyme such as a peroxidase that converts a nob-absorbing substrate to an absorbing product (color-forming enzyme). For example, the ARP can be a biotinylated O-alkyl hydroxylamine such as N-biotinyl N'-(aminooxy)acetyl hydrazide. After aldehyde-specific complexation with the nucleic acid, the biotinylated ARP can be tagged with a biotinylated color-forming enzyme through avidin and the activity of the enzyme measured as an indication of the quantity of aldehyde-containing sites in the nucleic acid. Alternatively, the ARP can be detected immunologically by using an antibody specific to an epitope formed by the reaction of the ARP with the nucleic acid and measuring the complex-specific antibody bound as an indication of the quantity of aldehyde-containing lesions.

The invention also includes kits for performing the methods to quantify aldehyde-containing lesions in nucleic acids. For colorimetric assays, the kit can include a biotinylated ARP, a biotinylated enzyme, its chromogenic substrate and avidin. For an immunological assay, the kit can include an antibody

specific for the epitope formed by complexation of the ARP and the nucleic acid and appropriate reagents for detection of the antibody.

The methods and reagents of this invention provide for simple and sensitive quantification of aldehyde-containing lesions in nucleic acids. The methods can be used to test for genotoxicity of various agents.

Brief Description of the Drawings

Figure 1 illustrates the structure of N-biotinyl N'(aminooxy)acetyl hydrazide of the invention.

Figure 2 illustrates that the signal measured by colorimetric detection of a biotinylated ARP was proportional to the number of apyrimidinic/apurinic (AP) sites in the DNA sample.

Figure 3 illustrates that the reagent of the invention was able to detect AP sites in both single stranded DNA and double stranded DNA.

Figure 4A illustrates that the number of AP sites detected in uracil-containing f1 single stranded DNA was proportional to the amount of uracil N-glycolsylase added.

Figure 4B illustrates that the number of AP sites detected after endonuclease III digestion of calf thymus DNA containing thymine glycols was linearly proportional to the original number of thymine glycols in the DNA sample.

Figure 5A illustrates that when f1 single stranded DNA was exposed to ionizing radiation, the number of ARP sites detected was proportional to the x-ray dose.

Figure 5B illustrates that the number of AP sites in ϕX RF DNA was proportional to increasing doses of x-rays.

Figure 5C illustrates that the number of AP sites in calf thymus DNA was proportional to increasing doses of x-rays.

Figure 6 illustrates the rate of production of ARP sites and thymine glycols in heat-inactivated E. coli cells.

Figure 7 illustrates that the steady state production and subsequent repair of AP sites in E. coli cells was detectable by the assay of the invention.

Figure 8 illustrates hapten inhibition of anti-NBHA-dRp (5' phosphodeoxyibosyl O-4-nitrobenzyl hydroxylamine) antibody reactivity.

Figures 9A and 9B illustrate inhibition of anti-NBHA (O-4-nitrobenzyl hydroxylamine) antibody reactivity by calf thymus DNA containing NBHA-modified AP sites.

Figure 10 illustrates inhibition of anti-NBHA antibody reactivity by NBHA residues in DNA. The symbols are the same as for Figures 9A and 9B, and data were calculated from Figure 9B.

Figure 11 illustrates specificity of anti-NBHA-dRp antibody.

Figure 12A illustrates the effect of DNA denaturation on antibody reactivity.

Figure 12B illustrates the reactivity of anti-NBHA-dRp antibody toward single stranded and duplex DNA.

Figure 13 illustrates that the number of antibody reactive sites detected in f1 single stranded DNA modified with NBHA and exposed to ionizing radiation, was proportional to the x-ray dose.

Detailed Description of the Invention

According to the methods of this invention, aldehyde-containing lesions in nucleic acid are detected by reacting the nucleic acid with an ARP and then detecting the amount of ARP complexed as an indication of the amount of aldehyde-containing sites in the nucleic acid. The ARP is chemically specific for aldehyde groups in the nucleic acid. The amount of ARP complexed to the nucleic acid is directly proportional to the quantity of aldehyde-containing lesions in the nucleic acid.

Aldehyde-containing lesions detectable by the method of this invention may be naturally occurring, or may be induced by such methods as x-irradiation, chemical treatment, or genetic alteration. Aldehyde-containing lesions are primarily apyrimidinic and apurinic (abasic) sites (AP). These occur in nucleic acids as a result of cleavage of the N-glycosylic bond between the pyrimidine or purine base and the ribose. Other aldehyde-containing lesions such as formamidopyrimidine, which is an imidazole ring opened purine induced by ionizing radiation, can be detected by the method of this invention.

Aldehyde-containing lesions detected in, for example, x-ray irradiated DNA, probably include AP sites, such as alkali labile sugar lesions, strand

breaks containing AP sites at the terminus, as well as formamidopyrimidine. Certain classes of AP sites and strand breaks which do not contain an aldehyde group would not be detected by these methods. Since more than one type of x-ray damage is detectable by these methods, they provide a simple and sensitive method for detecting the presence of x-ray damage.

The nucleic acid can include DNA, RNA, rRNA, tRNA, and the like. Typically the nucleic acid is DNA. DNA is generally prepared for the assay as described by Ausubel et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York.

The ARP of this invention are reagents which specifically react with the aldehyde group of a lesion of a nucleic acid. In preferred embodiments, the ARP is an O-alkyl hydroxylamine. O-alkyl hydroxylamines react specifically with aldehyde groups in the nucleic acid. A particularly preferred ARP is O-4-nitrobenzyl hydroxylamine (NBHA), which forms an iminic linkage with abasic sites. This reaction stabilizes the phosphodiester bond 3' to the original aldehyde-containing site, producing O-NBHA residues at the site.

The ARP and the nucleic acid are reacted by incubating them under conditions conducive for reaction of the ARP with the aldehyde groups of the nucleic acid. In general, the incubation is performed in a physiological buffer at room temperature.

The ARP reacts with the aldehyde group to form a complex which can be detected in various ways. The ARP can be tagged with a detectable label such as an enzyme, radioisotope or fluorescent compound. In a

preferred embodiment, the ARP is biotinylated. For example, a biotinylated O-alkyl hydroxylamine can be formed by reacting O-carboxymethyl hydroxylamine and biotin hydrazide in the presence of carbodiimide. The biotinylated ARP can be detected by means of a biotinylated enzyme bound to the complexed ARP through an avidin-mediated linkage. Preferred enzymes are peroxidases, such as horseradish peroxidase, but other types of enzyme, such as phosphatases, oxidases, galactosidases and ureases may be used.

The complexed ARP can also be detected immunologically. This can be done with an antibody that specifically binds to an epitope formed by the complexation of the ARP to the nucleic acid. For example, antibodies can be raised against an epitope formed by the complex of NBHA and an abasic site.

A monoclonal antibody is preferred. Monoclonal antibody specific for an ARP complex can be prepared by standard somatic cell hybridization technqiues (Milstein and Kohler (1975) Nature 256:495-497). Briefly, an appropriate animal, such as a mouse, is immunized with an analog of the ARP complex. For example, to produce an antibody against NBHA sites in DNA, 5'phosphodeoxyribosyl O-4-nitrobenzyl hydroxylamine can be conjugated to a protein such as BSA and used as the immunogen. After immunization, the animal is sacrificed, and cells are removed from its spleen and fused with myeloma cells or other immortalizing cells. The resulting antibody-producing cells (hybridomas) are screened to idenfity the antibodies of desired specificity. The antibodies can be screened for reactivity with the

immunogen by an ELISA or similar assay. The antibody can then be evaluated for appropriate sensitivity and specificity in the detection of ARP sites in nucleic acids.

For detection, the antibody can be labelled with an enzyme, a radioisotope, fluorescent compound or other detectable label. The enzyme can be peroxidase alkaline phosphatase, glucose oxidase, galactosidase, urease, and the like. In a preferred embodiment, the antibody is detected by means of a second antibody directed against it. For example, the second antibody can be directed against the Fc region of the first antibody (e.g., if the primary antibody is murine, a goat anti-mouse antibody can be used). In this format, the second antibody is labelled for detection as described above.

For convenience and standardization, the reagents for performing the assays of this invention may be assembled into assay kits. These kits contain, in individual containers, predetermined quantities of the reagents necessary to perform a quantitation of aldehyde-containing lesion sites in nucleic acid samples. The kits would also include appropriate standards and controls.

The components of the kit depend on the format of the assay. For example, a kit for performing an enzymatic, colorimetric assay can comprise biotinylated ARP (e.g., biotinylated NBHA), a biotinylated enzyme (e.g., horseradish peroxidase), a chromogenic substrate for the enzyme (e.g., O-phenylenediamine) and avidin. For an immunological assay, a kit can contain the ARP, an antibody specific for the ARP complex with the nucleic acid

and a secondary antibody detectably labelled, for example, with an enzyme for colorimetric determination.

The methods of this invention offer several advantages depending on the embodiment of the assay. In embodiments which employ a biotinylated hydroxylamine as ARP, the chemical specificity for the aldehyde group in AP sites can be combined with the ease and sensitivity of the avidin/biotin system in an enzyme assay (e.g., Elisa-like microtiter plate assay). For example, employing this method, 1 AP site per 10^5 nucleotides can be detected where approximately 80 ng of DNA is bound per microtiter well. This detection limit translates to about 1.5 femtomoles of AP sites. Thus, the sensitivity of the assay is comparable to the current available methods for detection of AP sites. However, in contrast to other procedures, the assay is rapid and simple to perform, and offers the possibility of processing a large number of DNA samples through automation.

The antibody-based embodiment demonstrates about the same sensitivity as above and as other procedures for measuring AP sites including alkali elution (Brent et al. (1978) DNA Repair Mechanisms, eds. Hanawalt, Friedberg, and Fox (Academic Press, New York), pp. 19-22), DNA unwinding (Kohn et al. (1981) DNA Repair: A Laboratory Manual of Research Procedures, eds. Friedberg and Hanawalt (Marcel Dekker, New York), Vol. 1., Part B, pp. 379-401), ^{32}P-postlabeling (Weinfeld et al. (1990) Biochemistry 29:1737-1743). Some of these methods involve laborious procedures (Brent et al. (1978) DNA Repair Mechanisms, eds. Hanawalt, Friedberg and Fox

(Academic Press, New York), pp. 19-22 ; Kohn et al. (1981) DNA Repair: A Laboratory Manual of Research Procedures, eds. Friedberg and Hanawalt (Marcel Dekker, New York), Vol. 1, Part B, pp. 379-401; Birnboim and Jevcak (1981) Cancer Res. 41:1889-1892 and Weinfeld et al. (1990) Biochemistry 29:1737-1743) and others require the use of radioactive materials (Weinfeld. et al. (1990) Biochemistry 29:1737-1743; Talpaert-Borle and Liuzzi (1983) Biochem. Biophys. Acta 740:410-416 and Liuzzi and Talpaert-Borle (1988) Int. J. Radiat. Biol. 54:709-722). In addition, the modification of AP sites by NBHA helps to stabilize the AP sites in the nucleic acid, and therefore, helps to reduce the loss of labile AP sites during purification of DNA from cells. Thus, this embodiment of the method should have broad applicability for detecting abasic sites in the DNA extract from damaged cells.

The methods of this invention have many different utilities. They allow researchers to detect and quantitate different types of damage to nucleic acids, thus facilitating the study of DNA repair and the role of DNA repair in the etiology of different diseases. In the clinic, the assays can be used to monitor the extent of genotoxicity resulting from radiotherapy or chemotherapy. The amount of lesions in DNA from individuals (in lymphocytes from blood samples) undergoing a specific therapy can be quantitated. The DNA can be treated with a battery of damage-specific enzymes, which will remove the lesions, leaving behind an AP site that is detectable. The assays will allow one to quantitatively assess the extent of DNA damage (and

DNA repair). Information derived from these measurements should help physicians establish an effective regime of chemotherapy for a cancer patient optimizing therapeutic gain while minimizing potential side effects of the chemotherapeutic agent.

The methods of this invention can be used to monitor individuals for possible occupational exposure to genotoxic agents. People exposed to chemicals or radioactivity have a higher risk of genetic damage. DNA from individual workers (lymphocytes from blood) can be monitored for DNA damage. This information can be used to set guidelines and to limit individual exposure.

The methods of this invention provide a useful supplement for the Ames' Test for mutagenicity. Every year, thousands of new chemicals are synthesized, and their genotoxicity is not known. The Ames' Test has been used to screen chemicals that might be mutagenic before they are released into the public sector. The assay of this invention can be used in conjunction with and, in some cases, in place of the Ames' Test. DNA in vitro or in vivo can be treated directly with the new chemical and damage to DNA can be quantitated. In addition, sensitivity can be increased by exposing sensitive cells (prokaryotes and eukaryotes) to these agents, and DNA can be used to assess the damage. Most importantly, human cell lines can be used as a model to assess a chemical's genotoxic effect on human DNA.

The methods can also be used to test water supplies for possible contamination with genotoxic chemicals and to detect genotoxic chemicals in the

environment, including possible contaminations in air or water resources.

The invention is illustrated further by the following examples.

EXEMPLIFICATION

EXAMPLE 1

Chemicals

O-4-nitrobenzyl, O-methyl, and O-carboxymethyl hydroxylamine were purchased from Aldrich Chemical Company Inc., 1001 West Saint Paul Avenue, Milwaukee, WI 53233. 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDAC), biotin hydrazide and methylmethanesulfonate were obtained from Sigma Chemical Company, P.O. Box 14508, St. Louis, MO 63178.

Enzymes.

Endonuclease III was prepared as described in Katcher and Wallace (1983) Biochemistry 22:4071-4081, and Kow and Wallace (1987) Biochemistry 26:8200-8206, from an E. coli strain harboring an overproducing endonuclease III plasmid pHIT1 obtained from R.P. Cunningham, SUNY, Albany. The procedure was changed by replacing the DNA agarose step with FPLC mono S chromatography, for which the reagents are available from Pharmacia Biochemicals, 800 Centennial Avenue, Piscataway, NJ 08854, following the methods of Asahara et al. (1989) Biochemistry 28:4444-4449. Uracil N-glycosylase was prepared from an E. coli strain carrying an overproducing plasmid of uracil

N-glycosylase according to the method of Lindahl et al. (1977) J. Biol. Chem. 252:3286-3294.

Nucleic acids.

Calf thymus DNA was obtained from Pharmacia Biochemicals, 800 Centennial Avenue, Piscataway, NJ 08854. PM2, ϕX-174 RF and f1 DNA are prepared routinely in the laboratory. Calf thymus DNA containing AP sites was prepared by heat/acid depurination (Lindahl, T. and Nyberg, B. (1972) Biochemistry 11:3610-3618) and the number of AP sites was estimated by correlation with the rate of heat/acid depurination of PM2 DNA treated under the same conditions (Kow, Y.W. (1989) Biochemistry 28:3280-3287). The number of AP sites in PM2 DNA was assayed by alkali fluorometry (Kow, Y.W. (1989) Biochemistry 28:3280-3287 and Futcher, S.D. and Morgan, A.R. (1979) Can. J. Biochem. 57:932-938). DNA containing other lesions including thymine glycols, pyrimidine dimers, uracil, urea residues, O-methyl hydroxylamine residues and reduced AP sites were prepared as previously described in Futcher and Morgan (1979) Can. J. Biochem. 57:932-938, and Kow et al. (1990) Mutat. Res. 235:147-156. E. coli DNA was isolated after lysis of the cells with sodium dodecyl sulfate (SDS) and lysozyme (Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York). After the RNAse and protease K treatment described in the Futcher and Kow references, the DNA was extracted with phenol twice and extensively dialyzed in 10 mM Tris-HCl (pH 7.5) plus 1 mM EDTA.

N-biotinyl N'-(aminooxy)acetyl hydrazide, a biotinylated aldehyde reactive probe (ARP).

Biotin hydrazide initially was used as a reagent for detecting AP sites in DNA. However, due to the reactivity of hydrazide, the signal to noise ratio was poor for DNA containing AP sites. In order to overcome this problem, biotin hydrazide was converted to an O-alkyl hydroxylamine derivative by the reaction of O-carboxymethyl hydroxylamine with biotin hydrazide in the presence of carbodiimide (Rosenberg, M.B. et al. (1986) J. Neurochem. 46:641-648). O-Carboxymethyl hydroxylamine-HCl (1.365 g) was dissolved in 125 ml of distilled water. EDAC (2.395g) was added, and the pH of the solution was adjusted with pyridine to between 4.0 and 5.0. Then 320 mg of biotin hydrazide was added and the reactive solution was incubated overnight at room temperature. Following incubation, the pH was adjusted to 7.0 by NaOH and the solution was extracted three times with an equal volume of chloroform ($CHCl_3$). The chloroform phase was discarded and the aqueous phase was evaporated to dryness under reduced pressure. The residue was extracted again with chloroform. The chloroform-extracted residue was then dissolved in 50 ml of distilled water, and the pH was adusted to neutrality.

50 g of AG1-X8 (Cl^- form) (Bio-Rad Laboratories, 1414 Harbour Way South, Richmond, CA 94804) was prepared by washing the sample resin twice with 200 ml of 2 M NaCl, and packing the washed resin into a column (2.5 x 50 cm). The resin was washed with four bed volumes of distilled water. Crude ARP solution

was loaded onto the column to remove unreacted O-carboxymethyl hydroxylamine. The first 50 ml of eluate was discarded and subsequent fractions of 50 ml were collected. Each of the fractions was tested for its ability to produce a colorimetric reading with calf thymus DNA containing AP sites previously bound to a microtiter plate (for colorimetric quantitation, see following section).

Active fractions were pooled (totaling approximately 300 ml) and used as a reagent for AP sites without further purification. The chemical structure of the ARP reagent is tentatively assigned as shown in Figure 1.

Microtiter plate method for assaying AP sites.

Immulon I microtiter plates, available from Dynatech Laboratories, 14340 Sully Field Circle, Chantilly VI 22021, were irradiated overnight with 40 watts from an unfiltered germicidal lamp (254 nm) at a distance of 30 cm from the source. The irradiated plates could be used for up to two months with no loss in DNA binding efficiency. 200 µl of calf thymus DNA containing AP sites (10 µg/ml) in phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4.7H_2O$, and 1.4 mM KH_2PO_4) was added to each well, and incubated at 37°C for 2-3 hrs. Alternatively, the microtiter plates were incubated at 4°C overnight to facilitate DNA binding. The plates were then washed four times with PBS-Tween buffer (PBS plus 0.5% Tween 20, Tween is available from Sigma Chemical Company, P.O. Box 14508, St. Louis, MO 63178. Next, 100 µl of ARP reagent was added and the plates were further incubated at 37°C

for an hour. The plate was washed four times with PBS-Tween to remove any unreacted ARP reagent, and 50 µl of horseradish peroxidase substrate (O-phenylenediamine) was added. Color development was stopped with 50 µl of 5N H_2SO_4 after an appropriate time interval and the absorbance at 490 nm was taken. Standard curves were determined with either f1 or calf thymus DNA containing known amounts of AP sites.

For determining the number of intermediary AP sites produced after the action of endonuclease III, 200 µl of calf thymus DNA containing different amounts of thymine glycol (10 µg/ml solution) was added to each of the wells of a microtiter plate as described above. After incubation at 37°C for two hours, the plate was washed four times with PBS-Tween, after which 100 µl of diluted endonuclease III was added, and the sample incubated at 37°C for 10 minutes. In this example, endonuclease III was diluted into 10 mM Tris-HCl, 1 mM EDTA and 0.1 M KCl at 37°C. The reaction was terminated by washing the microtiter plate four times with PBS-Tween after which 100 µl of ARP reagent was added and the number of AP sites determined as described above.

For determining the number of AP sites produced after the action of uracil glycosylase, a procedure was used which is similar to that described for endonuclease III. The procedure was modified by using uracil-containing f1 single stranded DNA, and uracil N-glycosylase was diluted appropriately into 10 mM Tris-HCl (pH 7.5), and 1 mM EDTA.

X-irradiation.

Calf thymus, ϕX RF, and f1 DNA were x-irradiated at a concentration of 30 μg/ml in 10 mM potassium phosphate buffer (pH 7.5) using a Philips x-ray generator (Philips Electronic Instruments, Mount Vernon, NY) with a beryllium window Machlett tube (Machlett Laboratories, 1063 Hope Street, Stamford, CT 06907) operated at 50 kVp and 2 mA (9.6 Gy per minute). The dose rate is routinely determined in the laboratory by a Fricke ferrous sulfate dosimetry (Fisher Scientific Company, Fair Lawn, NJ) and phage T4 survival.

For irradiation of E. coli cells, 20 ml of strain AB1157 cells (laboratory strain) in PBS (1 x 10^9 per ml) were inactivated by heating at 90°C for 2 minutes and quick cooled to 4°C. The dead cells were then diluted to a cell density of 2 x 10^8 per milliliter and x-irradiated at 10°C at the doses indicated in Figures 5A and 5B. Due to the length of the x-irradiation, heat inactivated E. coli cells were employed so as to minimize any possible post-irradiation repair of the x-ray-induced lesions. Irradiated cells were then collected by centrifugation at 5000 rpm for 10 minutes and the DNA was isolated as described by Ausubel et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York.

Methylmethanesulfonate treatment.

10 ml (5 x 10^8 cells) of actively growing E. coli (AB1157) cells in PBS buffer were treated with different concentration of methylmethanesulfonate (MMS) and incubated at 37°C for 30 minutes. The MMS

was then removed by centrifugation (2000 rpm, SS34) at 4°C, and the resulting cell pellet was washed once with 10 ml of PBS. The cell pellet was collected at 2000 rpm (SS34) and resuspended in 10 ml PBS. For each treatment, 5 ml of the washed cells were further incubated at 37°C for 1 hour, and 5 ml was centrifuged at 2000 rpm, with the resulting cell pellet frozen at -20°C. After one hour, the post-MMS treated cells were collected by configuration at 2000 rpm and the cell paste frozen at -20°C. The frozen cell paste was later thawed and the DNA purified according to the published procedures of Ausubel et al. noted above.

RESULTS

Specificity and sensitivity of the ARP assay.

The ARP is a biotin-tagged derivative of O-carboxymethyl hydroxylamine, thus the chemical specificity of ARP was expected to be similar to that of O-alkyl hydroxylamine. To determine whether this was the case, DNA containing different unique base modifications was prepared. These included thymine glycols, urea residues, uracil, pyrimidine dimers, O-methyl hydroxylamine-modified AP sites, $NaBH_4$-reduced AP sites, as well as the simple AP sites produced by heat/acid depurination. Microtiter plates coated with calf thymus DNA containing each of these lesions were incubated with the ARP reagent. The amount of biotin bound in each of the wells was measured by an avidin/biotin complex (ABC) conjugated to horseradish peroxidase for colorimetric determination. The amount of signal measured was taken as a measure of the reactivity of the lesion with the ARP reagent.

Table I shows that ARP reacted only with the simple AP site. No reactivity of the ARP reagent was observed with other lesions even after prolonged incubation with the reagent (data not shown). Modification of the AP sites by sodium borohydride ($NaBH_4$) reduction or methoxyamine completely eliminated the signal.

TABLE I

Specificity of the ARP reagent

DNA substrates	O.D. 490 nm
Undamaged calf thymus DNA	0.00[a]
Uracil containing f1 DNA[b]	0.00
UV (254 nm) irradiated calf thymus DNA[c]	0.00
Calf thymus DNA containing 1 AP site per molecule	0.52
Calf thymus DNA containing reduced AP sites[d]	0.05
Calf thymus DNA containing methoxyamine modified AP sites[e]	0.00
Calf thymus DNA containing 10 thymine glycol per molecule	0.05

[a] control DNA

[b] DNA contained approximately 200 uracil per molecule

[c] DNA contained approximately 10 UV dimers per molecule

[d] DNA contained approximately 10 AP sites per molecule was reduced with $NaBH_4$

[e] DNA contained approximately 10 AP sites per molecule was modified with methoxyamine

As illustrated in Figure 2, the signal measured with the ABC-horseradish peroxidase assay was proportional to the number of AP sites in the DNA. In that case, f1 DNA containing AP sites was prepared by heat/acid depurination. Reduced AP sites were prepared by sodium borohydride reduction. The f1 DNA containing different numbers of AP sites (solid square) or reduced AP sites (solid triangle) was adsorbed to microtiter plates as described above. The DNA on the plate was then allowed to react with the ARP reagent and the amount of biotin on the plate was measured with avidin/biotin complexed to horseradish peroxidase as described above. Again, once the AP site was reduced with sodium borohydride, no reactivity with the ARP reagent was observed.

The ARP reagent was able to detect AP sites in both single and double stranded DNA, as illustrated in Figure 3. In that case, 200 µl of partially depurinated f1 single stranded DNA (solid circle) or calf thymus duplex DNA (solid square) were used to coat the wells of a microtiter plate at 4°C overnight as described above. The ARP reagent (100 µl) was added and the amount of biotin on the plate was measured with avidin/biotin-horseradish peroxidase as described above. The reaction was much faster with single stranded DNA than with double stranded DNA (data not shown).

Figure 3 further shows that the signal produced in depurinated f1 DNA is approximately twice the amount produced in depurinated calf thymus DNA at all depurinating time intervals. This is in agreement with the fact that the rate of depurination in single stranded DNA is two-fold higher than double stranded DNA.

AP sites formed in DNA after the action of DNA N-glycosylases are detectable by the ARP assay of the invention. Base excision repair enzymes, such as endonuclease III and uracil N-glycosylase, excise the damaged base from DNA, leaving behind an AP site containing aldehyde reactive sites. Figures 4A and 4B show that the ARP reagent exhibited no reaction with thymine glycol or uracil. In these cases, microtiter plates were precoated with f1 DNA containing 2% uracil (Figure 4A) or calf thymus DNA containing different amounts of thymine glycols (Figure 4B). For Figure 4A, uracil N-glycosylase was added and incubated at 37°C for 10 minutes. The enzyme was removed by washing the plates and the amount of ARP signal was then measured by the ARP assay. For Figure 4B, 100 µl of endonuclease III (solid square) was added in a buffer containing 0.1M KCl, 10 mM Tris-EDTA (pH 7.5) while to control plates, 100 µl of the same buffer was added without endonuclease III (solid triangle). The plates were incubated at 37°C for 30 minutes and the wells were washed. The amount of ARP signal was measured by the ARP assay.

Upon digestion with endonuclease III and uracil N-glycosylase respectively, the intermediary AP sites formed were readily detectable with the ARP reagent. Using uracil-containing f1 single stranded DNA, the number of AP sites detected was proportional to the amount of uracil N-glycosylase added (Figure 4A). Similarly, with calf thymus DNA containing thymine glycols, the ARP signal observed after endonuclease III digestion was linearly proportional to the original number of thymine glycols in the DNA (Figure 4B).

AP sites detected in x-irradiated DNA.

Ionizing radiation produces a wide spectrum of DNA damages including DNA strand breaks, base damages as well as AP sites. The results illustrated in Figures 5A, 5B and 5C were obtained by irradiating f1 single stranded (Figure 5A), ϕX-RF (Figure 5B) and calf thymus duplex (Figure 5C) DNA in phosphate buffer at a concentration of 30 µg/ml at the indicated x-ray dose. The DNA was then adsorbed onto the microtiter wells and the number of AP sites produced was determined by the ARP assay (solid square). For Figure 5A, the amount of thymine glycols (solid circle) was also measured by using anti-thymine glycol antibody following the procedures described by Hayes et al. (1988) J. Mol. Biol. 201:239-246.

Figure 5A shows that when f1 single stranded DNA was exposed to ionizing radiation, the number of ARP sites detected was proportional to the x-ray dose. Similarly, production of ARP sites in ϕX RF (Figure 5B) and calf thymus DNA (Figure 5C) was proportional to increasing doses of x-rays. The estimated detectable limit of ARP sites was obtained with an x-ray dose of approximately 50 rad. The rate of production of thymine glycols by x-rays, as detected by an anti-thymine glycol antibody, was significantly less than the rate of production of ARP sites. We estimated that at one Krad (10 Gray), seven ARP sites were produced as compared to one thymine glycol per f1 DNA molecule.

ARP sites in DNA extracted from X-irradiated E. coli cells.

When actively growing wild type E. coli cells (starved at 37°C for 30 minutes in PBS buffer) were x-irradiated, little or no detectable ARP signal was observed, suggesting that the lesions may have been repaired during processing of the irradiated cells to obtain the DNA sample. To circumvent this problem, E. coli cells (AB1157, 2 x 10^9 per ml in PBS buffer) were preheated at 90°C for two minutes to inactivate the repair enzymes, then irradiated with the appropriate dose of x-rays. Figure 6 shows the rate of production of ARP sites and thymine glycols in heat-inactivated E. coli cells. ARP sites were detectable at about one krad, while thymine glycol was barely detectable even at one Krad of x-irradiation. The rate of production of ARP sites was proportional to the x-ray dose.

Production and Repair of AP sites in MMS-treated E. coli cells.

Methylmethanesulfonate (MMS) alkylates both adenine and guanine at multiple positions. Upon alkylation of the purines, the N-glycosylic bond becomes more labile, leading to an increased production of AP sites. Furthermore, repair of these alkylated DNA bases by N-glycosylases, such as 3-methyladenine DNA N-glycosylase, should lead to a further increase in the production of AP sites. To determine the number of AP sites formed in DNA after MMS treatment of E. coli cells, DNA was isolated from cells treated with different concentrations of MMS for 30 minutes at 37°C. MMS was then removed by

centrifugation followed by washing with PBS buffer. Figure 7 illustrates the results, wherein DNA was extracted from half of the cell pellet (solid square), while the other half (solid triangle) was further incubated at 37°C for an additional hour to allow for repair. Cells were then collected by centrifugation and the DNA extracted and the number of AP sites formed determined by the ARP assay of this invention.

Figure 7 shows that the steady state production and subsequent repair of AP sites in E. coli cells was detectable by the ARP assay of this invention. At 25 μM MMS, where the survival of E. coli cells was at about 80% (data not shown), a strong signal was observed. Since a five-fold lower signal can easily be measured, by extrapolation, ARP sites should be detectable in DNA from cells treated with 5 μM MMS. At this concentration, little or no killing of cells was observed (data not shown).

EXAMPLE 2

Materials and Methods

Chemicals.

Deoxyribose 5-phosphate, deoxyribose, ribose, O-4-nitrobenzyl hydroxylamine, and O-benzyl hydroxlamine were purchased from Aldrich Chemical Company, 1001 West Saint Paul Avenue, Milwaukee, WI 53233.

Preparation of haptens and antigen.

5'phosphodeoxyribosyl O-4-nitrobenzyl hydroxlyamine (NBHA-dRp) was prepared by incubating O-4-nitrobenzyl hydroxylamine (NBHA, 10 mM solution at pH 6-7) with 15 mM deoxyribose 5-phosphate overnight at room temperature. The solution was then used for hapten inhibition studies without further purification. Solutions of 5'phosphodeoxyribosyl O-benzyl hydroxylamine, deoxyriboxyl O-4-nitrobenzl hydroxylamine and deoxyribosyl O-benzyl hydroxylamine, each at 10 mM, were prepared in a similar manner without further purification. The reaction of O-alkyl hydroxylamine with aldehyde is known to be quantitative. NBHA-dRp was coupled to BSA through the 5'-phosphate by the carbodiimide procedure described by Rosenberg et al. (1986) J. Neurochem. 46:641-648.

Preparation of the monoclonal antibody.

Immunization of BALB/c mice was performed according to Holmdahl et al. (1985) J. Immunol. Methods 83:379-384. The NBHA-dRp-BSA conjugate was dissolved in 0.05 mM sodium phosphate buffer (pH 7.5) to a concentration of 2 mg/ml. The solution was emulsified with an equal volume of complete Freund's adjuvant, available from GIBCO BRL, P.O. Box 68, Grand Island, NY 14072. A 50 µl aliquot of the emulsion was injected subcutaneously into each hind foot pad. Nine days after immunization, lymphocytes were isolated from the draining popliteal lymph nodes and fused with mouse myeloma P3 x 63-AG63-653.

Screening of antibody production was done by ELISA assay using NBHA-dRp-RSA conjugate as the antigen. Ten out of 440 wells tested positive. The positive supernatants were subsequently assayed by ELISA for binding efficiency to calf thymus DNA containing NBHA-modified AP sites (200 sites per 100,000 bp). Of the ten positive lines, one reacted equally well with undamaged calf thymus DNA as with NBHA-modified AP DNA, while two clones bound specifically to calf thymus DNA containing NBHA-modified AP sites. The latter were subcloned twice by limited dilution. Four monoclonal antibody clones were obtained and one of them, 3-12G-12H-12H, was chosen for further study. The antibody was IgM, kappa as determined by Ouchterlony gel diffusion.

Preparation of NBHA-modified depurinated DNA.

Calf thymus DNA was alkylated with 0.3M MMS at 37°C for one hour. After exhaustive dialysis against Tris-HCl (pH 7.5), 1 mM EDTA, the DNA was partially depurinated by heating at 50°C for an appropriate amount of time to produce DNA containing approximately 200 AP sites per 10,000 bp. DNA containing 2 to 16 AP sites per 10,000 bp was prepared according to Lindahl and Nyberg (1972) Biochemistry 11:3610-3618. Calf thymus DNA was dissolved in 0.1M NaCl, 0.01M sodium citrate (pH 5.0, 100 µg/ml) at 70°C for 15, 30, 60 or 120 minutes to produce approximately 2, 4, 8 or 16 AP sites per 10,000 bp, respectively.

The depurinated DNA preparations were modified by reaction with 5 mM NBHA in phosphate buffer (pH 7.2) at room temperature for two hours. The product

was then ethanol precipitated in the presence of 2.5M sodium acetate. The precipitated DNA was washed once with 75% ethanol, then precipitated again to remove traces of NBHA. The precipitated DNA was re-dissolved in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA. Undamaged calf thymus DNA was treated similarly with NBHA to serve as the control.

Direct binding assay.

Direct binding assays were carried out as previously described by Haugen et al. (1981) Proc. Nat. Acad. Sci. U.S.A. 78:4124. Polystryene 96 wells microtiter plates, available from Corning 25855, Corning Incorporated Science Products Division, Corning NY 14831, or Immulon 1, available from Dynatech Laboratories, Inc., 14340 Sully Field Circle, Chantilly, VI 22021, were UV-irradiated overnight and coated with 200 µl of NBHA-modified calf thymus DNA in phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, $7H_20$, and 1.4 mM KH_2PO_4) for 2 hours at 37°C. After blocking the plate with 1% horse serum in PBS-Tween (PBS plus 0.5% Tween) for one hour at room temperature, the monoclonal antibody, diluted 250 fold in PBS-Tween, was added to the coated wells and incubated at 37°C for 2 hours. After washing with PBS-Tween, bound antibody was detected with horseradish peroxidase-conjugate goat-anti-mouse IgG + IgM (1:3000 dilution, TAGO, Inc., Burlingame, CA 94010) by the addition of hydrogen peroxide (H_2O_2) and O-phenylenediamine. Absorbance at 490 nm was then taken after stopping the reaction with 50 µl 8N H_2SO_4 at appropriate times.

Inhibition assay.

Competitive enzyme immunoassay by haptens of calf thymus DNA containing NBHA-modified AP sites was performed with microtiter plates which were precoated with calf thymus DNA containing NBHA-modified AP sites (200 sites per 10,000 bp, 270 ng/ml) according to procedures previously described by Hubbard et al. (1989) Radiat. Res. 118:257-268.

Results

Selection of monoclonal antibody clones for study.

Four monoclonal antibody clones specific for NBHA-modified AP sites were obtained and examined for their reactivities with NBHA-modified AP sites using a competitive enzyme immunoassay. Microtiter wells were preadsorbed with calf thymus DNA containing 40 NBHA-modified AP sites per 10,000 bp. In addition, the NBHA-modified DNA was diluted at various concentrations and 100 µl of aliquots were incubated at 37°C for one hour with 100 µl of each of the antibody preparations. Aliquots of the incubated DNA samples were removed and added to the individual wells containing the adsorbed NBHA-modified AP DNA. ELISA assays were performed using horseradish peroxidase as the indicator enzyme. The results are shown in Table II.

TABLE II

% Inhibition of Antibody Reactivity

Monoclonal Clone	Inhibitor Concentration (µg / ml)						
	5	2.5	1.25	0.625	0.312	0.16	0.08
3-12G-12H-12H	--	64	55	44	33	31	21
4-8H-2D-12G	52	51	48	43	28	21	12
4-8H-3H-6H	46	48	41	43	15	12	0
4-8H-12H-9A	29	23	14	13	12	0	0

As shown in Table II, clone 3-12G-12H-12H was consistently more reactive with competing NBHA-modified AP sites at all DNA concentrations tested. At 0.08 µg/ml of competing NBHA-modified AP sites in DNA, 21% inhibition was observed for clone 3-12G-12H-12H as compared to 12%, 0% and 0% for clones 4-8H-2D-12G, 4-8H-3H-6H and 4-8H-12H-9A, respectively. Based on these data, clone 3-12G-12H-12H was selected for further characterization.

Antibody specificity.

Various haptens were used as competitors of the binding of antibody to NBHA-modified AP sites in DNA. The results are shown in Figure 8 and Table III.

TABLE III

Hapten	IC_{50}
5'Phosphoribosyl-NBHA	0.3 μM
Deoxyribosyl-NBHA	5.0 μM
Ribosyl-NBHA	5.0 μM
NBHA	7.0 μM
5'Phosphoribosyl-BHA	No inhibit. at 1mM
Deoxyribosyl-BHA	No inhibit. at 1mM
Ribosyl-BHA	No inhibit. at 1mM
BHA	No inhibit. at 1mM
Deoxyribose 5-phosphate	No inhibit. at 10mM
Ribose	No inhibit. at 10mM
Deoxyribose	No inhibit. at 10mM

IC_{50} is the concentration of inhibitor to effect a fifty percent inhibition of the antibody response. In this case, UV-irradiated microtiter plates were previously adsorbed with calf thymus DNA (200 ng/well) containing 200 NBHA-modified AP sites per 10,000 bp. Serial dilutions of haptens were incubated with 1:1000 dilution of antibody ($(NH_4)_2SO_4$-precipitated supernatant), and 100 μl of the hapten-antibody reaction mix was added to each well. After washing with PBS-Tween, the amount of

bound antibody was determined with horseradish peroxidase-conjugated goat anti-mouse IgM (1:3000 dilution) in PBS-Tween. Haptens used in the experiments were 5'phosphoribosyl-NBHA (solid square), deoxyribosyl-NBHA (solid circle), ribosyl-NBHA (open circle), and NBHA (open square).

As shown in Table III and Figure 8, 5'deoxyribosyl-NBHA was approximately 20 times more effective as a competitor (IC_{50} = 0.3 μM) than NBHA, deoxyribosyl-NBHA and ribosyl-NBHA, which gave IC_{50} values of 5μm, 5μm and 7 μM, respectively. Haptens lacking a nitro group, such as benzyl hydroxylamine, ribosyl, deoxyribosyl, and 5'-phosphoribosylbenzyl hydroxylamine, showed no inhibition up to 1 mM. Deoxyribose, ribose, and deoxyribose 5-phosphate showed no inhibition even up to 10 mM.

When DNA containing NBHA-modified AP sites was used to inhibit antibody reactivity, inhibition occurred with DNA containing as few as two NBHA-modified AP sites per 10,000 bp, at a concentration of 50 μg/ml of DNA. Inhibition of anti-NBHA antibody reactivity by calf thymus DNA containing NBHA-modifed AP sites was tested using UV-irradiated microtiter plates which were previously adsorbed with 200 μl of 160 ng/ml of calf thymus DNA containing NBHA-modified AP sites (200 sites per 10,000 bp). Dilutions of calf thymus DNA containing various amounts of NBHA-modified AP sites were then incubated with a 1:1000 dilution of the antibody preparation.

Figure 9A illustrates the results of calf thymus DNA containing 0 (solid triangle), 2 (open circle), 4 (open square), 8 (solid circle), or 16 (solid square)

NBHA-modified AP sites per 10,000 bp used to adsorb to the microtiter plate. Similarly, calf thymus DNA containing 25 (open circle), 50 (open square), 100 (closed circle), or 200 (solid square) NBHA-modified AP sites per 10,000 bp were used to obtain the results illustrated in Figure 9B. In both cases, 100 µl of the NBHA-treated DNA/antibody reactions were added to each well, and the bound antibody was detected with horseradish peroxidase-conjugated anti-mouse IgM diluted 1:3000 in PBS-Tween as described above.

The results reported in Figure 10 demonstrate that inhibition was proportional to the total number of AP sites regardless of the extent of depurination of DNA used. The symbols used in Figure 10 are the same as those used for Figures 9A and 9B, and the data were calculated from test results illustrated in Figure 9B.

The specificity of the anti-NBHA-dRp antibody of this invention was also tested. UV-irradiated microtiter plates were coated with 200 µl of DNA containing AP sites (solid circle), reduced AP sites (open square), NBHA-modified AP sites (solid square), and thymine glycols (open triangle). After blocking the microtiter plate with 1% calf serum in PBS-Tween, the monoclonal antibody (1:250 dilution of the ascites) was added to the wells and the bound antibody was detected with horseradish peroxidase-conjugated goat anti-mouse IgG + IgM at a dilution of 1:3000 as described above. The results, reported in Figure 11 demonstrate, using a direct ELISA assay with native calf thymus DNA, that the

antibody reacted only with NBHA-modified AP sites but not with DNA containing AP sites, reduced AP sites or thymine glycols.

The effect of DNA denaturation on antibody reactivity, and the reactivity of the antibody of this invention toward single stranded and duplex DNA were also tested. The results are illustrated in Figures 12A and 12B. To test the effect of DNA denaturation on antibody reactivity, native (solid square) or heat denatured (solid circle) calf thymus DNA containing NBHA-modified AP sites adsorbed to microtiter plates. Bound antibody was detected with the horseradish peroxidase-conjugated goat anti-mouse IgM secondary antibody as described above. As illustrated in Figure 12A, the immunoreactivity of the antibody NBHA-modified AP sites was about two fold higher with heat denatured calf thymus DNA as compared to native DNA.

This preference was substantiated by comparing single stranded f1 (solid square) and double stranded calf thymus (solid circle) DNA, as illustrated in Figure 12B. In this case, DNA containing NBHA-modified AP sites (10 mg/ml) were adsorbed on the UV-irradiated microtiter plate. The amount of detectable NBHA residues were measured with anti-NBHA-dRp monoclonal antibody, and the signal obtained with NBHA-modified f1 DNA was two-fold higher than NBHA-modified calf thymus DNA. This suggests that the antibody of this invention reacts more effectively with the lesion in single stranded DNA than in double stranded DNA.

Antibody sensitivity.

Calf thymus DNA containing two NBHA-modified AP sites per 10,000 bp gave an optical density reading of 0.65 after a two minute reaction time. Therefore, by extrapolation, the antibody should easily detect one AP site per 10,000 nucleotides. Since only approximately 70 to 100 ng of DNA was absorbed on each well, the antibody thus can detect approximately 10 femtomoles of AP sites.

Reaction of the antibody with damaged f1 DNA.

To determine the number of AP sites detected by the antibody of this invention, f1 single stranded DNA was x-irradiated in phosphate buffer at a concentration of 30 µg/ml at the indicated x-ray dosage. The x-irradiated DNA was treated with 10 mM NBHA for three hours at room temperature, and excess NBHA was removed by ethanol precipitation. The DNA was then adsorbed onto microtiter wells and the number of anti-NBHA reactive sites were determined. As illustrated in Figure 13, the number of antibody reactive sites detected was proportional to the x-ray dose. The estimated detectable limit of the antibody reactive sites was obtained with an x-ray dose of approximately 50 rad.

Equivalents

Those skilled in the art will recognize, or be able to ascertain, by no more than routine experimentation, many equivalents of the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

CLAIMS

1. A method of quantifying aldehyde-containing lesions in nucleic acid, comprising the steps of:
   a. reacting the nucleic acid with a reagent that specifically reacts with the aldehyde group of the lesion in the nucleic acid to form a detectable complex of the reagent and the nucleic acid; and
   b. detecting the amount of reagent complexed to the nucleic acid as an indication of the quantity of aldehyde-containing lesions in the nucleic acid.

2. The method of claim 1, wherein the aldehyde-containing lesion is an abasic site.

3. The method of claim 1, wherein the nucleic acid is DNA.

4. The method of claim 1, wherein the aldehyde-specific reagent is an O-alkyl hydroxylamine.

5. The method of claim 4, wherein the O-alkyl hydroxylamine is O-4-nitrobenzylhydroxylamine.

6. The method of claim 1, wherein the aldehyde-specific reagent is detectably labeled and the amount of reagent complexed with the nucleic acid is determined by detecting the amount of label associated with the nucleic acid.

7. A method of claim 1, wherein the aldehyde-specific reagent is biotinylated and the amount of biotinylated reagent complexed with the nucleic acid is determined by detecting the activity of a biotinylated enzyme bound to the biotinylated reagent through avidin mediated linkage.

8. A method of claim 1, wherein the amount of aldehyde-specific reagent complexed with nucleic acid is determined by i) reacting the nucleic acid with an antibody that specifically binds to the complex of the reagent and the nucleic acid; and ii) detecting the amount of antibody bound to the nucleic acid.

9. A method of quantifying abasic sites in DNA, comprising the steps of:

a. reacting the DNA with a biotinylated, aldehyde-specific O-alkyl hydroxylamine to form a complex of the O-alkyl hydroxylamine and the abasic sites of the DNA;

b. reacting the DNA with avidin to bind the biotinylated O-alkyl hydroxylamine complexed to the abasic sites;

c. reacting the avidin-bound DNA with a biotinylated enzyme to enzymatically label the abasic sites; and

d. measuring activity of the enzyme as indicative of the amount of abasic sites in the DNA.

10. The method of claim 9, wherein the biotinylated, aldehyde-specific O-alkyl hydroxylamine is N'-biotinyl N'(aminooxy)acetyl hydrazide.

11. The method of claim 9, wherein the biotinylated enzyme is horseradish peroxidase.

12. The method of claim 11, wherein enzymatic activity is measured colorimetrically.

13. A method of quantifying abasic sites in DNA, comprising the steps of:

a. reacting the DNA with an aldehyde-specific O-alkyl hydroxylamine to form a complex of the O-alkyl hydroxylamine and the abasic sites of the DNA;

b. reacting the DNA with an antibody that binds specifically to the complex of the O-alkyl hydroxylamine and the abasic sites; and

c. detecting the amount of complex-specific antibody bound to the DNA as indicative of the amount of the abasic sites in the DNA.

14. The method of claim 13, wherein the aldehyde-specific O-alkyl hydroxylamine is O-4-nitrobenzyl hydroxylamine.

15. The method of claim 13, wherein the antibody is monoclonal.

16. The method of claim 13, wherein the amount of complex-specific antibody bound to the DNA is detected by i) reacting the DNA with a detectably labeled second antibody against the complex-specific antibody; and ii) measuring the activity of the label associated with the DNA as indicative of the amount of complex-specific antibody bound to the DNA.

17. The method of claim 16, wherein the antibody is labeled with an enzyme.

18. A detectably labeled reagent which specifically complexes to the aldehyde group of an abasic site of a nucleic acid.

19. A reagent of claim 18, comprising a biotinylated O-alkyl hydroxylamine, N-biotinyl N'-(aminooxy)acetyl hydrazide.

20. A reagent of claim 19, comprising O-4-nitrobenzyl hydroxylamine.

21. A reagent which specifically complexes with the aldehyde group of an abasic site of DNA and forms an epitope for antibody binding to the complex.

22. A reagent of claim 21, comprising O-4-nitrobenzylhydroxylamine.

23. A monoclonal antibody specific for the complex of a reagent which specifically reacts with aldehyde groups in nucleic acid.

24. A monoclonal antibody of claim 23 wherein said reagent includes an O-alkyl hydroxylamine.

25. A monoclonal antibody of claim 24, wherein the hydroxylamine is O-4-nitrobenzylhydroxylamine.

26. A kit for performing a colorimetric assay to quantify aldehyde-containing lesions of DNA, comprising:
a. a biotinylated reagent specific for the aldehyde group of the lesion of DNA;
b. avidin;
c. biotinylated enzyme, and
d. chromogenic substrate for the enzyme.

27. The kit of claim 26, wherein the aldehyde-specific reagent is a biotinylated O-alkyl hydroxylamine.

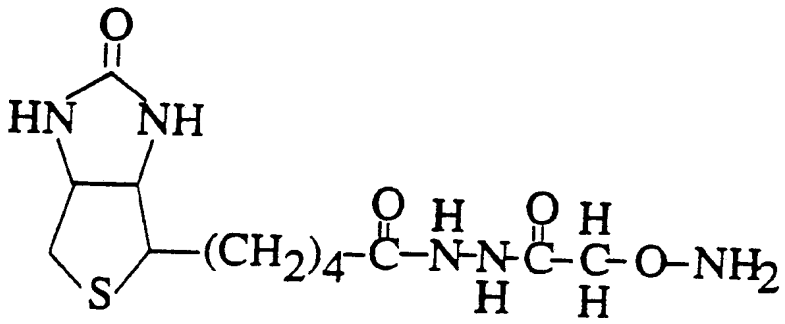
28. A kit of claim 26, wherein the enzyme is horseradish peroxidase.

29. A kit for performing an assay to quantify abasic sites in DNA, comprising:
a. a reagent which specifically complexes to an aldehyde group of an abasic site of DNA and forms an epitope for antibody binding to the complex; and
b. an antibody specific for the epitope.

30. The kit of claim 29, wherein the aldehyde-specific reagent is an O-alkyl hydroxylamine.

31. The kit of claim 39, wherein the O-alkyl hydroxylamine is O-4-nitrobenzylhydroxylamine.

Figure 1



IDT Exhibit 1020.1015

Figure 2

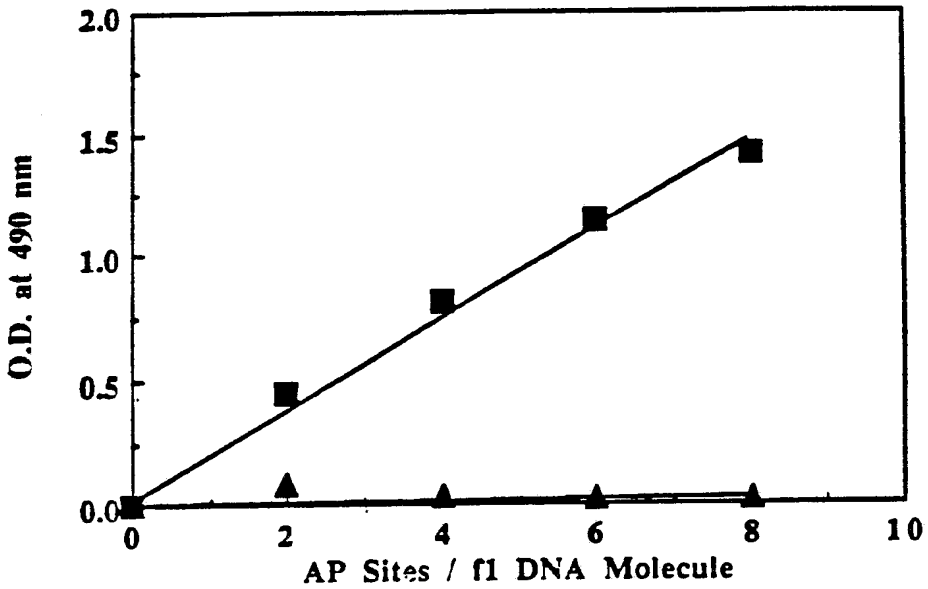


Figure 3

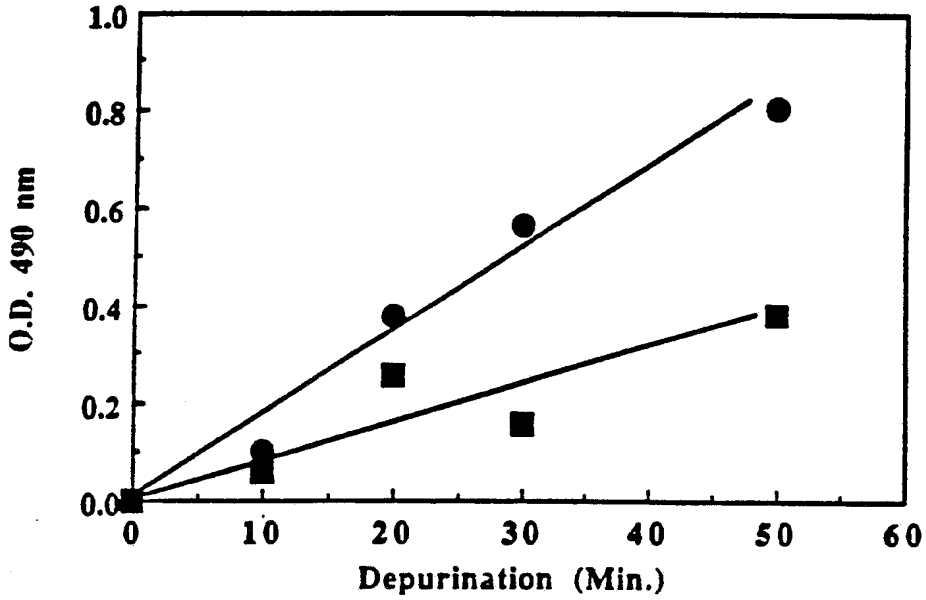


Figure 4A

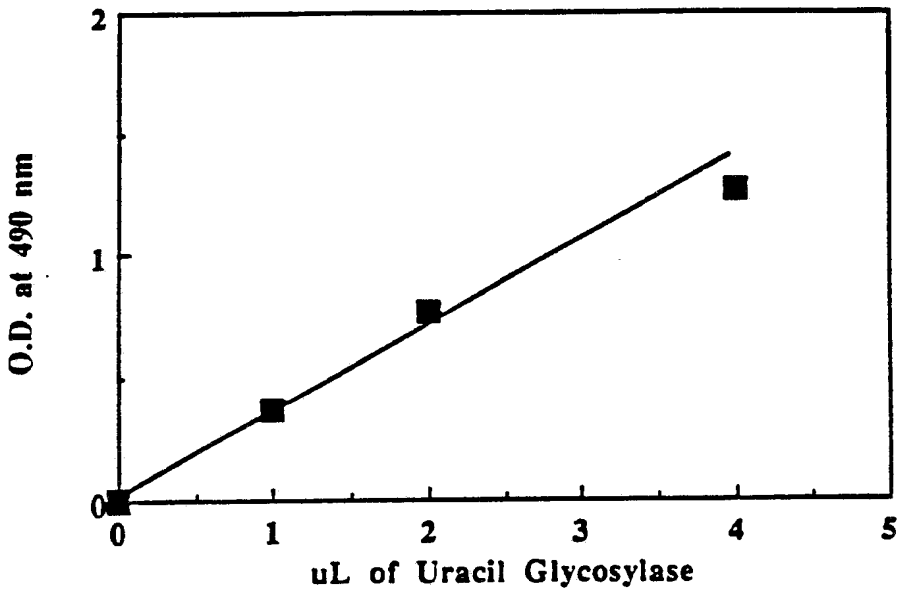


Figure 4B

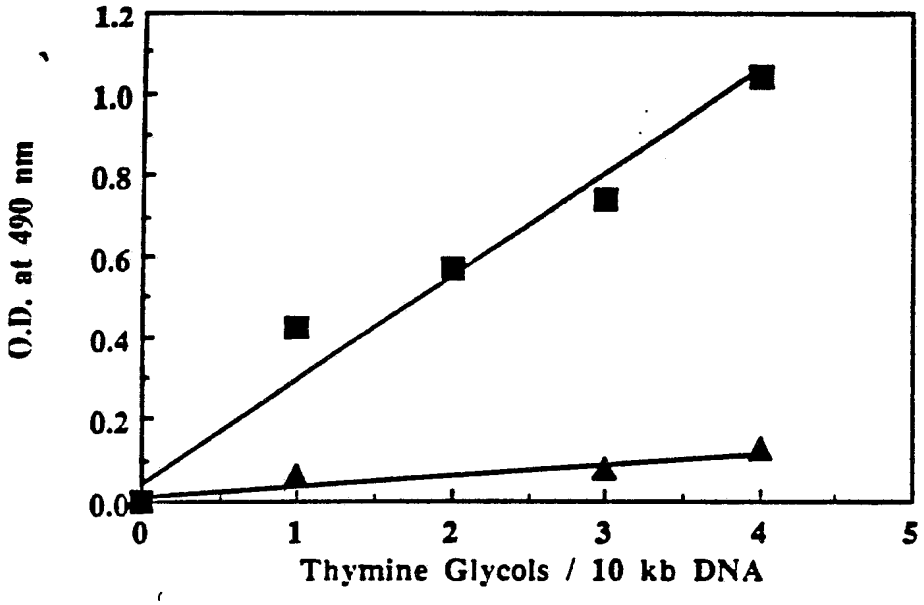

1.2
1.0
0.8
0.6
0.4
0.2
0.0
O.D. at 490 nm
0
1
2
3
4
5
Thymine Glycols / 10 kb DNA

Figure 5A

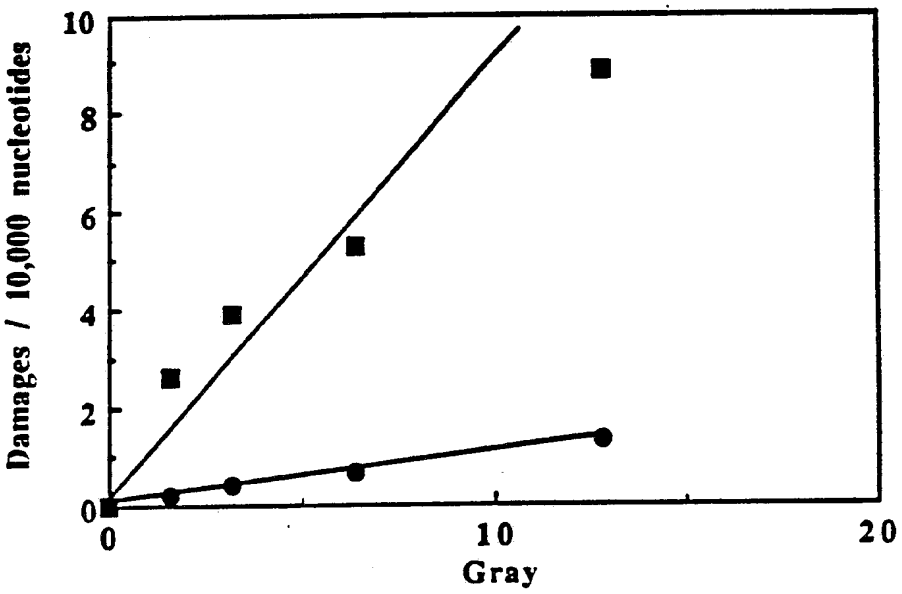


Figure 5B

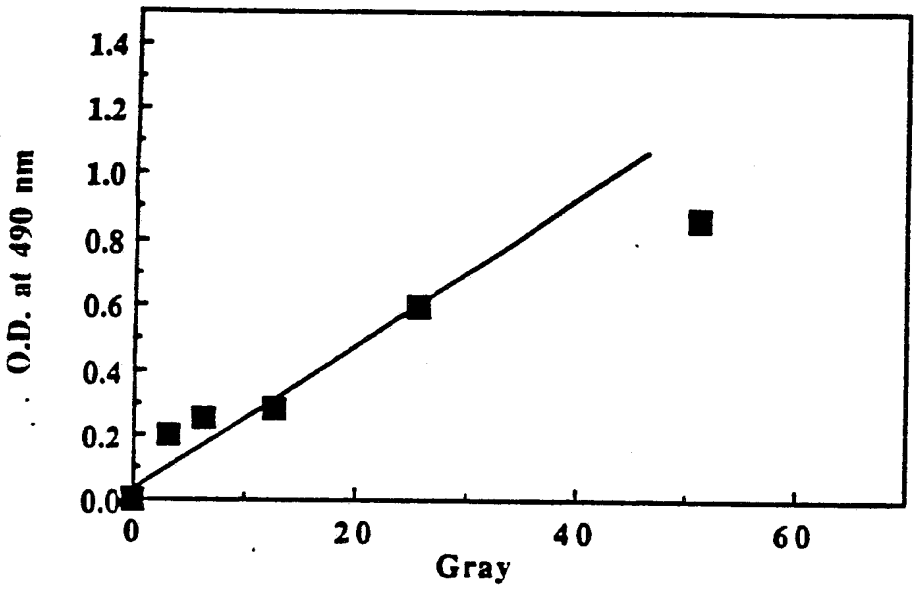


Figure 5C

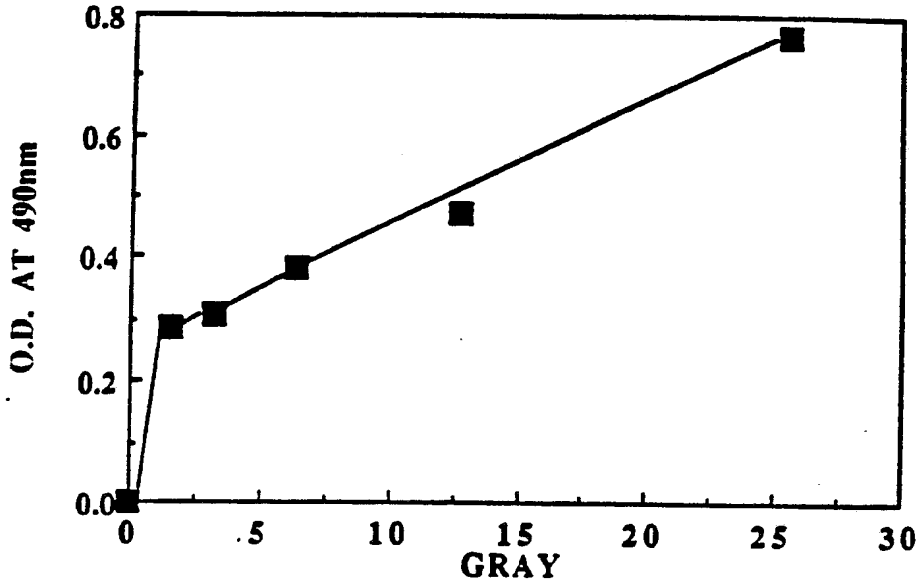


Figure 6

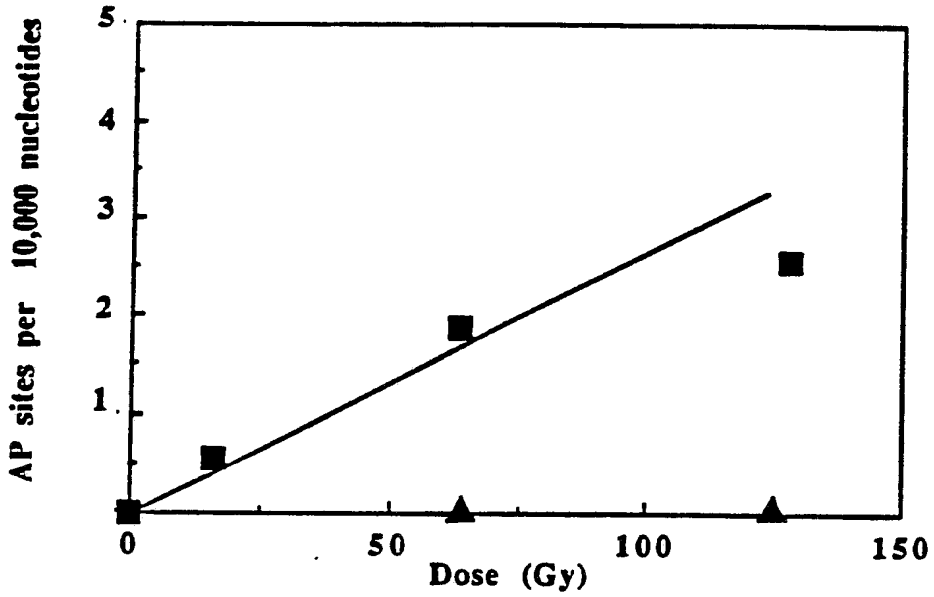


Figure 7

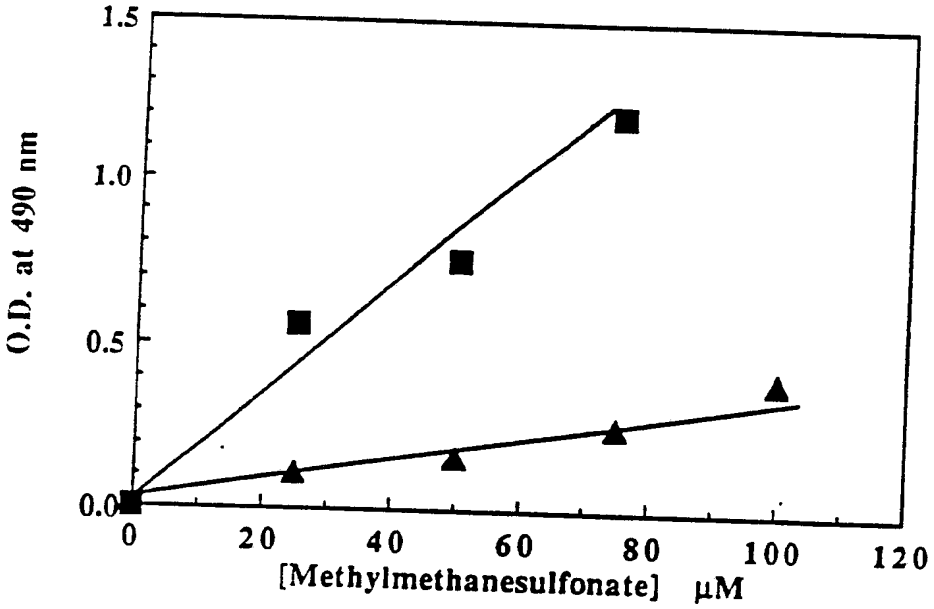

1.5
1.0
0.5
0.0
O.D. at 490 nm
0
20
40
60
80
100
120
[Methylmethanesulfonate] µM

Figure 8

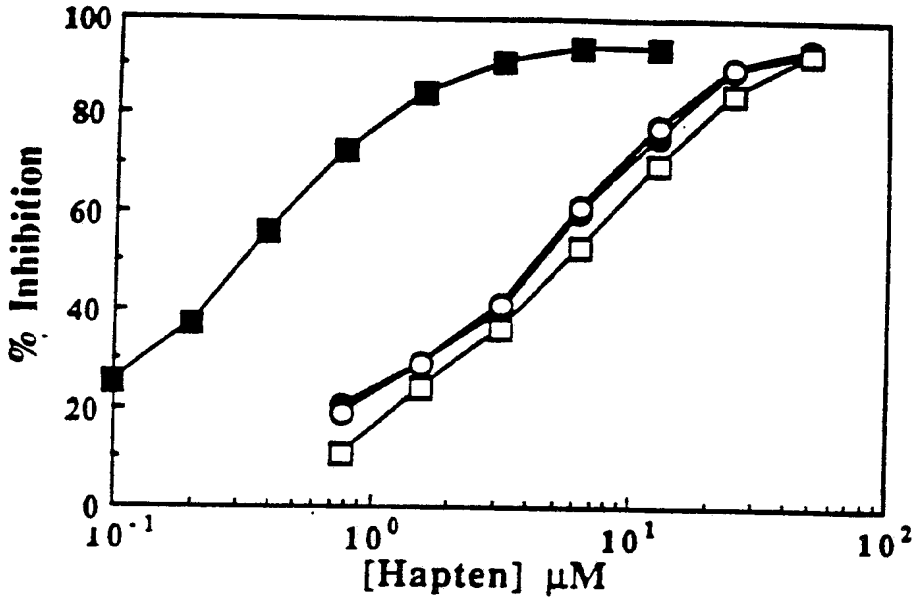


Figure 9A

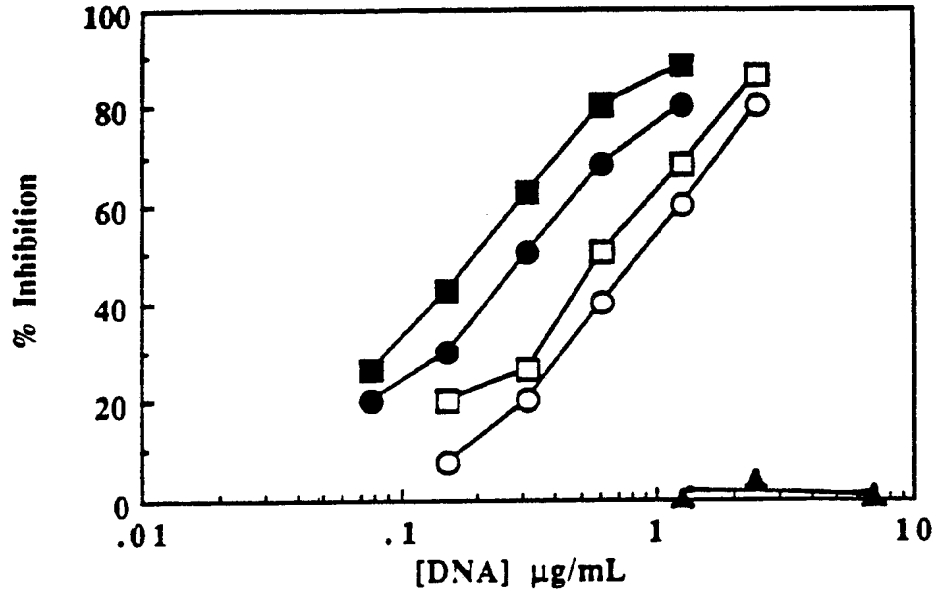


Figure 9B

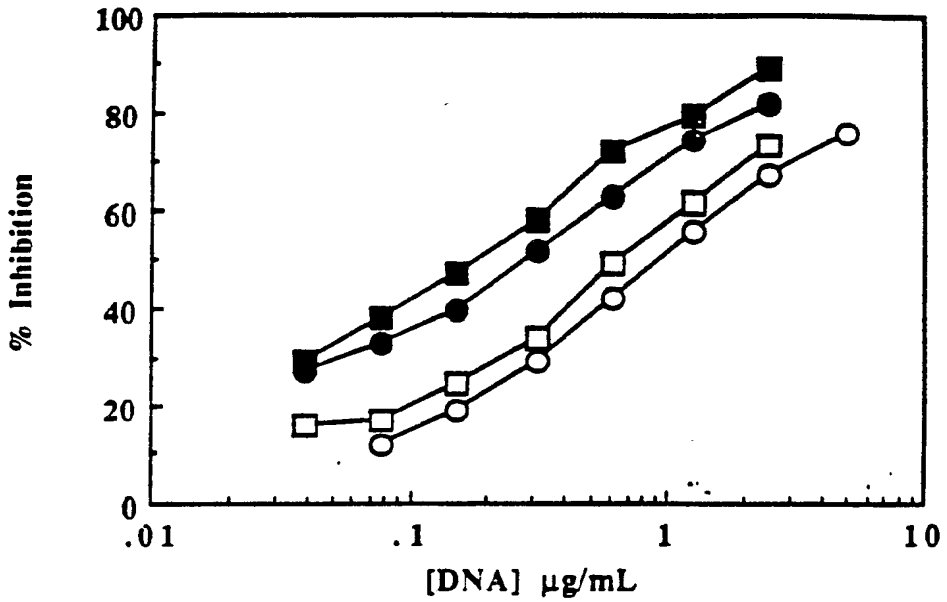


Figure 10

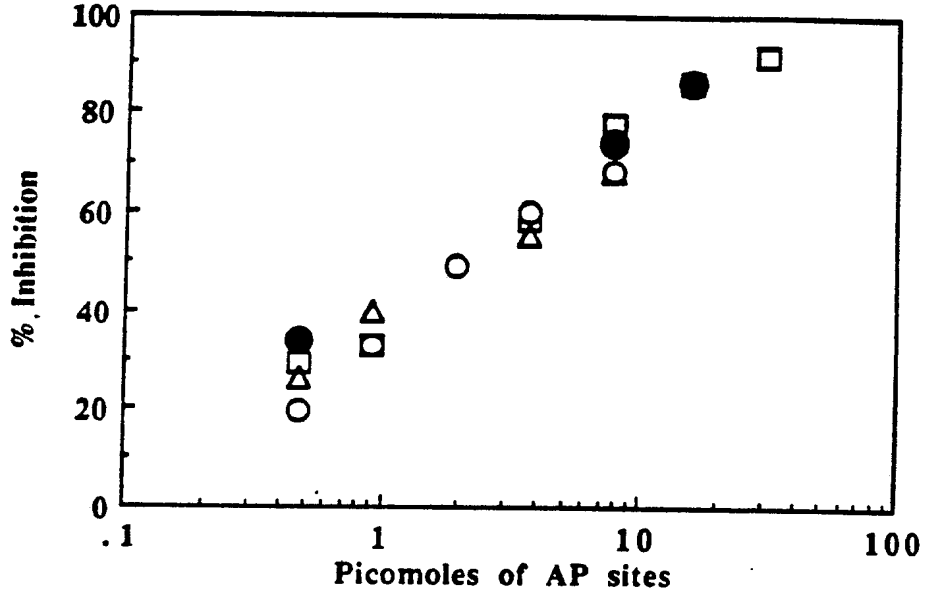


Figure 11

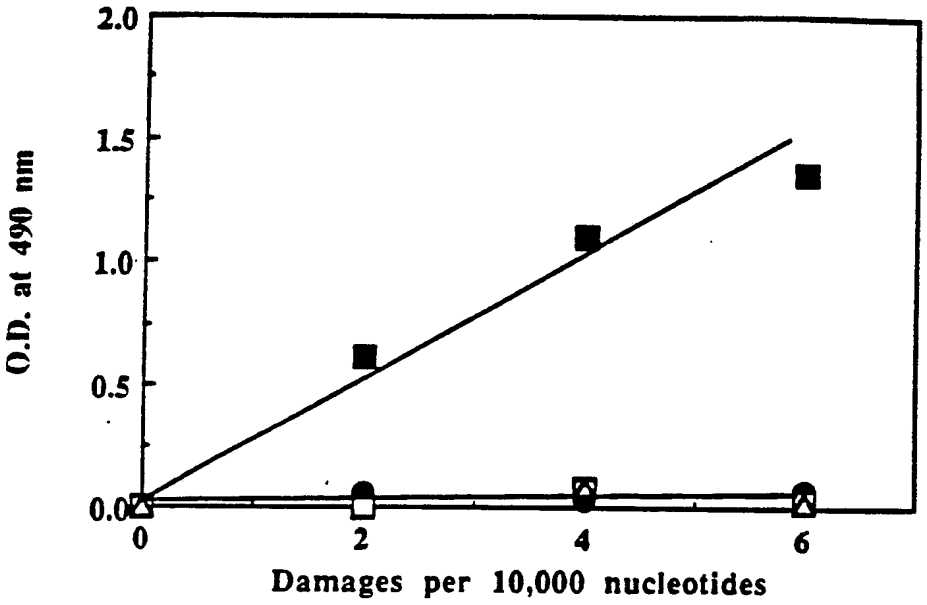


Figure 12A

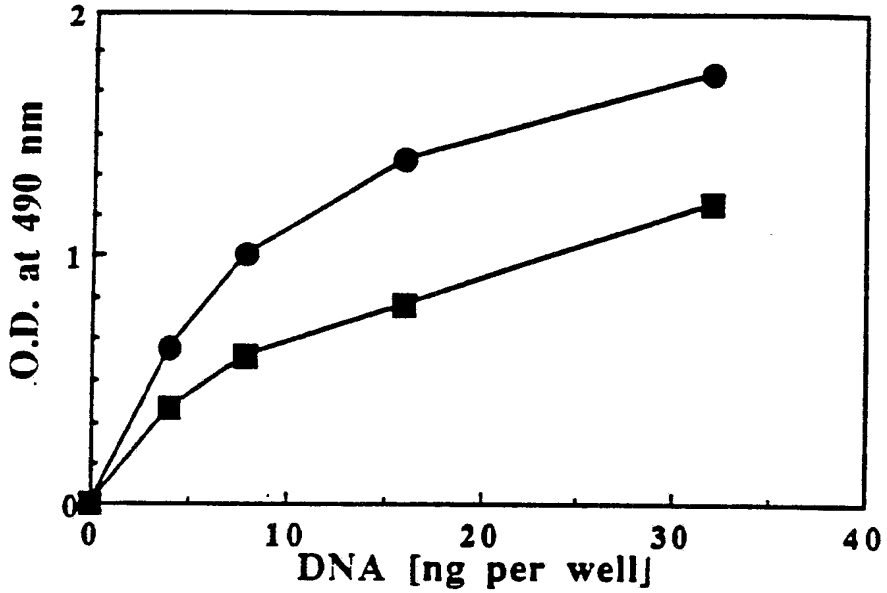


Figure 12B

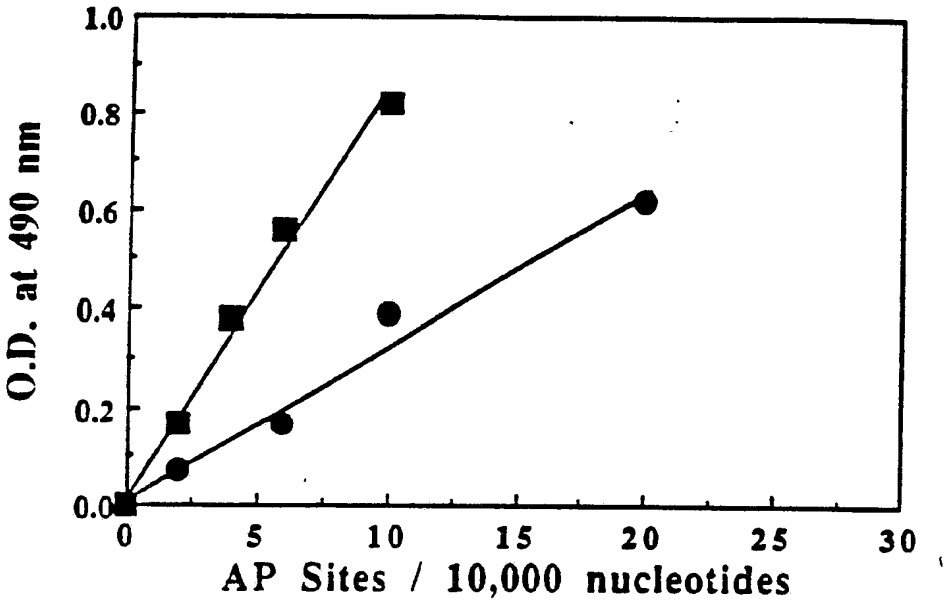
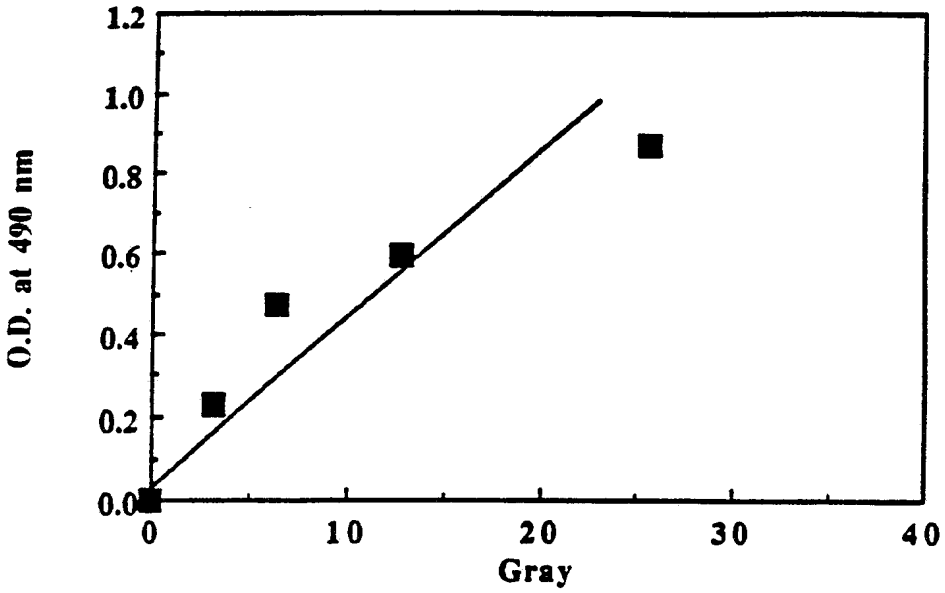

1.0
0.8
0.6
0.4
0.2
0.0
O.D. at 490 nm
0
5
10
15
20
25
30
AP Sites / 10,000 nucleotides

Figure 13



INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/07813

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C12Q 1/68; C07H 15/12; C12N 15/00
US CL : 435/6; 536/27; 935/77,78

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	435/6; 536/27; 935/77,78

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

CAS ONLINE

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
X	Biochimica et Biophysica Acta, Volume 740, issued 1983, Talpaert-Borle et al, "Reaction of Apurinic/Apyrimidinic Sites with [14C]Methoxyamine - A Method for the Quantitative Assay of AP Sites in DNA", pages 410-416, see entire document.	1-4, 6, 18, 21
X/Y	US, A 4,863,847 (Talpaert-Borle et al) 05 September 1989, see entire document.	1-4, 6, 18, 21/5, 7-17, 19, 23-31
X	Sigma Chemical Company Catalog, issued February 1986, see page 882.	20, 22
Y	Analytical Biochemistry, Volume 169, issued February 1988, Matthews et al, "Analytical Strategies for the Use of DNA Probes" pages 1-25,. see entire document.	7-17, 19, 23-31

* Special categories of cited documents:[15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
30 JANUARY 1992	12 FEB 1992
International Searching Authority[1]	**Signature of Authorized Officer[20]**
ISA/US	Mindy B. Fleisher

Form PCT/ISA/210 (second sheet)(May 1986) ß

IDT Exhibit 1020.1033

Espacenet

Bibliographic data: WO2010003153 (A2) — 2010-01-07

METHYLATION ANALYSIS OF MATE PAIRS

Inventor(s): MCKERNAN KEVIN J [US]; SCHROEDER BENJAMIN G [US]; BOYD VICTORIA L [US] ± (MCKERNAN, KEVIN, J, ; SCHROEDER, BENJAMIN, G, ; BOYD, VICTORIA, L)

Applicant(s): LIFE TECHNOLOGIES CORP [US]; MCKERNAN KEVIN J [US]; SCHROEDER BENJAMIN G [US]; BOYD VICTORIA L [US] ± (LIFE TECHNOLOGIES CORPORATION, ; MCKERNAN, KEVIN, J, ; SCHROEDER, BENJAMIN, G, ; BOYD, VICTORIA, L)

Classification: - international: *C12N15/11; C12Q1/68*
- cooperative: **C12Q1/6827**; **C12Q1/686** more

Application number: WO2009US49724 20090706 Global Dossier

Priority number(s): US20080133891P 20080703 ; US20090149976P 20090204

Also published as: WO2010003153 (A3) US2010120034 (A1)

Abstract of WO2010003153 (A2)

Various embodiments of the present teachings relate to methods for the methylation analysis of nucleic acids. The subject methods include methods that result in the preparation of mate-pair libraries suitable for highly multiplexed DNA sequencing. Embodiments include methods of preparing mate-pair libraries comprising a first tag sequence and a second tag sequence, wherein one of the tag sequences has been converted by a methylation conversion agent and the other tag sequence has not been converted by the methylation conversion agent. Other embodiments provided

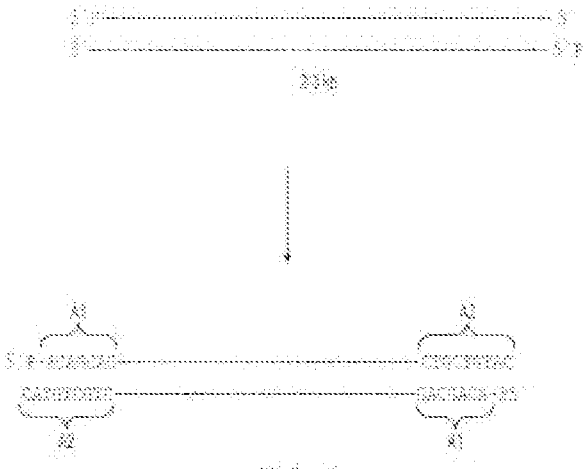


FIG. 1

IDT Exhibit 1020.1034

include intermediates for making the mate-pair library and kits for making the mate-pair libraries. Also provided is software and computer systems for analyzing the methylation levels of genomic DNA from which the tag sequences were derived.

IDT Exhibit 1020.1035

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 January 2010 (07.01.2010)

PCT

(10) International Publication Number
WO 2010/003153 A2

(51) **International Patent Classification**:
C12Q 1/68 (2006.01) *C12N 15/11* (2006.01)

(21) **International Application Number**:
PCT/US2009/049724

(22) **International Filing Date**:
6 July 2009 (06.07.2009)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/133,891 3 July 2008 (03.07.2008) US
61/149,976 4 February 2009 (04.02.2009) US

(71) **Applicant** *(for all designated States except US)*: **LIFE TECHNOLOGIES CORPORATION** [US/US]; C/o Intellevate, P. O. Box 52050, Minneapolis, MN 55402 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **MCKERNAN, Kevin, J.** [US/US]; 10 Manley Street, Marblehead, MA 01945 (US). **SCHROEDER, Benjamin, G.** [US/US]; 1964 Shoreview Avenue, San Mateo, CA 94401 (US). **BOYD, Victoria, L.** [US/US]; 199 Fir Street, San Carlos, CA 94070 (US).

(74) **Agents**: **FINN, Andrew, K.** et al.; C/o Intellevate, P. O. Box 52050, Minneapolis, MN 55402 (US).

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

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

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: METHYLATION ANALYSIS OF MATE PAIRS

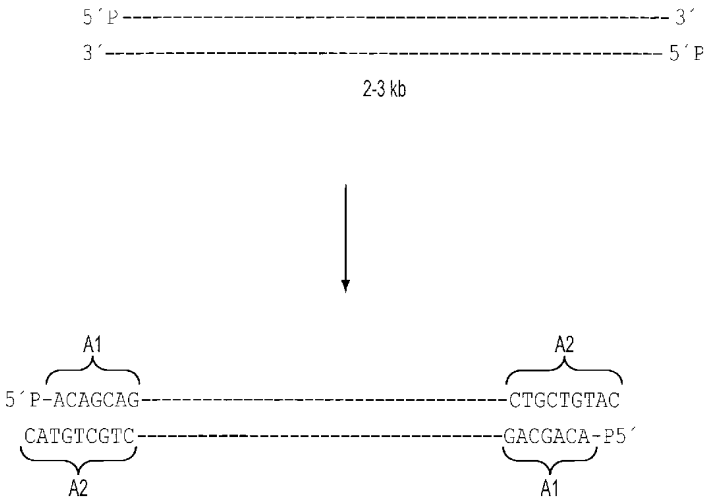


FIG. 1

(57) **Abstract**: Various embodiments of the present teachings relate to methods for the methylation analysis of nucleic acids. The subject methods include methods that result in the preparation of mate-pair libraries suitable for highly multiplexed DNA sequencing. Embodiments include methods of preparing mate-pair libraries comprising a first tag sequence and a second tag sequence, wherein one of the tag sequences has been converted by a methylation conversion agent and the other tag sequence has not been converted by the methylation conversion agent. Other embodiments provided include intermediates for making the mate-pair library and kits for making the mate-pair libraries. Also provided is software and computer systems for analyzing the methylation levels of genomic DNA from which the tag sequences were derived.

WO 2010/003153 A2

METHYLATION ANALYSIS OF MATE PAIRS

Priority Claims

[0001] This application claims the benefit of priority to U.S. Provisional Application No. 61/133,891, filed July 3, 2008, entitled, Methylation Analysis of Mate Pairs, and U.S. Provisional Application No. 61/149,976, filed February 4, 2009, entitled, Methylation Analysis of Mate Pairs, which are incorporated herein by reference.

Field

[0002] This invention is in the field of analysis of a methylated nucleic acid by means of high throughput nucleic acid sequencing techniques.

Background

[0003] Regions of genomic DNA are frequently methylated. The base 5-methyl cytosine is the most frequently encountered methylated base in the DNA derived from eukaryotic cells. 5-methyl cytosine results from methylation of the number 5 carbon in the pyrimidine ring of cytosine. The methylation of genomic DNA, which is reversible, is well-known to have important biological significance. Such areas of biological significance include the activation and inactivation of genomic regions for transcription. For example, carcinogenesis may occur by the methylation of tumor suppressing genes, which may deactivate the genes. Consequently, the analysis of methylation patterns in cancer cells is a major area of research.

[0004] Most conventional methods of nucleic acid methylation analysis involve treatment of the nucleic acid of interest with a methylation conversion agent. Exemplary of such conversion agents is sodium bisulfite. Sodium bisulfite converts the nucleic acid base cytosine to uracil. 5-methylcytosine, however, is not converted by sodium bisulfite under conditions employed for methylation analysis. Thus, sequencing the sodium bisulfite-treated DNA will result in the detection of an uracil when the cytosine was not methylated, and the detection of a cytosine when the cytosine was methylated. Many methods exist for manipulating and detecting sequence variations in genomic DNA that has been treated with a methylation conversion agent such as

IDT Exhibit 1020.1037

sodium bisulfite. Such techniques include DNA sequencing, real-time PCR, and the oligonucleotide ligation assay (OLA).

[0005] There are many methods of high throughput sequence analysis that result in extremely high numbers of relatively short stretches of DNA being sequenced, e.g., the SOLiD™ sequencing system sold by Applied Biosystems or the Genome Analyzer sold by Illumina.

[0006] One method of extracting more information from such short DNA sequences is to use mate-pair sequence tags, wherein the approximate distance between the mate-pair sequences on the genome is known. Mate-pairs of sequence tags can be derived from a single polynucleotide fragment. Such genomic fragments used to generate mate-pairs are typically of a length within a pre-determined range of possible lengths, such as, for example 2-3kb. This length information can be used to help map the sequence information to a genomic reference sequence. Given the relatively short lengths of the sequence reads, such matching back to a reference sequence can be important for assembling accurate sequence information. The use of mate-pair analysis with a methylation conversion agent for methylation analysis can be problematic for mapping back to genomic reference sequences because of reduced sequence complexity after exposure to the methylation conversion. Sequence complexity is reduced because of the loss of cytosines caused by exposure to sodium bisulfite, which results in mate-pairs rich in adenine, thymine, and guanine following amplification.

[0007] There is thus a long-felt need in the industry for sequencing methylated DNA quickly and accurately. Methods, reagents, genetic constructs, kits, data analysis systems, and software for addressing the problems associated with reduced sequence complexity arising from the use of methylation conversion agents are provided herein.

Summary

[0008] Various embodiments of the present teachings relate to methods of analyzing the methylation state of genomic DNA. The methods involve fragmenting genomic DNA. In at least one embodiment, the DNA fragments are circularized to produce a double-stranded circular DNA comprising a nick on one strand. A nick translation in the presence of methylation conversion agent resistant nucleotide triphosphate is then performed. The circular genetic construction can be linearized

IDT Exhibit 1020.1038

prior to the nick translation reaction. After the nick translation step, two tag regions of a mate-pair are created, wherein the first tag region may comprise methylation conversion resistant nucleotides and the second tag region may lack methylation conversion resistant nucleotides and not be methylation conversion agent resistant. The construction can, in some embodiments, be amplified. The circular genetic construction can in some embodiments comprise a specific binding pair member so as to facilitate strand separation and purification. The tag regions can be sequenced to provide information about the methylation state of the genomic DNA from which the clone was derived.

[0009] The present teachings also relate to methods of analyzing the methylation state of genomic DNA comprising fragmenting a genomic DNA and using the fragmented DNA to form linear genetic constructions, each construction having a first tag sequence and a second tag sequence, wherein the first tag and the second tag are derived from a single genomic DNA fragment. In certain embodiments, the first tag sequence may be converted by a methylation conversion agent, while the second tag sequence is not converted by a methylation conversion agent. The constructs can be clonally amplified to provide templates for sequencing.

[0010] The present teachings also relate to polynucleotide constructions comprising a first tag sequence and a second tag sequence, wherein the first tag sequence and the second tag sequence are derived from a single fragment of genomic DNA. The first tag may comprise methylation conversion resistant nucleotides that have been incorporated into the construction by an *in vitro* reaction and, in certain embodiments, the second tag does not comprise incorporated methylation conversion resistant nucleotides. In some embodiments, the genetic construction comprises a specific binding pair member. In some embodiments, the genetic construction comprises primer-binding sites.

[0011] Embodiments of the present teachings also include kits comprising an adapter having a first strand having methylation conversion resistant nucleotides and a second strand complementary to the first strand, wherein the second strand optionally comprises methylation conversion resistant nucleotides. Kits can further comprise oligonucleotide primers specific for a strand of the adapter. Kits can also comprise one

IDT Exhibit 1020.1039

or more additional reagents for use in carrying out one or more embodiments of the methods disclosed herein, such as a DNA polymerase, a DNA ligase, methylation conversion resistant nucleotides, etc.

[0012] The present teachings further relate to methods of matching a DNA sequence to a genomic sequence database, the methods comprising comparing a data record comprising (1) a first tag sequence that corresponds to a DNA sequence that has not been modified by a methylation conversion agent, (2) a second tag sequence that corresponds to a DNA sequence that may have been modified by a methylation conversion agent, and (3) a distance value indicative of the approximate distance in the genome between the first tag sequence and the second tag sequence, with DNA sequence information in the genomic database. Such methods can be implemented by general purpose computers. Embodiments include systems and software for implementing such methods.

[0013] Further embodiments of the present teachings relate to methods of amplifying polynucleotides converted by a methylation conversion agent in which primer-adapters may be ligated to fragments of genomic DNA. The adapters may comprise a double-stranded polynucleotide having a first stand and second strand complementary to the first strand, wherein the first strand may comprise methylation conversion resistant nucleotides and, in certain embodiments, the second strand lacks methylation conversion resistant nucleotides. The adapter modified polynucleotide may then be amplified using primers specific for the sequences in the second strand of the adapter, after the sequences have been converted. In at least one embodiment of the present teachings, the first strand may comprise methylation conversion resistant nucleotides and the second strand may optionally lack methylation conversion resistant nucleotides. The second strand of the adapter may optionally be converted into a methylation resistant sequence during a nick translation step with dNTPs comprising 5-methylcystosine (5mC dNTPs), or other methylation conversion resistant nucleotides to generate adapters that are fully methylation conversion resistant on both strands of the DNA. Adapters that are fully methylation conversion resistant on both strands of the DNA will be the same before and after bisulfite conversion.

IDT Exhibit 1020.1040

[0014] Embodiments of the present teachings also relate to methods of analyzing the methylation state of a polynucleotide bound to a solid support. In at least one embodiment, the methods involve fragmenting genomic DNA and circularizing a fragment with two cap adapters that create sticky ends and an internal adapter comprising a specific binding moiety. A nick translation may then be performed and the circularized polynucleotide linearized to create two tag regions of a mate-pair. The polynucleotide can be bound to a solid support using a cognate specific binding moiety to bind the specific binding moiety. The double-stranded polynucleotide can be denatured, and the unbound strand may be eluted and collected. One or both of the bound or unbound strands may be exposed to a methylation conversion reagent, such as sodium bisulfite. The converted strand may then be amplified and sequenced to analyze the methylation of the polynucleotide.

Brief Description of the Drawings

[0015] Figure 1 is an example of a 2-3kb fragment of genomic DNA undergoing ligation to add adapters (cap adapters), wherein the cap adapters comprise an EcoP151 restriction endonuclease recognition site;

[0016] Figure 2 shows an adapter modified genomic DNA circularized by sticky end ligation to an internal adapter comprising a biotin on one strand;

[0017] Figure 3 shows the circular DNA construction linearized by incubation with the restriction endonuclease EcoP15I;

[0018] Figure 4 shows the linearized fragment incubated with a nick translation enzyme and the conversion resistant nucleotide 5-methylcytosine (5mC);

[0019] Figure 5 shows the location of the 5mC's in one strand after the nick translation reaction;

[0020] Figure 6 shows the addition of the primer-adapters to the linearized fragment;

[0021] Figure 7 shows the construct in the bottom of Figure 6 following the removal of the nicks after nick translation;

[0022] Figure 8 shows the selectively recovered strand, i.e., the strand lacking the biotin;

IDT Exhibit 1020.1041

[0023] Figure 9 shows the treatment of the construct with the methylation conversion agent, sodium bisulfite;

[0024] Figure 10 shows the addition of P2 adapters to one end of the bisulfite converted construction containing the two tag regions, wherein PCR is used to fill in the second strand of the P2 region;

[0025] Figure 11 shows the sequence of the internal adapter, the P1-A/P1-B adapter and the P2-A tail;

[0026] Figure 12 shows the sequences of the internal adapter, the 5mC P1-A/P1B adapter, and the P2-A-tailed library amplification primer used in the method illustrated in Figures 1-11; and

[0027] Figures 13-16 show an exemplary method of preparing long mate-pairs using a double-stranded, circularized polynucleotide having a nick on each strand.

Definitions and Embodiments

[0028] The section headings used herein are for organizational purposes only and are not to be construed as limiting the described subject matter in any way. All literature and similar materials cited in this application, including but not limited to, patents, patent applications, articles, books, treatises, and internet web pages are expressly incorporated by reference in their entirety for any purpose. When definitions of terms in incorporated references appear to differ from the definitions provided in the present teachings, the definition provided in the present teachings shall control. It will be appreciated that there is an implied "about" prior to the temperatures, concentrations, times, etc. discussed in the present teachings, such that slight and insubstantial deviations are within the scope of the present teachings. In this application, the use of the singular includes the plural unless specifically stated otherwise. Also, the use of "comprise", "comprises", "comprising", "contain", "contains", "containing", "include", "includes", and "including" are not intended to be limiting. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the present teachings.

[0029] Unless otherwise defined, scientific and technical terms used in connection with the present teachings described herein shall have the meanings that are

IDT Exhibit 1020.1042

commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Generally, nomenclatures utilized in connection with, and techniques of, cell and tissue culture, molecular biology, and protein and oligo- or polynucleotide chemistry and hybridization described herein are those well known and commonly used in the art. Standard techniques are used, for example, for nucleic acid purification and preparation, chemical analysis, recombinant nucleic acid, and oligonucleotide synthesis. Enzymatic reactions and purification techniques are performed according to manufacturer's specifications or as commonly accomplished in the art or as described herein. The techniques and procedures described herein are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the instant specification. *See, e.g.*, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 2000). The nomenclatures utilized in connection with, and the laboratory procedures and techniques described herein are those well known and commonly used in the art.

[0030] As utilized in accordance with the embodiments provided herein, the following terms, unless otherwise indicated, shall be understood to have the following meanings:

[0031] The term "nucleotide" refers to a phosphate ester of a nucleoside, as a monomer unit or within a nucleic acid. "Nucleotide 5'-triphosphate" refers to a nucleotide with a triphosphate ester group at the 5' position, and is sometimes denoted as "NTP", or "dNTP" and "ddNTP" to particularly point out the structural features of the ribose sugar. The triphosphate ester group can include sulfur substitutions for the various oxygens, *e.g.* .alpha.-thio-nucleotide 5'-triphosphates. For a review of nucleic acid chemistry, see Shabarova, Z. and Bogdanov, A., Advanced Organic Chemistry of Nucleic Acids, VCH, New York, 1994.

[0032] The term "nucleic acid" refers to natural nucleic acids, artificial nucleic acids, analogs thereof, or combinations thereof.

IDT Exhibit 1020.1043

[0033] As used herein, the terms "polynucleotide" and "oligonucleotide" are used interchangeably and mean single-stranded and double-stranded polymers of nucleotide monomers (nucleic acids), including, but not limited to, 2'-deoxyribonucleotides (nucleic acid) and ribonucleotides (RNA) linked by internucleotide phosphodiester bond linkages, *e.g.* 3'-5' and 2'-5', inverted linkages, *e.g.* 3'-3' and 5'-5', branched structures, or analog nucleic acids. Polynucleotides may have associated counter ions, such as H^+, NH_4^+, trialkylammonium, Mg^{2+}, Na^+ and the like. A polynucleotide can be composed entirely of deoxyribonucleotides, entirely of ribonucleotides, or chimeric mixtures thereof. Polynucleotides can be comprised of nucleobase and sugar analogs. Polynucleotides typically range in size from a few monomeric units, *e.g.* 5-40 when they are more commonly frequently referred to in the art as oligonucleotides, to several thousands of monomeric nucleotide units. Unless denoted otherwise, whenever a polynucleotide sequence is represented, it will be understood that the nucleotides are in 5' to 3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes deoxythymidine.

[0034] Polynucleotides are said to have "5' ends" and "3' ends" because mononucleotides react to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotide or polynucleotide is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also can be said to have 5' and 3' ends.

[0035] The phrases "DNA fragment of interest," "polynucleotide of interest," "target polynucleotide," "DNA template," "template polynucleotide," and variations thereof mean the DNA fragment or polynucleotide that one is interested in identifying, characterizing, or manipulating. As used herein, the terms "template" and "polynucleotide of interest" refer to a nucleic acid that is acted upon, such as, for

IDT Exhibit 1020.1044

example, a nucleic acid that is to be mixed with polymerase. In some embodiments, the polynucleotide of interest is a double stranded polynucleotide of interest ("DSPI").

[0036] As used herein, the phrases "different strand of a polynucleotide," "different strand of a nucleic acid molecule," and variations thereof refer to a nucleic acid strand of a duplex polynucleotide that is not from the same side as another strand of the duplex polynucleotide.

[0037] As used herein, the phrase "paired tag," also referred to as a "tag mate-pair," "mate-pair," or "paired-end," contains two tags (each a nucleic acid sequence) that are from each end region of a polynucleotide of interest. Thus, a paired tag includes sequence fragment information from two parts of a polynucleotide. In some embodiments, this information can be combined with information regarding the polynucleotide's size, such that the separation between the two sequenced fragments is known to at least a first approximation. This information can be used in mapping where the sequence tags came from.

[0038] As used herein, the term "nick" refers to a point in a double stranded polynucleotide where there is no phosphodiester bond between adjacent nucleotides of one strand of the polynucleotide.

[0039] The term "nick translation" as used herein refers to a coupled polymerization/degradation or strand displacement process that is characterized by a coordinated 5' to 3' DNA polymerase activity and a 5' to 3' exonuclease activity or 5' to 3' strand displacement. As will be appreciated by one of skill in the art, a "nick translation," as the term is used herein, can occur on a nick or to a gap. As will be appreciated by one of skill in the art, in some embodiments, the "nick translation" of a gap entails the insertion of appropriate nucleotides in order to form a traditional nick that lacks a phosphodiester bond, which is then translated.

[0040] As used herein, the phrases "nick is translated into the DNA fragment of interest," "nick is translated into the polynucleotide of interest," and variations thereof refer to the translocation of a nick to a position in the strand that includes the nick that is within the DNA fragment or polynucleotide of interest.

[0041] An "analog" nucleic acid or nucleotide is a nucleic acid or nucleotide that is not normally found in a host to which it is being added or in a sample that is

IDT Exhibit 1020.1045

being tested. The target sequence may not comprise an analog nucleic acid because it is the sequence that is to be identified, modified, or manipulated. Nucleic acid analogs include artificial nucleic acids, synthetic nucleic acids, or combination thereof. Thus, for example, in one embodiment, PNA (peptide nucleic acid) is an analog nucleic acid, as is L-DNA and LNA (locked nucleic acids), iso-C/iso-G, L-RNA, O-methyl RNA, or other such nucleic acids. In at least one embodiment, any modified nucleic acid will be encompassed within the term analog nucleic acid. In other embodiments, an analog nucleic acid can be a nucleic acid that will not substantially hybridize to native nucleic acids in a system, but will hybridize to other analog nucleic acids; thus, in those embodiments, PNA would not be an analog nucleic acid, but L-DNA would be an analog nucleic acid. For example, while L-DNA can hybridize to PNA in an effective manner, L-DNA will not hybridize to D-DNA or D-RNA in a similar effective manner. Thus, nucleotides or nucleic acids that can hybridize to a probe or target sequence but lack at least one natural nucleotide characteristic, such as susceptibility to degradation by nucleases or binding to D-DNA or D-RNA, may be analog nucleotides or nucleic acids in some embodiments. Of course, the analog nucleotide or nucleic acid need not have every difference.

[0042] The term "nucleic acid sequencing chemistry" as used herein refers to a type of chemistry and associated methods used to sequence a polynucleotide to produce a sequencing result. A wide variety of sequencing chemistries are known in the art. Examples of various types of sequencing chemistries useful in various embodiments disclosed herein include, but are not limited to, Maxam-Gilbert sequencing, chain termination methods, dye-labeled terminator methods, sequencing using reversible terminators, sequencing of nucleic acid by pyrophosphate detection ("pyrophosphate sequencing" or "pyrosequencing"), and sequencing by ligation. Such sequencing chemistries and corresponding sequencing reagents are described, for example, in U.S. Patent Nos. 7,057,026; 5,763,594; 5,808,045; 6,232,465; 5,990,300; 5,872,244; 6,613,523; 6,664,079; 5,302,509; 6,255,475; 6,309836; 6,613,513; 6,841,128; 6,210,891; 6,258,568; 5,750,341; and 6,306,597; and PCT Publication Nos. WO 91/06678 A1; WO 93/05183 A1; WO 06/074351 A2; WO 03/054142 A2; WO 03/004690 A2; WO 07/002204 A2; WO 06/084132 A2; and WO 06/073504 A2.

[0043] As used herein, the term "polymerase chain reaction" (PCR) refers to the method described by K. B. Mullis in U.S. Patent Nos. 4,683,195 and 4,683,202, which describe a method for increasing the concentration of a segment of a polynucleotide of interest sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the polynucleotide of interest sequence comprises introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired polynucleotide of interest sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective strands of the double stranded polynucleotide of interest sequence. To effect amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the polynucleotide of interest molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing and polymerase extension can be repeated many times (i.e., denaturation, annealing and extension constitute one "cycle"; there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired polynucleotide of interest. The length of the amplified segment of the desired polynucleotide of interest is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of repeating the process, the method is referred to as the "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the polynucleotide of interest sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified".

[0044] "Clonal amplification" refers to the generation of many copies of an individual molecule. Various methods known in the art can be used for clonal amplification. For example, emulsion PCR is one method, and involves isolating individual DNA molecules along with primer-coated beads in aqueous bubbles within an oil phase. A polymerase chain reaction (PCR) then coats each bead with clonal copies of the isolated library molecule and these beads are subsequently immobilized for later sequencing. Emulsion PCR is used in the methods published by Marguilis *et al.* and Shendure and Porreca *et al.* (also known as "polony sequencing"). *See* Margulies, et al. (2005) *Nature* 437: 376-380; Shendure et al., *Science* 309 (5741):

IDT Exhibit 1020.1047

1728-1732. Another method for clonal amplification is "bridge PCR," where fragments are amplified upon primers attached to a solid surface. *See, e.g.*, PCT Publication No. WO 98/44151 and U.S. Patent No. 6,090,592. These methods, as well as other methods of clonal amplification, both produce many physically isolated locations that each contains many copies derived from a single molecule polynucleotide fragment.

[0045] As used herein, "binding moiety" means a molecule that can bind to a purifying moiety under appropriate conditions. The interaction between the binding moiety and purifying moiety is strong enough to allow enrichment and/or purification of the binding moiety and a molecule associated with it, for example, a paired tag clone. Biotin is an example of a binding moiety. In some embodiments, by coupling a binding moiety to an adapter, binding of the binding moiety to a purifying moiety target allows purification of the paired tag clone. In some embodiments, the purifying moiety can be present on a solid support, such as, for example, streptavidin bound to a polystyrene bead.

[0046] As used herein, the term "specific binding pair member" means a member of a pair of molecules that specifically bind to one another with sufficient specificity so as to avoid the binding of interfering quantities of background compounds. A "binding moiety" can be a specific binding pair member. A least one member of a specific binding pair, and possibly both members, are biological molecules or analogs thereof, such as proteins, carbohydrates, polynucleotides, metabolic intermediates and the like. Exemplary of such specific binding pairs are biotin and avidin, biotin and streptavidin, lectins and carbohydrates, antibodies and antigens, complementary nucleic acids and nucleic acid analogues. When referring to a pair of specific binding pair members, the second binding pair member can be referred to as the cognate pair member or cognate specific binding pair member. For example, when referring to biotin attached to a nucleic acid, it may be said that the nucleic acid is purified by binding to the cognate specific binding pair member, e.g., avidin. Conversely, biotin could be said to be the cognate specific binding pair member for avidin.

[0047] The term "solid support" refers to any solid phase material upon which an oligonucleotide is synthesized, attached, or immobilized. Solid support

IDT Exhibit 1020.1048

encompasses terms such as "resin", "solid phase", and "support". A solid support can be composed of organic polymers such as polystyrene, polyethylene, polypropylene, polyfluoroethylene, polyethyleneoxy, and polyacrylamide, as well as co-polymers and grafts thereof. A solid support can also be inorganic, such as, for example, glass, silica, controlled-pore-glass (CPG), or reverse-phase silica. The configuration of a solid support can be in the form of beads, spheres, particles, granules, a gel, a surface, or combinations thereof. Surfaces can be planar, substantially planar, or non-planar. Solid supports can be porous or non-porous, and can have swelling or non-swelling characteristics. A solid support can be configured in the form of a well, depression or other container, vessel, feature, location, or position. A plurality of solid supports can be configured in an array at various locations, *e.g.*, positions, addressable for robotic delivery of reagents, or by detection means including scanning by laser illumination and confocal or deflective light gathering.

[0048] The term "distance value" means a value indicative of the approximate physical distance in the genome between the first tag sequence and the second tag sequence.

[0049] The term "nick translation enzyme" means an enzyme with DNA polymerase activity that also has 5' to 3' exonuclease activity, thus giving the appearance of a moving or "translating" a nick (or gap) in a double-stranded region of DNA from one location to another as polymerase and exonuclease activity proceed in concert with one another. Methods for performing nick translation reactions are known to those of skill in the art. *See, e.g.,* Rigby, P. W. et al. (1977), J. Mol. Biol. 113, 237. A variety of suitable polymerases can be used to perform the nick translation reaction, including for example, *E. coli* DNA polymerase I, Taq DNA polymerase, Vent DNA polymerase, Klenow DNA polymerase I, and phi29 DNA polymerase. Depending on the enzyme used, nick translation can occur by 5' to 3' exonuclease activity or by 5' to 3' strand displacement.

[0050] The term "methylation conversion agent" means a chemical reagent that modifies the chemical structure of a nucleotide base so as to produce a nucleotide base with different base pairing specificity. Exemplary of such reagents is sodium bisulfite (and other bisulfite salts) that deaminates cytosine to produce uracil.

IDT Exhibit 1020.1049

[0051] As used herein, the phrases "converted nucleotide," "converted nucleic acid," and variations thereof mean any nucleotide base or nucleic acid that has been chemically modified by a methylation conversion agent so as to produce a nucleotide base or nucleic acid with different base pairing. An example of a converted base is the deamination of cytosine to uracil by sodium bisulfite. Thus, cytosine is said to be converted by sodium bisulfite to uracil.

[0052] The term "methylation conversion agent resistant nucleotide" means a nucleotide comprising a nucleic acid base that is not chemically altered by the methylation conversion agent (used in a given embodiment) so as to change the base pairing specificity of the nucleotide base. Methylation conversion agent resistant nucleotides are capable of being incorporated by a nick translation enzyme in a primer extension reaction. Exemplary of methylation conversion resistant nucleotides is 5-methylcytosine (5mC) used in conjunction with sodium bisulfite. Thus, 5-methylcytosine is not deaminated when exposed to sodium bisulfite.

[0053] The term "adapter" means a synthetic double-stranded polynucleotide. Adapters can be ligated to a polynucleotide so as to facilitate further structural or physical manipulations of the polynucleotide. Adapters can be used to do one or more of the following: introduce amplification primer binding sites, introduce sequencing primer binding sites, introduce restriction endonuclease recognition sites, introduce specific binding pair members, or facilitate the circularization of a linear polynucleotide molecule.

[0054] As used herein, the phrase "a full set of dNTPs" means a set of at least 4 nucleotides capable of supporting a nick translation reaction, e.g., dATP, dCTP, dGTP, and dTTP. Various analogs can also be employed in addition to or in place of any one of dATP, dCTP, dGTP, and dTTP, including, but not limited to, methylated bases such as 5-methylcytosine. The phrase "a full set of regular dNTPs" means a set of nucleotides consisting of dATP, dCTP, dGTP, and dTTP.

[0055] The terms "tag," "tag region," and "tag sequence" as used herein refer to each of the two polynucleotide sections of mate-pair clone that are derived from polynucleotide sequences at the termini of a genomic fragment. Tag regions and tag sequence can be sequenced to produce base pair sequences representative of the actual

IDT Exhibit 1020.1050

tag regions. The terms can be used to refer to a sub-sequence of a polynucleotide of interest.

Description

[0056] Various embodiments of the present present teachings relate to methods for the methylation analysis of nucleic acids. The subject methods include methods that may result in the preparation of mate-pair libraries suitable for highly multiplexed DNA sequencing. Subject embodiments include methods of preparing mate-pair libraries comprising a first tag sequence and a second tag sequence, wherein one of the tag sequences may be converted by a methylation conversion agent and the other tag sequence may not be converted by the methylation conversion agent. Other embodiments provided include intermediates for making the mate-pair library and kits for making the mate-pair libraries. It also be appreciated that while much of the description provided herein focuses on the use of methylation conversion resistant nucleotides to generate tag regions that are resistant to conversion by methylation conversion agents, the embodiments provided herein can be adapted to take advantage of the inability of many methylation conversion agents to convert nucleotide bases that are base paired, i.e., in double-stranded form.

[0057] In various embodiments, genomic DNA obtained from cells of interest is fragmented. Methods of DNA fragmentation and the selection of the proper fragmentation method(s) are well-known to persons of ordinary skill in the art. Such methods include, for example, sonication, shearing, digestion with restriction endonucleases, random chemical degradation, and the like. DNA can be obtained from a variety of different cell types, including both eukaryotic and prokaryotic. DNA can be obtained from a variety of different tissues in higher organisms. In some embodiments, DNA can be obtained from tumors.

[0058] In at least one embodiment, the fragmented DNA can be size selected so as to produce a fraction of DNA fragments of the desired size range. Fractionation of DNA fragments according to size is well known to persons of ordinary skill in the art, and such fractionation techniques may include electrophoresis, size exclusion gel chromatography, HPLC, centrifugation, and the like. The use of size fractionated DNA

IDT Exhibit 1020.1051

fragments can be used to produce mate-pair libraries in which the approximate distance between the mate-pairs on the genome of interest is known, thereby facilitating matching of the mate-pairs to pre-existing genomic sequence information.

[0059] In some embodiments, DNA fragments can be circularized in order to provide for the generation of mate-pair libraries. DNA fragments can be modified so as to enable circularization. Adapters can be added to the ends of the genomic fragments so as to facilitate circularization. Such adapters can be blunt-ended, sticky-ended, or comprise a sticky-end and a blunt-end. After the addition of adapters to the ends of the DNA fragment, the modified fragment can be circularized. Circularization can be achieved by enzymatic or chemical ligation of the ends of the genetic construction to one another or through an intermediate polynucleotide. In some embodiments, the adapter modified fragment can be circularized by ligation to an internal adapter fragment. Internal adapter fragments can optionally comprise a specific binding pair member, e.g., biotin, digoxygenin, and the like.

[0060] Internal adapter fragments can be used to facilitate the generation of mate-pair libraries. Internal adapter fragments, in some embodiments, can comprise restriction endonuclease recognition sites for restriction endonucleases that cleave at a site distal to the recognition sequence, e.g., type IIs or type III restriction endonuclease recognition sites. For example, the type IIs or type III restriction recognition sites can be oriented so as to enable the enzyme to cut the genomic DNA in the proximity of the junction between the internal adapter and the genomic DNA so as to generate tag sequences between the cut sites and the junctions. The internal adapter fragments can further comprise a specific binding moiety attached to one strand of the internal adapter. In at least one embodiment, the specific binding moiety is biotin. In some embodiments of the present teachings, the specific binding moiety can be used to remove an undesired strand of a nucleic acid construction in subsequent steps. In other embodiments of the present teachings, the specific binding moiety can be used to isolate a desired strand of a nucleic acid construction. Guidance on the creation of mate-pair libraries can be found in, among other places, PCT Published Application No. WO 05/42781 A2.

IDT Exhibit 1020.1052

[0061] In some embodiments of the present teachings, the circular genetic construction formed by circularizing the genomic DNA fragment for analysis will comprise a nick located in one strand of the circular genetic construction. The nick can be located at the junction between the genomic DNA for analysis and an adapter added to the genomic DNA. The nick can be formed by not phosphorylating a 5' terminus of a strand of the internal adapter, thereby preventing a ligation event from taking place.

[0062] After circularization, the circular DNA construction can be linearized so as to produce a genetic construction having a first tag sequence and a second tag sequence at opposite ends of the linear nucleic acid molecule. Generating the tag regions can, in certain embodiments, occur in the same step as the linearization step. In at least one embodiment, the double-stranded cleavage of the circular DNA construction can be achieved by an enzymatic or chemical cleavage. Linearization can be achieved, for example, by making a double-stranded cut in the circular genetic construction in one or more locations. One example of such methods of cleaving the circular genetic constructions is to use a type IIs or type III restriction endonuclease (or equivalents thereof) that is specific for restriction endonuclease recognition sites in the internal adapter.

[0063] According to at least one embodiment of the present teachings, the circular genetic construction formed between the genomic DNA fragment of interest and the internal adapter comprises a single-stranded nick. The nick can be subsequently translated during later steps in various embodiments of the present teachings. The nick can be located at the junction between the internal adapter and the genomic DNA fragments, or at a junction between the internal adapter and the adapter-modified genomic fragment. The nick may be located 3' relative to the tag region that is to remain susceptible to conversion by a methylation conversion reagent. The nick can be created by using an internal adapter that is not phosphorylated at one of its two 5' termini, thus creating a nick at the desired position during the circularization step. Alternatively, the nick (or nicks if both strands contain a nick) can be introduced by other enzymatic means or chemically, or by a combination of chemical and enzymatic means.

IDT Exhibit 1020.1053

[0064] Subsequent to the linearization of the circular genetic construction, the nick can be translated by incubating the genetic construction in the presence of a nick translation enzyme, a suitable buffering environment, and a full set of dNTPs, wherein the set of dNTPs comprises at least one methylation conversion resistant nucleotide. Exemplary of such methylation conversion resistant nucleotides is 5-methylcytosine. In at least one embodiment, one or more of the dNTPs in the full set of dNTPs can be a methylation conversion resistant nucleotide.

[0065] During the process of nick translation, DNA synthesis proceeds through only one of the tag sequence regions. The DNA synthesis can, in some embodiments, proceed through the internal adapter region of the linearized construction. In some embodiments, after nick translation, a portion of one strand can comprise methylation conversion resistant nucleotides incorporated during the nick translation reaction. In at least one embodiment, the methylation conversion resistant nucleotides are in one of the tag regions, but not the other. The strand of the linear genetic construction that is not modified by the nick translation enzyme does not comprise the incorporated methylation conversion resistant nucleotides.

[0066] According to at least one embodiment, the linear double-stranded genetic constructions that remain after the nick translation reaction can be modified with primer-adapters so as to facilitate manipulation of a strand or strands comprising the tag regions. Primer-adapters can be joined to the linearized genetic construction either before or after treatment of the linearized genetic construction with a methylation conversion agent. In at least one embodiment, the primer-adapters are joined to the linearized genetic construction before treatment with a methylation conversion agent. Primer-adapters can be ligated to the termini of the linear genetic construction. The primer-adapters can comprise a primer binding site for use in amplifications or selective binding to complementary sequences for enrichment of desired products. The primer-adapters do not require 5' phosphorylated ends, but in some embodiments can have 5' phosphorylated ends. In at least one embodiment, the ligation product formed between the linearized construction and the primer-adapters can be subjected to a nick translation reaction to remove nicks formed between the 5' ends of the strands and the primer-adapter and the linearized construction. In at least one embodiment, the nick

IDT Exhibit 1020.1054

translation reaction can take place in the absence of methylation conversion resistant nucleotides.

[0067] In at least one embodiment, the primer-adapter can contain methylation conversion resistant nucleotides in one strand of a double-stranded adapter used to introduce amplification primer binding sites. As used herein, the primer-adapters containing methylation conversion resistant nucleotides in one strand are referred to as "partially protected primer-adapters." Partially protected primer adapters can be used to preferentially amplify polynucleotides that have been converted by a methylation conversion agent. The methylation conversion agents, such as sodium bisulfite, do not always completely react with all polynucleotides and nucleic acid bases in a conversion reaction. By having a strand that is converted by the methylation conversion agent and a strand that is resistant to conversion, it is possible to employ complementary oligonucleotide primers specific for the converted primer binding regions of the partially protected primer-adapter so as to enrich or selectively amplify for those polynucleotides that have been converted by the methylation conversion agent. The inventors have discovered that conversion of the nucleotide bases in the primer-adapter by a methylation conversion agent is correlated with conversion of the unprotected bases located in between the primer adapters, e.g., the tag regions and the internal adapters.

[0068] After addition of the primer-adapters to the linear genetic construction comprising the tag regions, the strand containing the protected tag regions and the unprotected tag regions can be isolated from the complementary strand, so as to be prepared for subsequent manipulations and analysis, e.g. sequencing. The strands of the linearized genetic construction can be denatured and the desired strand retained. Such purification of the desired member of the denatured polynucleotide strands can be achieved by numerous methods well known to the person of ordinary skill in the art of molecular biology, e.g., electrophoresis, chromatography, and the like. In embodiments employing internal adapters comprising a specific binding pair member, the strand comprising the specific binding pair member may be conveniently separated from the other strand by contacting the specific binding pair member with its cognate specific binding pair member that has been immobilized on a solid support. Examples of such

IDT Exhibit 1020.1055

solid supports include glass, plastic, and the like, that are capable of being modified so as to attach the cognate specific binding pair member or moiety to the surface. The free strand in the solution can be easily purified away from the balance trend so as to be available for subsequent manipulations, e.g., sequencing or amplification. In at least one embodiment, the specific binding pair member comprises biotin and its cognate specific binding pair member comprises streptavidin bound to polystyrene beads.

[0069] The strand of the linearized genetic construction comprises two tag regions: (1) a first tag region comprising methylation conversion agent resistant nucleotides, and (2) a second tag region that lacks methylation conversion agent resistant nucleotides. In at least one embodiment of the present teachings, the strand of the linearized genetic construction is incubated with at least one methylation conversion agent, such as sodium bisulfite. The use of methylation conversion agents for analysis of DNA is well known to the person skilled in the art. The methylation conversion reaction proceeds as long as necessary to provide reasonable certainty that the majority of accessible unprotected bases are converted. Detailed protocols for the use of bisulfite as a methylation conversion agent can be found, for example, in U.S. Patent Nos. 7,371,526; 7,368,239; and 7,262,013; and U.S. Patent Application Publication No. US 2006/0286577A. In embodiments employing bisulfite salts as a methylation conversion agent, formamide can be used as a denaturant instead of NaOH, the traditional denaturant for bisulfite methylation analysis.

[0070] In at least one embodiment of the present teachings, the methylation conversion reaction can be performed while the linearized genetic construction is bound to a solid support. For example, when the internal adapter comprises biotin as a specific binding moiety, the linearized genetic construction may be bound to streptavidin on a solid support, such as, for example, polystyrene beads. The inventors have discovered that sodium bisulfite conversion can be carried out on bound constructions. In at least one embodiment, the streptavidin polystyrene beads may be non-magnetic. Without wishing to be bound by theory, it is believed that the use of non-magnetic beads may prevent the oxidation of the nucleic acids by the iron present in magnetic beads. It is also believed that converting either the bound or unbound nucleic acid separate from their complement may improve the efficiency of the reaction

IDT Exhibit 1020.1056

with sodium bisulfite rendering the nucleic acids fully single stranded. The nucleic acid can be denatured and the unbound nucleic acid collected for subsequent use. In at least one embodiment, the bound nucleic acid, the unbound nucleic acid, or both can be subjected to sodium bisulfite conversion. In embodiments where only one of the bound nucleic acid and the unbound nucleic acid is converted by sodium bisulfite, the unconverted strands can be used as a reference or control sample, as an archive sample, or as another test sample. For example, if the unbound nucleic acid is converted using sodium bisulfite, the bound sample may be kept in its original form for later analysis or testing.

[0071] The converted strands exposed to the methylation conversion agent can be amplified prior to DNA sequencing. The standard nucleic amplification technologies such as PCR, rolling circle amplification, whole genome amplification, LCR and the like can be employed. Primer sites located within the primer-adapters can be used as priming sites for PCR and similar primer based amplification techniques. By suitable placement of the primer binding sites, the first tag region and second tag region can be simultaneously amplified in the same amplification reaction. In embodiments employing partially protected primer-adapters, amplification can be achieved using amplification primers specific for primer binding sites that have been converted by the methylation conversion agent, thereby permitting the preferential amplification of nucleic acids that have been converted by the methylation conversion agent. Amplification primers specific for converted primer binding sites can be used to introduce additional primer binding sites. These additional primer binding sites can be used for, among other things, amplification or sequencing.

[0072] The converted strands can be used as sequencing templates and may be sequenced using DNA sequencing procedures that are well-known to persons skilled in the art. The methods provided here in produce templates for analysis by a wide variety of DNA sequencing methods. Such methods include traditional DNA sequencing techniques employing in electrophoresis, e.g., Sanger sequencing or Maxim and Gilbert sequencing. The templates produced by the methods provided herein can also be sequenced by so-called "next-generation" sequencing techniques that may be amenable to performing large numbers of sequencing reactions in parallel. Such techniques

IDT Exhibit 1020.1057

include pyrosequencing, nanopore sequencing, single base extension using reversible terminators, ligation-based sequencing, single molecule sequencing techniques, and the like, as described in, for example, U.S. Patent Nos. 7,057,056; 5,763,594; 6,613,513; 6,841,128; and 6,828,100; and PCT Published Application Nos. WO 07/121489 A2 and WO 06/084132 A2. Many of the next-generation sequencing techniques employ a clonal amplification step, wherein individual template molecules are amplified in such a way as to maintain separate clones during the amplification. Exemplary of such clonal amplification methods are emulsion PCR (ePCR) and solid phase PCR. The use of suitable adapters for the amplification of templates produced by the methods provided herein may facilitate the use of such clonal amplification techniques as preparation of templates for sequencing.

[0073] Sequencing of the converted strands containing the first and second tag regions may be performed so as to determine the nucleotide sequence of all or part of both tag regions. The converted tag sequence polynucleotide sequences may be difficult to match to a reference sequence in a genomic database because of the presence of a reduced amount of sequence complexity, e.g., in some samples the converted tag sequence will only have three different nucleotide bases due to the conversion of cytosine to uracil, which base pairs with adenosine and thus reads as thymine. The protected tag sequence can, in some cases, be easier to unambiguously match to a reference sequence in the genomic database because of the greater nucleotide base complexity. As the converted tag region and the protected tag region are part of a mate-pair derived from the same genomic fragment, the approximate physical distance in the genome between the 2 tag regions in the mate-pair is known, and thus can be used to help match the tag regions into the reference sequences and to help provide for the assembly of overlapping regions to produce a larger DNA sequence. Accordingly, in at least one embodiment, the protected tag sequence is matched to a genomic database and then the match may be used as an "anchor" (or location of high certainty) to determine the possible location of the converted tag sequence in the genome based, in part, on the approximate physical distance of the tag regions in the mate-pair so as to find a match for the converted tag sequence. It will be appreciated by those skilled in the art that a match between the nucleotide sequence of

IDT Exhibit 1020.1058

the converted tag region and the reference sequence is not necessarily a perfect sequence match, but can take into account some of the changes in nucleotide bases caused by the partial or complete conversion of the bases caused by the methylation conversion agent. Additionally, it will also be understood that a match between the protected tag region and the reference genomic sequence can be other than a match for 100% identity, but can include various SNPs, insertions, deletions, substitutions, and the like. Furthermore, it will be understood that while a given genetic locus can be methylated or unmethylated on a single nucleotide of genomic DNA, preparations of a genomic DNA are derived from multiple cells in a sample, e.g., a tissue sample, and that the some of the genomic DNA can be methylated and some may not be methylated at the same locus within a sample. As noted in U.S. Patent No. 7,112,404, genomic methylation analysis of genomic DNA in a sample does not necessarily yield a simple choice of methylated vs. unmethylated for a given locus; sometimes, a more quantitative answer is required. By using multiple tag sequences from the same genetic locus, i.e., the same or overlapping converted tag regions, a single base position can be interrogated multiple times so as to produce a composite value indicative of the degree of methylation at a given genetic locus in a sample derived from one or more different cells. For example, a tumor sample can comprise identical regions of DNA, but differing in methylation state between the different cells that are with the tissue sample; sequencing such an aggregate of different cells can give data indicative of methylation state that is neither 100% methylated nor 100% unmethylated at the locus of interest.

[0074] Various embodiments of the present teachings also relate to software and computers configured for the implementation of such methods of matching converted tag sequences and protected tag sequences to a database of genomic DNA sequences. The genomic database used comprises genomic data, including in some embodiments the entire genome or genomes of the organism from which the mate-pair library was derived. The nucleotide base sequence information obtained from sequencing the tag regions (or portions thereof) of a mate-pair can conveniently be stored as a data record in a form easily manipulated by an electronic computer. The data record can optionally comprise a value indicative of the approximate physical distance between the tag regions on the genome. However, since in a given genetic

IDT Exhibit 1020.1059

library the approximate physical distance between the tag regions may be essentially the same, the physical distance information can be kept as a separate record. The matching of sequence to genomic DNA database can be achieved by using well-known methods of sequence searching algorithms, e.g., BLAST, Smith-Waterman, and the like.

[0075] Embodiments of the present teachings can be implemented in digital electronic circuitry, or in computer hardware, firmware, software, or in combinations thereof. Apparatus of the present teachings can be implemented in a computer program product tangibly embodied in a machine-readable storage device for execution by a programmable processor; and method steps of the present teachings can be performed by a programmable processor executing a program of instructions to perform functions of the present teachings by operating on input data and generating output. The present teachings can be implemented advantageously in one or more computer programs that are executable on a programmable system including at least one programmable processor coupled to receive data and instructions from, and to transmit data and instructions to, a data storage system, at least one input device, and at least one output device. Each computer program can be implemented in a high-level procedural or object-oriented programming language, or in assembly or machine language if desired; and in any case, the language can be a compiled or interpreted language. Suitable processors include, by way of example, both general and special purpose microprocessors. Generally, a processor will receive instructions and data from a read-only memory and/or a random access memory. Generally, a computer will include one or more mass storage devices for storing data files; such devices include magnetic disks, such as internal hard disks and removable disks; magneto-optical disks; and optical disks. Storage devices suitable for tangibly embodying computer program instructions and data include all forms of non-volatile memory, including by way of example semiconductor memory devices, such as EPROM, EEPROM, and flash memory devices; magnetic disks such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM disks. Any of the foregoing can be supplemented by, or incorporated in, ASICs.

[0076] Other embodiments of the present teachings include methods for analyzing the methylation state of genomic DNA. These methods may be applied to the

IDT Exhibit 1020.1060

mate-pair generation techniques discussed above or used for other forms of methylation analysis that do not involve the creation of mate-pair libraries. One such embodiment includes methods of analyzing the methylation state of genomic DNA in which the genomic DNA is denatured with formamide, rather than sodium hydroxide. Sodium hydroxide is typically used to denature DNA for sodium bisulfite treatment so as to provide for the methylation analysis of DNA. However, strong bases, such as sodium hydroxide, may have unwanted side effects such as depurination of the DNA. The use of formamide as a denaturant has been shown to be effective in permitting bisulfite to efficiently modify genomic DNA for methylation analysis purposes. The use of formamide as a denaturant has also been shown to be effective in permitting bisulfite to efficiently modify genomic DNA obtained from formalin fixed paraffin embedded tissues samples. Formalin fixed paraffin embedded tissues are commonly used to store tissue samples, e.g., as prepared by pathologists.

[0077] In at least one embodiment of the present teachings, the methylation state of the genomic DNA sample can be ascertained by mixing the genomic DNA with formamide whereby a mixture is formed. The mixture can then be heated to a temperature sufficient to denature the DNA, and a bisulfite salt, such as, for example, sodium bisulfite, can be added to the mixture so as to allow the bisulfite to react with the free amines on the cytosine in the DNA, thereby sulfonating the DNA. The DNA can then be desulfonated, thereby converting the non-methylated cytosines to uracils.

[0078] According to at least one embodiment, the formamide solution employed for denaturation in the subject methods can be in the range of 50 to 100% formamide. The formamide can be in an aqueous solution. In at least one embodiment, the method uses formamide solutions having a concentration of at least 50%, such as at least 75%, at least 90%, or at least 95% formamide.

[0079] In at least one embodiment of the present teachings, independent of the use of mate-pair library generation, the DNA for analysis can be present in a gel matrix, such as a polyacrylamide gel. In at least one embodiment, the use of DNA present in a gel matrix may facilitate the ease with which a given technique can be performed and may increase the yield of bisulfite treated DNA because DNA that has been size separated in an electrophoresis separation gel matrix can be bisulfite treated prior to

IDT Exhibit 1020.1061

removal of the DNA from the gel matrix. In at least one embodiment, the bisulfite treated DNA can also be amplified in the gel matrix. Amplification may be achieved by a variety of standard nucleic amplification techniques, such as PCR, rolling circle amplification, and the like. Amplification of nucleic acids with gel matrices is well-known to person of ordinary skill in the art and is described, for example, in U.S. Patent Nos. 6,001,568; 5,958,698; and 5,616,478.

Examples

Example 1

[0080] An embodiment of the subject method as applied to the generation of mate-pair libraries for sequencing using the methods described in PCT Published Application No. WO 06/084132 A2, which is herein incorporated by reference for at least the purpose of describing mate-pair library formation and sequencing by ligation with an emulsion PCR preparation step, is provided by way of example. The figures described herein illustrate the preparation and sequencing of a mate-pair library containing clones having first and second tag regions, wherein one of the tag regions has been protected from conversion by bisulfite and is suitable for amplification by emulsion PCR. In the example shown in Figures 1-12, the mate-pair library was prepared using EcoP151 cuts, which resulted in short mate-pairs.

[0081] Figure 1 is an example of a 2-3kb fragment of genomic DNA. In the figure, adapters A1 and A2 are added by ligation. The cap adapters comprise an EcoP151 restriction endonuclease recognition site.

[0082] Figure 2 shows an adapter-modified genomic DNA circularized by ligation to an internal adapter comprising a biotin on one strand. A sticky end ligation was used to join the adapter modified genomic fragment to the internal adapter. The 5' phosphate on the non-biotinylated strand of the internal adapter was not ligated to the corresponding A2 adapter.

[0083] Figure 3 shows the circular DNA construction linearized by incubation with the restriction endonuclease EcoP15I. The nick N in one strand can be seen at the arrow indicating the relative position on the linear genetic construction. Tag regions T1 and T2 are indicated. Tag regions T1 and T2 are approximately 25-27 bp each.

IDT Exhibit 1020.1062

[0084] Figure 4 shows the linearized fragment incubated with a nick translation enzyme and the conversion resistant nucleotide 5-methylcytosine (5mC). Tag T1 also comprises 5mC.

[0085] Figure 5 shows the location of the 5mCs in one strand after the nick translation reaction. The 5mCs in this figure and the following figures are underlined. The box around segment 501 comprises 5mC at all cytosines and preserves the actual genomic sequence resistant to sodium bisulfite. The segment at 502 has native methylation status.

[0086] Figure 6 shows the addition of the primer-adapters P1-A and P1-B (partially protected primer-adapters) to the linearized fragment. The location of nicks N caused by absence of 5' terminal phosphates on the adapters is also shown.

[0087] Figure 7 shows the removal of the nicks after nick translation of the construct shown in the bottom of figure 6.

[0088] Figure 8 shows the selectively recovered strand, i.e., the strand lacking the biotin.

[0089] Figure 9 shows treatment with the methylation conversion agent, sodium bisulfite. P1-B, adapter A2 and tag T2 were converted by bisulfite to produce A2' and T2', respectively. The internal adapter, P1-A, and tag T1 were 5mC protected.

[0090] Figure 10 shows the addition of P2 adapters to one end of the bisulfite converted construction containing the tag regions T1 and T2. PCR was used to fill in the second strand of the P2 region.

[0091] Figure 11 shows the sequence of the internal adapter, the P1-A/P1-B adapter and the P2-A tail.

[0092] Figure 12 shows the internal adapter, the 5mC P1-A and P1-B adapters, and the P2-A-Tailed library amplification primer used in the process shown in Figures 1-11.

Example 2

Mate-Pair library Generation

Shearing and End-Repair of the Genomic DNA

IDT Exhibit 1020.1063

1) DNA shearing of 45 ug of E.coli DH10B chromosomal DNA was performed by nebulization in 750 ul of 10 mM Tris pH7.5 as follows:
pressure: 10 psi
time: 2 min 30 sec
on ice in Nebulizer (Invitrogen)
After nebulization 92% of initial volume was recovered (approx 41 ug DNA, measured by UV absorbance in NanoDrop). 1 ul was analyzed in Bioanalyzer (Agilent) using DNA 7500 Assay. Sheared DNA had a peak at 2, 950 bp:
2) DNA concentration.
DNA was concentrated by ultrafiltration in Nanosep 30K Omega spin cartridge: Column was loaded with 500 ul of nebulized DNA and spin at 5,000 rcf for 3 min; then the rest was loaded and spun for an additional 4 min. DNA was concentrated to 172 ul (233 ug/ul, UV absorbance, NanoDrop). Thus, 40 ug (98%) of DNA was recovered after ultrafiltration.
3) Repair of DNA Ends and Purification of Sample
Repaired and purified as in SOLiD System Mate-Paired Library Preparation, except 13 ul of End-It Enzyme mix (instead of 10 ul) was used to adjust for higher DNA input (40 ug instead of 30 ug). Combined and mixed the following components: Sheared DNA (40 ug) – 170 ul; 10X End-It Buffer – 30 ul; End-It ATP (10 mM) – 30 ul; End-It dNTPs (2.5 mM) – 30 ul; Nuclease-free water – 27 ul; End-It Enzyme Mix – 13 ul Total: 300 ul. Incubated 30 min at room temperature.
4) Purify the DNA using QIAquick spin columns in the QIAquick Gel Extraction Kit: total of 4 columns were used; DNA was eluted with 25 ul of EB from each column resulting in total of 187 ul of eluate containing 34 ug of DNA.
Methylation of the Genomic DNA EcoP15I Sites: performed as in SOLiD System Mate-Paired Library Preparation except reaction was performed in larger volume to adjust all reaction components to 34 ug DNA input:
1) Methylation reaction:
Sheared, End-Repaired DNA – 187 ul
10X NEBuffer 3 – 35 ul
100X BSA – 3.5 ul

IDT Exhibit 1020.1064

EcoP15I Enzyme (10U/ul) (NEB) – 34 ul
S-adenosylmethionine (32 mM) – 4.2 ul
Nuclease-free water – 86.3 ul
Total: 350 ul

Incubated at 37°C for 5 hours

2) Purified the methylated DNA using 4 QIAquick spin columns. After elution with EB buffer, 23.6 ug of DNA was recovered, as measured by UV absorbance (NanoDrop). Ligated the EcoP15I CAP Adapters. Ligated as in SOLiD System Mate-Paired Library Preparation. To ligate CAP adapters to 14.4 pmoles DNA in sample 1440 pmoles of adapter were needed (28.8 ul of 50 pmole/ul CAP stock)

1) Ligation reaction:
DNA – 115 ul
2X NEB Quick ligase buffer – 150 ul
NEB Quick Ligase – 8 ul
CAP adapter (ds)(50 pmoles/ul) – 28.8 ul
Total 301.8 ul

Incubated at room temperature for 10 min.

2) Purified DNA using three QIAquick columns, eluted with 30 ul of EB per column. Pooled eluates.

Size-selection of DNA with 1% Agarose Gel
Size-selected as in SOLiD System Mate-Paired Library Preparation. The DNA band of approximately 3 kb (tight size selection) was excised; DNA was extracted from agarose gel using QIAquick Gel Extraction Kit. DNA was eluted from column in 120 ul of EB and analyzed in BioAnalyzer (Agilent) using DNA 7500 Assay:
Mean peak size was found to be at 2845 bp (2.8 kb) (see, for example, Figure 1 above).

IDT Exhibit 1020.1065

DNA concentration was measured by UV absorbance (NanoDrop): 41.7 ng/ul. Thus total 41.7 ng/ul X 106 ul = 4.42 ug DNA was recovered after this step.

DNA Circularization

Circularized as in SOLiD System Mate-Paired Library Preparation, except modified internal adapter, NonPhosIA, was used to generate a nick after circularization by ligation. Preparation of the NonPhosIA (SEQUENCE ID NO: 1) which was the same DNA sequence as per the SOLiD protocol, but no 5'P:

Internal adapter, bottom strand without a 5' P

NonPhosIAb 5' GGCCAAGGCGGATGTACGGT (SEQUENCE ID NO: 1)

1. Prepared 1 mM stock of special oligo NonPhosIAb in Low TE buffer.

2. Mixed equal volumes of 1 mM oligonucleotides Top strand normal SOP (biotinylated) internal adapter and NonPhosIAb. Added enough 5X Ligase buffer for a final concentration of 1X Ligase buffer.

Preparation of 200 uL of 50 uM ds-adapter in 1X Invitrogen Ligase buffer

Mix:

12.5 uL of the 800 uM biotinylated internal adapter

12.5 uL of the 800 uM modified bottom strand Internal adapter minus a 5'Phos

40 uL of 5X Ligase buffer

135 uL of water

[12.5 X 10-6 X 800 X 10-6 = .00000001 which divided by 200 uL = .00005 or 50 uM]

3. Hybridized the oligonucleotides by running the following program on a PCR machine:

Temperature (°C)	Time (min)
95	5
72	5
60	5

IDT Exhibit 1020.1066

50	3
40	3
30	3
20	3
10	3
4	forever

Note: For the 200 uL total volume, it was divided into two equal portions (100 uL) and the above thermalcycling program was followed.

To obtain 95% of circularization efficiency, 4.42 ug of DNA was diluted during circularization reaction to approximately 2.1 ng/ul.
There were 2.34 pmoles of DNA in 4.42 ug of sample of 2.8 kb (0.53 pmoles of DNA/ug X 4.42 ug = 2.34 pmoles)
Total of 7.02 pmoles of internal adapter were needed (2.34 pmoles X 3 = 7.02), or 3.5 ul of internal adapter stock (2 pmoles/ul).

1) Ligation reaction was set:
DNA (4.4 ug) – 106 ul
2X NEB Quick Ligase Buffer – 1100 ul
NonPhosIA internal adapter (ds) (2 pmoles/ul) – 3.5 ul
Quick Ligase (NEB) – 55 ul
Nuclease-free water – 935.5 ul
Total: 2200 ul

Incubated 10 min at room temperature.

2) Purified the DNA using QIAquick column. Eluted 2x30 ul of EB.

3) Treated DNA with Plasmid-Safe ATP-dependent DNase:
DNA – 60 ul

IDT Exhibit 1020.1067

25 mM ATP – 5 ul
10X Plasmid-Safe Buffer – 10 ul
ATP-dependent Plasmid-Safe Dnase (10U/ul) – 1.5 ul
Nuclease-free water – 23.5 ul
Total: 100 ul

Incubated 40 min at 37°C, followed by 20 min at 70°C

4) Purified DNase treated circularized DNA using QIAquick column. Eluted DNA with 40 ul of EB. Quantitated DNA by UV absorbance (NanoDrop): 7.9 ng/ul. Total: 304 ng of circularized DNA.

EcoP15I Digestion of Circularized DNA
Digestion as in SOLiD System Mate-Paired Library Preparation, except after EcoP15I digestion step, DNA was cleaned up using ultrafiltration device instead of heat inactivation of enzyme. Heat inactivation was avoided to prevent strand separation, since one of the "circles" of the ds construct was "nicked" due to use of the Non-phosphorylated-internal adapter (NonPhosIA).

1) EcoP15I digestion reaction:
Circularized DNA (304 ng) – 38 ul
10X NEBuffer 3 – 10 ul
100X BSA – 1 ul
10 mM Sinefungin – 1 ul
10X ATP – 20 ul
EcoP15I (10U/ul) – 1.5 ul (5 U per 100 ng of 2-6 kb long DNA)
Nuclease-free water – 28.5 ul
Total: 100 ul

Incubated at 37°C overnight. Then added additional 1 ul 10 mM Sinefungin, 2 ul 10X ATP, and 0.5 ul EcoP16I and continued incubation for additional 1 hour at 37°C.

IDT Exhibit 1020.1068

2) Purified DNA using Microcon 10 ultrafiltration spin device. Reconstituted in 100 ul of NEBuffer 2.

Nick-translation

1) Assembled on ice the nick-translation reaction:
DNA in NEBuffer 2 – 100 ul
5mC-dNTP mix (25 mM each) – 1.5 ul
E.coli DNA Polymerase I (10U/ul) – 2 ul

Incubated 30 min at 16°C

2) Purified the nick-translated DNA with the Qiagen MinElute Reaction Cleanup kit. Eluted in 40 ul EB. Ligation of partially methylated adapter (SEQUENCE ID NO: 2) (only one adapter was ligated to both ends; adapter has one strand with 5mC). The 5mC positions are underlined:
5mC-P1-A (ss): 5'CCA CTA CGC CTC CGC TTT CCT CTC TAT GGG CAG TCG GTG AT 3' Length: 41 (SEQUENCE ID NO: 2)

1. Prepared 800 uM stock of special oligo 5mC-P1-A.

2. Prepared 1 mM (1000uM) stock of Normal SOP adapter P1-B in Low TE buffer .

Preparation of 200 uL of 50 uM ds-adapter in 1X Invitrogen Ligase buffer
Mixed:
12.5 uL of the 800 uM 5mC-P1-A
12.5 uL of the 1000 uM P1-B
40 uL of 5X Ligase buffer
135 uL of water

IDT Exhibit 1020.1069

[12.5 X 10-6 X 800 X 10-6 = .00000001 which divided by 200 uL = .00005 or 50 uM]

3. Hybridized the oligonucleotides by running the following program on a PCR machine:

Temperature (°C)	Time (min)
95	5
72	5
60	5
50	3
40	3
30	3
20	3
10	3
4	forever

Note: For 200 uL total volume, it was divided into two equal portions (100 uL), and the thermalcycling program was followed.

After EcoP15I digestion, 304 ng of circularized DNA was reduced approximately 29 times. Thus, there were 0.01 ug DNA available for linker ligation. This was 0.01 ug x 17.8 pmoles = 0.178 pmoles DNA available for ligation. 0.178 pmoles X 60 = 10.68 pmoles adapter needed, or 0.22 ul of 50 uM adapter

1) Ligation reaction:
Nick-translated DNA – 38 ul
5mC-P1-A/P1-B adapter (50 uM) – 0.44 ul
2X Quick Ligase Buffer – 50 ul
NEB Quick Ligase – 2.5 ul
Nuclease-free water – 9 ul

IDT Exhibit 1020.1070

Incubated 10 min at room temperature. Purification of library molecules from side products (Streptavidin- Biotin pull out) was performed as in SOLiD System Mate-Paired Library Preparation. Nick-translation of DNA was performed as in SOLiD System Mate-Paired Library Preparation.

1) Nick-translation reaction:
Adapter ligated DNA-Bead complex - 37.7 ul
GeneAmp dNTP Blend (100 mM) – 0.8 ul
DNA Polymerase I (10U/ul)
Total: 40 ul

Incubated at 16°C for 30 min.

2) Washed DNA-Bead complex using magnet in EB. Resuspended DNA-Bead complex in 40 ul EB buffer.

Removal of Biotinylated Strand and Bisulfite Convertion
The last step of the library preparation before the bisulfite conversion was the capture of the fragments with the biotin on magnetic beads. Only 1-2 ng of fragments was estimated to be present. There were changes to the bisulfite conversion that were used:

- Due to the low concentration of DNA for bisulfite conversion, a carrier DNA was spiked into the bisulfite conversion, DH10b, and was not denatured, so it remained double stranded through the bisulfite conversion
- The non-biotinylated strand was eluted with base denaturation from the magnetic beads according to the protocol below, immediately prior to bisulfite conversion
- Because the non-biotinylated strand was eluted as single stranded, no further steps were needed for denaturation prior to bisulfite conversion - the carrier DNA was deliberately left double stranded
- Incubation in bisulfite at 50 degrees for 3 hours was likely sufficient due to short, single stranded fragments of DNA and not large complex genomes with secondary structure.

IDT Exhibit 1020.1071

- Microcon 10 as used for the purification to capture the small mate-pair library fragments

Elution of the non-biotinylated strand for the magnetic beads
Removed the buffer from the beads. Resuspended the beads in 20 µl of freshly prepared 0.15 M NaOH. Incubated at room temperature for 10 minutes. Put the tube in magnet stand for 1-2 minutes and transferred the supernatant to a new tube. The supernatant contained the non-biotinylated DNA strand. The 20 uL of 0.15 M NaOH solution containing the single-stranded library fragments was mixed with 100 uL of Zymo (reconstituted) CT conversion reagent. One uL of a 300 ng/uL solution of DH10B was added to supply a carrier DNA. No attempt was made to denature the carrier DNA. The reaction was incubated at 50 degrees for 3 hours. The bisulfite reaction was then purified with a Microcon 10 device following the steps below.

The Microcon 10 washes were as follows:

1. Diluted each bisulfite reaction (if multiples were done) with 100 uL of water. Transferred each diluted reaction to a Microcon 10 and centrifuge at 7000 rpm for 30-40 minutes
2. Removed flow-through and added 100 uL of water to the upper chamber of the M-10 and centrifuge for ~30 min at 7000 rpm
3. Repeated step 2.
4. Removed flow-through and add 100 uL of 0.1 M NaOH, let sit for 5 minutes at RT, and centrifuged at 7000 rpm for ~30 min.
5. Removed Flow-through, added 100 uL water, centrifuged for ~30 Min. at 7000 rpm.
6. Reconstituted the bisulfite converted library in TE (25-50 uL, depending of desired concentration)

Library Amplification

IDT Exhibit 1020.1072

1) PCR with modified P1 primer

Pre-emulsion Library amplification primer with P2-A tail (SEQUENCE ID NO: 3)
P2AtailbisP1 5'
CTGCCCCGGGTTCCTCATTCTAACCACTACACCTCCACTTTCCTCTCTATAAA
(SEQUENCE ID NO: 3)

Note: The P2 tail on this Bisulfite-P1 primer sequence (which is the reverse compliment to the bisulfite converted P1B sequence) introduced the P2 sequence recognized by the beads for ePCR according to the SOLiD protocol.
The two primers for library amplification were therefore the "normal" P1 primer and the bisulfite converted P1 primer.

Bisulfite converted library - 33 ul
P2A-tailbisP1 primer (50 uM) – 1 ul
Library PCR Primer 1 (50 uM) – 1ul
10X PCR Gold Buffer w/o Mg++ - 5 ul
MgCl2 (25 mM) – 3 ul
dNTP mix (25 mM each) – 0.4 ul
AmpliTaq Gold (10U/ul) – 2 ul
Nuclease-free water – 4.6 ul
Total: 50 ul

Thermal profile:
9 min at 95°C;
95°C 30 seconds, 55°C 30 seconds, 70°C 5 min for 2 cycles

2) Trial-PCR performed as in SOLiD System Mate-Paired Library Preparation

3) Large-scale PCR performed as in SOLiD System Mate-Paired Library Preparation

IDT Exhibit 1020.1073

Large-scale PCR was performed for 40 cycles. DNA was cleaned up with Qiagen MinElute column and eluted with EB buffer

Example 3

Fragment Library preparation

[0093] Human gDNA (10 μg) from a male individual of Yoruban ancestry [Coriell cell repository (http://locus.umdnj.edu): NA 18507] was sheared to give fragments (~60-90 bp) using a Covaris S2 system (Covaris, Woburn, MA, USA) as described in Chapter 1 of the SOLiD System 2.0 user guide (Applied Biosystems, Foster City, CA, USA). The sheared DNA was purified with a MinElute Reaction Cleanup kit (Qiagen, Valencia, CA, USA) as described in the user guide, and then quantified by UV using a NanoDrop ND 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). An End-It DNA end-repair kit (Epicentre Biotechnologies, Madison, WI, USA) was used according to manufacturer instructions to convert DNA with damaged or incompatible 5'- or 3'-protruding ends to 5'-phosphorylated, blunt-end DNA suitable for blunt-end ligation. Following purification of the resultant blunt-end fragments with aforementioned MinElute columns and then quantification by UV, as described above, the required volume of pre-annealed double-stranded adapters needed for ligation was calculated as described in the SOLiD user guide referenced above. The top strand (P1-A) (SEQUENCE ID NO: 4) of the double-stranded P1 adapter was synthesized (TriLink Biotechnologies, San Diego, CA, USA) with 5mC in place of C to protect the adapter from modification during bisulfite conversion. P1 and P2 adapter sequences were as follows wherein 5mC is underlined.

[0094] (Top strand) 5mC-P1-A: 5'CCA CTA CGC CTC CGC TTT CCT CTC TAT GGG CAG TCG GTG AT3' (SEQUENCE ID NO: 4)

[0095] (Bottom strand) P1-B: 3'TT GGT GAT GCG GAG GCG AAA GGA GAG ATA CCC GTC AGC CAC TA5' (SEQUENCE ID NO: 5)

[0096] (Top strand) P2-A: 3'TCT CTT ACT CCT TGG GCC CCG TC5' (SEQUENCE ID NO: 6)

[0097] (Bottom strand) P2-B: 5'AGA GAA TGA GGA ACC CGG GGC AGT T3' (SEQUENCE ID NO: 7)

IDT Exhibit 1020.1074

[0098] The single-stranded adapter-pairs of oligonucleotides 5mC-P1 and P1-B, and P2-A and P2-B were pre-annealed to form double-stranded adapters. During adapter ligation, only the top adapter strands were joined to the 5'- phosphorylated ends of the DNA fragments. After purification of the ligation products with aforementioned MinElute columns, the bottom adapter sequence was filled-in by extension with DNA polymerase during nick-translation. 2'-deoxycytidine-5'-triphosphate (dCTP) in the conventional mixture of four dNTPs was replaced with 5-methyl-2'-deoxycytidine-5'-triphosphate (5mC-dNTP) (TriLink Biotechnologies). This 5mC-dNTP containing mixture was prepared at 25 mM for each of the four nucleotides using 100 mM stock solutions that included commercially available dNTPs of A, G and T (GE HealthCare-Amersham Biosciences, Pittsburgh, PA, USA). Following nick-translation, 75 μL of the 80-μL reaction was electrophoresed using a 3 % cross-linked agarose gel (Bio-Rad Laboratories, Hercules, CA, USA) and fragments having the desired size-range (~150-200 bp) were excised and then purified with aforementioned MinElute columns. The resultant Yoruban SOLiD fragment-library suitable for bisulfite conversion was quantified by UV as described above, and found to be 12.1ng/μL or a total yield of 1.21 μg.

Semi-quantitative PCR to monitor Bis-PAGE

[0099] Preliminary studies of denaturing DNA embedded in a 6% cross-linked PAGE-slice (see below) compared formamide to NaOH by employing ~50-ng portions of an *Escherichia coli* (*E. coli*) DH10B genomic library) for construction of a SOLiD-60-90 bp fragment-library having 5mC-protected ends. The following four conditions were studied: (A.) 25 uL of formamide, (B.) 0.4 M NaOH prepared by us, (C.) NaOH ~0.4 M supplied as M-Dilution Buffer in the EZ DNA Methylation-Direct kit (Zymo Research) and (D.) ~0.2M NaOH as M-Dilution Buffer; denaturation with formamide was performed at 95 °C for 5 min. whereas denaturation with NaOH was performed at 37 °C for 15-20 min. Conditions (C.) approximated the commercial kit bisulfite-reaction conditions ignoring the volume of the PAGE-slice whereas condition D approximated the commercial kit bisulfite-reaction conditions taking into account the ~25-μl volume of the PAGE-slice. Following denaturation, 100 μL of freshly prepared sodium bisulfite

IDT Exhibit 1020.1075

obtained as CT Conversion Reagent (Zymo Research, Orange, CA, USA) was added to each of conditions (A.)-(D.), and the resultant PAGE-slices were incubated for 8 hr at 50 °C. Following post-bisulfite washes and desulfonation, each PAGE-slice was subjected to pre-emulsion PCR, all as described below. The number (n) of PCR cycles necessary for an amplicon-band to be visibly detected using FlashGel (Lonza, Basel, Switzerland) was found to be ~2 less for the library denatured with formamide. This approach was applied to an analogous 5mC-end-protected Yoruban fragment-library at 100-, 10- and 5-ng starting amounts, which gave n = 17, 22 and 22, respectively, thus indicating a rough, semi-quantitative, inverse relationship between starting amounts of fragment-library and values of n that appeared to be insensitive to a 2-fold difference between 10- and 5-ng. Despite the limited sensitivity of this approach, it was routinely used for monitoring various pilot experiments including 8 hr vs. overnight incubation with bisulfite at 50 °C, which indicated substantial loss of amplifiable fragment-library DNA during overnight conditions.

Solution bisulfite conversion

[00100] A 25-μL aliquot containing ~280 ng of the partially 5mC-end-protected Yoruban SOLiD fragment-library prepared as described above was bisulfite converted according to our reported [Anal Biochem 326 (2004) 278-80.] procedure except for the following modifications. Denaturation was performed by mixing the 25-μL aliquot of the library with 25 μL of highly deionized formamide (Hi-Di Formamide) (Applied Biosystems) and then heating at 95 °C for 5 min. To the resultant solution was added freshly prepared sodium bisulfite obtained as CT Conversion Reagent (Zymo Research), and the reaction mixture was incubated in a 96-well thermal cycler (Applied Biosystems) for 8 hr at 50 °C followed by a programmed hold at 4 °C overnight. A similarly prepared aliquot was incubated overnight for 17 hr at 50 °C. Each bisulfite-converted fragment-library was purified as reported [Anal Biochem 326 (2004) 278-80.] except for the following modifications. A Microcon 10 spin-column (Millipore, Billerica, MA, USA) was used in place of a Microcon 100 spin-column in order to retain the presently described fragment-libraries that are much smaller in size compared to conventionally processed and bisulfite-converted gDNA. In addition, centrifugation

IDT Exhibit 1020.1076

speed and time were increased to 7000 rpm and 45 min per wash and for the desulfonation step. Each bisulfite-converted SOLiD fragment-library was recovered in a final volume of 30 μL of sterile buffer (10 mM Tris-HCl, 1.0mM EDTA, pH 7.2) (Teknova, Hollister, CA, USA).

Bis-PAGE bisulfite conversion

[00101] For comparison of results obtained for solution bisulfite conversion described above, bisulfite conversion was performed directly in a gel-band from PAGE according to the following protocol referred to herein as Bis-PAGE. An aliquot containing ~100 ng of the final preparation of partially 5mC-end-protected Yoruban SOLiD fragment-library obtained as described above was electrophoresed into a 6 % cross-linked DNA Retardation Gel (Invitrogen, Carlsbad, CA, USA), and the band containing the library was excised using a razor blade. The PAGE slice was then cut into two, approximately equal, halves such that each piece was then small enough to fit into the bottom of a single MicroAmp tube (Applied Biosystems) and be fully immersed upon addition of 25 μL of Hi-Di Formamide (Applied Biosystems). Each ~50-ng portion of the original fragment-library embedded in the PAGE slice was heated in a 96-well thermal cycler (Applied Biosystems) at 95 °C for 5 min to denature the library fragments followed by cooling to 30 °C to allow addition of 100 μL of freshly prepared CT Conversion Reagent (Zymo Research) and then heating at 50 °C. One of these two samples was heated for 8 hr with a programmed hold at 4 °C until the following morning, and the other sample was incubated at 50 °C overnight for 17 hr. Bisulfite reagent was removed by pipet from each Bis-PAGE sample, and then 180 μL of molecular biology-grade water (Sigma, St. Louis, MO, USA) was added, pipeted up and down several times and then removed. This step was repeated and third wash with fresh water included a 5-min wait before removal, and was repeated in a final, fourth wash. Desulfonation of each embedded Bis-PAGE sample was performed using 180 μL of 0.1 N NaOH that was allowed to stand for 15-20 min before removal. Each still fully intact PAGE slice was then washed twice with 180 μL of water, without a wait step, followed by two washes that each included a 5-min wait time. Each resultant PAGE slice containing embedded bisulfite-converted fragment-library was then

IDT Exhibit 1020.1077

immediately used for library amplification prior to emulsion-PCR (pre-emulsion PCR) as described below.

Library amplification (pre-emulsion PCR)

[00102] The following standard P1 and P2 primers were used for SOLiD fragment-library amplification according to the SOLiD System 2.0 user guide (Applied Biosystems).

[00103] P1: 5'CCA CTA CGC CTC CGC TTT CCT CTC TAT G3' (SEQUENCE ID NO: 8)

[00104] P2: 5'CTG CCC CGG GTT CCT CAT TCT3' (SEQUENCE ID NO: 9)

[00105] Note that, following bisulfite conversion, double-strand DNA is rendered single stranded and is no longer complementary. Only the strand with bisulfite-resistant ends 5mC-P1-A and 5mC-P2-B is amplified during PCR.

Amplification of bisulfite-converted libraries in solution

[00106] The master mix specified in the SOLiD System 2.0 user guide (Applied Biosystems) was supplemented as follows with additional AmpliTaq Gold DNA Polymerase to ensure "reading" of U, i.e., deaminated C. For each 1X reaction, 50 μL of Platinum SuperMix (Invitrogen) was mixed with fragment-library PCR primers P1 and P2 (1 μL of 50 μM), 3 μL of the bisulfite-converted DNA (that was recovered as described above in 30 μL of 10 mM Tris-HCl, 1.0mM EDTA, pH 7.2 sterile buffer) and 0.25 μL of AmpliTaq DNA Polymerase, LD (Applied Biosystems). This 1X PCR reaction was scaled-up 8-fold and dispensed into eight separate tubes to accommodate ~24 μL of the solution-based bisulfite-converted fragment-library. The 8-hr and overnight bisulfite-conversion samples were processed identically. Thermal cycling as described in the SOLiD System 2.0 user guide (Applied Biosystems) was interrupted periodically (3, 5, 8 and 13 cycles) and 2-μL aliquots of the PCRs were analyzed by FlashGel (Lonza) until amplicon was detected. Thermal cycling was stopped after 13 cycles and PCRs were purified using an AMPure kit (Agencourt, Beverly, MA, USA) and then quantitatively characterized using a Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). A 1-μL aliquot (22 ng or 35 ng for the 8-hr and

IDT Exhibit 1020.1078

overnight samples, respectively) was removed for capillary electrophoretic fragment analysis and QC by Sanger sequencing, and the remainder was saved for emulsion-PCR and then SOLiD sequencing.

Amplification of Bis-PAGE libraries

[00107] Each thoroughly washed and desulfonated Bis-PAGE slice from 8-hr or overnight heating at 50 °C was PCR-amplified in the same MicroAmp tube used for the bisulfite conversion, as described above, using AmpliTaq Gold DNA Polymerase-supplemented conditions identical to those specified in the preceding section on amplification of the bisulfite-converted library in solution. A 2-μL aliquot of each sample was analyzed by FlashGel every other cycle. PCR thermal cycling was stopped after 17 cycles and the concentration of the amplified library was determined using a Bioanalyzer 2100 following purification using an AMPure kit.

Size-analysis of smPCR amplicons from bisulfite-converted fragment-libraries

[00108] A ~1-ng/μL aliquot of each minimally amplified library obtained as described in the preceding sections was serially diluted to give 1-mL of a working solution that was ~1 copy/μL. The following components were scaled for distribution into multiple 96-well plates for 5-μL PCR: common primers [0.25-μL FAM-short-P1 primer, 0.25-μL normal-P2 primer, 5-μM each; see sequences below incorporating 6-FAM DYE (Applied Biosystems)] were combined with 1.0 μL of the ~1 copy/μL bisulfite-converted amplified library, 0.5-μL AmpliTaq Gold 10X buffer, 0.4-μL dNTP (2.5 mM each), 0.4-μL $MgCl_2$ (25 mM), 0.1-μL AmpliTaq Gold DNA Polymerase (5U/μL), 1.6-μL molecular biology- grade water and 0.5-μL bovine serum albumin-glycerol solution [prepared by mixing 250 μL of a 20 mg/mL bovine serum albumin solution (Sigma, St. Louis, MO, USA), 700 μL of molecular biology-grade water (Sigma, see above) and 50 μL of Biology-Certified Glycerol (Shelton Scientific-IBI, Peosta, IA, USA)]. Thermal cycling conditions were as follows: 5 min at 95 °C (to activate the hot-start polymerase), 40 cycles at 95 °C /30 sec, 60 °C /2 min, 72 °C /45 sec; hold at 4 °C.

IDT Exhibit 1020.1079

[00109] FAM-short-P1: 5'(6-FAM)CGC CTC CGC TTT CCT CTC TAT G3' (SEQUENCE ID NO: 10)

[00110] normal-P2: 5'CTG CCC CGG GTT CCT CAT TCT3' (SEQUENCE ID NO: 11)

[00111] A 0.7-μL aliquot of the PCR reaction was added to 11 μL of Hi-Di Formamide (Applied Biosystems) containing 10% ROX 500 size-standard (Applied Biosystems), and heated at 95 °C for 5 min to denature the amplicon. Fragments were analyzed at 60 °C on a 96-capillary 3730xl DNA Analyzer (Applied Biosystems) using a 50-cm capillary array, POP 7 polymer and GeneMapper Software for data collection with run module GeneMapper50_POP7_1 with dye set Any5Dye (all from Applied Biosystems).

Sanger sequencing

[00112] In preparation for sequencing, unreacted dNTPs and primers were eliminated by addition of 1 μL of ExoSAP-IT (USB, Cleveland, OH, USA) to each PCR sample (after removing the 0.7-μL aliquot for fragment analysis) and incubation at 37 °C for 30 min. This was followed by heat-denaturation at 80 °C for 15 min and then storage at 4 °C. The resultant PCR samples were each diluted with 25 μL of water and a 0.5-μL aliquot of the diluted sample was used in BigDye Terminator v1.1 (Applied Biosystems) sequencing by adding 4-μL BigDye Terminator Ready Reaction Mix, 0.5 μL of unlabeled short-P1 primer, 5'CGC CTC CGC TTT CCT CTC TAT- G3' (SEQUENCE ID NO: 12) (5.0 μM) and 5 μL of water. Cycle sequencing employed 96 °C/1 min, followed by 25 cycles of 96 °C/10 sec, 50 °C/4 min and hold at 4 °C. Unincorporated BigDye Terminator and unused primers were removed using the Big Dye XTerminator Purification kit (Applied Biosystems) following manufacturer instructions. Sequencing was performed on a 96-capillary 3730xl DNA Analyzer (Applied Biosystems)

Results and discussion

[00113] Representative commercial kits and protocols using DNA-binding matrices for recovery have been shown to afford mostly 4.0-0.5 kb converted-DNA, and could thus lead to substantial loss of bisulfite-converted SOLiD fragment-libraries discussed above. Another concern was the possibly accelerated reannealing (driven by

IDT Exhibit 1020.1080

common-adapter sequences) during bisulfite treatment that could prevent complete bisulfite conversion, given the demonstrated requirement for single-stranded regions during the C-sulfonation step.

[00114] Nick-translation with 5mC-dNTP was performed in solution, rather than directly in the PAGE gel-slice, in order to better assess completeness of overall C→T conversion that was mentioned above as an acknowledged common source of error in bisulfite-based DNA methylation analyses. The influence of embedding DNA in a PAGE-slice during bisulfite conversion (Bis-PAGE) and subsequent PCR was compared to free-solution reactions in parallel experiments using aliquots of the same SOLiD fragment-library. A 100-ng aliquot of the fragment-library was electrophoresed into a 6% polyacrylamide gel, and the excised PAGE-slice was cut in half so that ~50-ng portions of the library were bisulfite converted in PAGE (Bis-PAGE) for either 8 hr or 17 hr ("overnight") at 50 °C. Free-solution bisulfite conversion of the same SOLiD fragment-library preparation was performed under each of these reaction conditions using larger, i.e., 240-ng, portions to compensate for expected lower recovery of relatively short fragment-library DNA. Bis-PAGE and free-solution bisulfite treatments bypassed conventional use NaOH to denature DNA by employing formamide, based on recent capillary sequencing results demonstrating that formamide denaturant gave more complete overall C→T conversion compared to NaOH. In this regard, it should be noted a commercially available, highly deionized grade of formamide was used to minimize potential problems due to ionic impurities known to be present in other common grades of formamide. Microcon 10 spin-columns having a lower molecular-weight cutoff range were used in place of previously reported Microcon 100 spin-columns as another means of increasing recovery of relatively short, ~150-200 bp converted DNA library-fragments. Appropriate spin-columns thus bypass use of typical DNA-binding matrices that have been found to provide mostly 4.0-0.5 kb converted-DNA.

[00115] *Semi-quantitative PCR comparison of denaturation with formamide vs. NaOH during Bis-PAGE*

[00116] Preliminary studies of denaturing ~50-ng of SOLiD fragment-library embedded in a 6% cross-linked PAGE-slice compared formamide at 95 °C for 5

IDT Exhibit 1020.1081

min with either 0.4 M NaOH or 0.2 M NaOH both at 37 °C for 15-20 min. This pre-denaturing was followed by addition of a solution of sodium bisulfite and then incubation at 50 °C for 8 hr. After sequential removal of sodium bisulfite, washing, desulfonation with NaOH and final washing, each PAGE-slice was subjected to PCR. The number (n) of PCR cycles necessary for an amplicon-band to be visibly detected using FlashGel (Lonza, Basel, Switzerland) was found to be ~2 less for the library denatured with formamide. An inverse relationship between values of n and amounts of starting fragment-library DNA indicates several-fold less PCR-amplifiable DNA in the case of NaOH, which could be due to degradation and/or loss of embedded DNA. Loss of PCR-amplifiable fragment-library DNA was also found for formamide during 50 °C incubation with bisulfite overnight vs. for 8 hr. In this regard, it should be noted that others have previously reported that heating DNA in formamide (without bisulfite) under more forcing conditions (e.g. 110 °C, 10 min) than those described herein leads to a low level of cleavage of DNA that was suggested as a chemical sequencing method. In view of this competing side-reaction, any protocol for denaturing and bisulfite conversion of DNA using formamide must avoid excessive heating.

[00117] The presently described Bis-PAGE protocol was developed as part of a streamlined sample-prep workflow to enable, for the first time, bisulfite sequencing of genome-wide SOLiD fragment-libraries that will be reported elsewhere. Completeness of overall C→T conversion was unambiguously established by smPCR for capillary sequencing as discussed below. Feasibility studies of extending Bis-PAGE to include conventional gDNA samples was performed. As a representative example, it has been determined that 1 μL containing 50 ng of commercially available (Applied Biosystems) gDNA (CEPH 13470-02) spotted onto a 6% cross-linked PAGE-slice and then air-dried for 5 min could be successfully subjected to the Bis-PAGE protocol described herein for a SOLiD fragment-library. This offers a simplified procedure relative to conventional methods or spin-columns or agarose-embedding using pre-denaturing in NaOH followed by formation of agarose beads in oil.

Fragment-library amplification (pre-emulsion PCR)

[00118] Comparison of bisulfite-converted SOLiD fragment-libraries involved PCR amplification using a limited number of cycles, as performed for

IDT Exhibit 1020.1082

conventional, i.e. non-bisulfite-converted SOLiD fragment-libraries, prior to emulsion-PCR of single molecules for attachment of "clonal" amplicon on beads. During limited amplification of a bisulfite-converted SOLiD fragment-library, the PCR reaction was supplemented with AmpliTaq LD, and the 5mC-protected universal primer-binding site in all members of the library remained unchanged during bisulfite conversion of genomic fragments of interest. Consequently, universal primers for this limited-PCR step amplify library-fragment regardless of whether bisulfite conversion of fragments was complete or not. It was determined to QC bisulfite-treated fragment-libraries derived from either free-solution reaction or Bis-PAGE by measurement of three variables. (1.) Yield was determined by relative recovery, as reflected by semi-quantitative limited PCR, while (2.) sequence and amplicon-size were each accurately determined by established capillary electrophoresis methods. Aliquots of limited-PCR samples were removed at two-cycle intervals for analysis by FlashGel to assess whether an amplicon band could be visually detected. This semi-quantitative discontinuous means of measuring a cycle threshold-like value ("Ct") akin to real-time PCR Ct-values was estimated to have a sensitivity of roughly ±2 "Ct" units. Free-solution bisulfite-conversion reactions were distributed into multiple wells at 28 ng of fragment-library/well assuming (for the sake of simplicity) 100% recovery, whereas Bis-PAGE samples (still embedded in PAGE-slices) had ~50 ng of bisulfite-converted fragment-library DNA assuming (for the sake of simplicity) 100% recovery. A representative well of free-solution fragment-library gave "Ct" = 13, whereas the Bis-PAGE fragment-library gave "Ct" = 15, which are roughly comparable values considering the assumptions about recovery and the estimated sensitivity of ±2 "Ct" units. In any case, these roughly comparable "Ct" values indicated that loss of short (~150-200 bp) library-fragments due to diffusion from 6% cross-linked PAGE-slices was insignificant in this first demonstration of Bis-PAGE workflow. Retention of these fragment-libraries was also demonstrated in separate experiments of the type described above starting with smaller amounts of fragment-library, i.e. 10- and 5-ng of input DNA for Bis-PAGE at 50 °C for 8 hr albeit with "Ct" = 22, which was consistent with less starting material for PCR. QC of resultant amplicons by capillary methods for size-analysis and sequencing are respectively discussed in the next two sections.

IDT Exhibit 1020.1083

QC of single-molecule library-fragment amplicons by capillary electrophoretic size-analysis

[00119] Bisulfite sequencing commonly involves capillary sequencing of bisulfite-converted DNA that has been either cloned to characterize individual molecules or amplified by PCR to characterize ensemble-average molecules. To overcome known sequence-bias during cloning or PCR, and to bypass tedious cloning entirely, recent publications have introduced smPCR for bisulfite sequencing. It was noted in the recent publications that a requirement for successful smPCR is very low occurrence of non-template-dependent amplification commonly referred to as primer-dimer. This problem is exacerbated during smPCR wherein primer concentrations vastly exceed that of a single-molecule in a PCR-well, is not entirely mitigated by use of hot-start reagents, and likely requires optimization of primer sequences. Applicants have found that during troubleshooting bisulfite sequencing that structures of primer-dimers can encompass molecules significantly longer than that of the starting PCR primers. Such primer-dimer related species formed after bisulfite conversion of the presently described fragment-library could therefore be mistaken for actual members of the fragment-library and thus incorrectly indicate incomplete C→T conversion. QC of all smPCRs by capillary electrophoretic sizing of all amplicons that was detected via use of a fluorescently labeled PCR primer, taking advantage of readily available and widely used GeneScan size-standards having a different fluorescent label. These size-standards can therefore be added to all smPCR wells prior to capillary electrophoresis, and interpolated sizes of PCR amplicons precisely calculated by automated GeneMapper software.

[00120] The size-range of the SOLiD fragment-library described herein was ~150-200 bp. Serial dilutions of aliquots of amplified fragment-libraries derived from various reaction conditions were carried out based on UV quantification of the starting amount of DNA in each case. For example, the calculated number of molecules in 1 μL of amplified fragment-library with a starting concentration equal to 2 ng/μL and an assumed ensemble-average fragment-size of 150 bp is 1.3×10^{10} copies, using an average of 600 g/mole per bp for double-stranded DNA. Serially diluting 1 μL into 1 mL provided 13 molecules/μL after 3 of such serial dilutions for further dilutions

IDT Exhibit 1020.1084

to in the single-molecule regime for pilot smPCRs ("range-finding"), prior to carrying out a relatively large number of smPCRs to obtain a reasonable Poisson distribution of PCR-wells each having 0 or 1 molecule (or more). A 6-carboxyfluorescein (FAM)-labeled forward (P1) primer was used for smPCR to provide FAM-labeled amplicons for capillary electrophoresis to determine interpolated sizes relative to added rhodamine (ROX)-labeled size-standards. Results confirmed that FAM-labeled amplicons had ~150-200 bp-sizes as expected for the 5mC-protected SOLiD fragment-library excised following PAGE, and that the number of such FAM-labeled amplicons detected in any given PCR-well decreased with lower concentrations of diluted stock solutions. Such range-finding results generally led to reasonable, Poisson-like single-molecule distributions (see below) that were with ~two-fold dilution of the ~1 molecule/μL concentrations calculated as described above. These optimized stock solutions were then used to prepare a total of ~1,500 5-μL smPCRs in 96-well microtiter plates in batches of 4 plates. Manually processing batches of 4 plates was easily performed on a daily basis and, moreover, was found to mitigate spurious non-template-dependent amplification or primer-dimer problems that occasionally necessitated discarding data plate-wise and repeating smPCRs of such plates.

[00121] In some cases, smPCR of a library-fragment gave rise to a group of FAM-labeled peaks, each separated by 1-bp and symmetrically distributed about a major peak that was within the expected range of ~150-200 bp. This phenomenon was attributed to polymerase slippage at oligo(T) or oligo(A) [or dinucleotide-repeats] regions of DNA during PCR, by analogy to the mechanism originally proposed to explain the observation of "shadow" bands in PCR of DNA having regions of oligo(CA). As has previously been discussed, Sanger-sequencing evidence for slippage at oligo(T) regions having >9 Ts in bisulfite-converted DNA in the context of avoiding such regions when designing PCR primers for amplification and Sanger sequencing. In the presently described SOLiD fragment-library, regions of oligo(T) or oligo(A) with >9 Ts or As within the fragment sequence are, unfortunately, unavoidable due to the random nature of fragment generation and use of universal, fixed-sequence primers for smPCR amplification of all library-fragments. smPCR-wells judged by visual inspection to contain either a single, appropriately sized (FAM-P1/P2)-derived library-

IDT Exhibit 1020.1085

fragment in the range of ~150-200 bp, and those smPCR-wells showing slippage that was not too extensive, were all subjected to Sanger sequencing as described in the next section.

QC of single-molecule library-fragment amplicons by capillary electrophoretic Sanger sequencing

[00122] Sanger-based sequence analysis of amplicons derived from smPCR of individual library-fragments after confirmatory sizing (see above) established the extent of C→T conversion achieved within each of such library-fragments that is randomly sampled. Sampling a relatively large number of bisulfite-converted library-fragments for this QC analysis thus provides a clear indication of % C→T achieved as a checkpoint for deciding whether or not to proceed with massively parallel, redundant ("deep") sequencing by means of SOLiD for genome-wide methylome analysis. The extent of genomic coverage achievable by this type of Sanger-sequencing QC analysis of a human genome-wide fragment-library derived from ~3 X 10^9 bp gDNA will represent an extremely small percentage of the genome even if many 1000s of library-fragments are randomly sampled by smPCR. On the other hand, even lesser numbers of Sanger-sequenced smPCR amplicons, such as ~200 discussed below, can provide compelling information on % C→T conversion in view of the following approximations. The ~150-200 bp range of fragments in the library implies an average of ~175 bases in a single-stranded fragment that has an average C-content of (~175 bases) X 25% = ~44 Cs, excluding for the sake of simplicity 5mCpG dinucleotides and various possible sources of bias. Thus, ~200 Sanger sequences that each covering an entire fragment provide ~44 Cs X ~200 = ~8,800 Cs that can each be detected as either a C (non-converted) or T (converted). This digital detection and counting therefore represents a dynamic range of nearly 10^4. In addition, exact sequence-contexts for any non-converted Cs that might be detected could possibly reveal particular sequences wherein Cs resist conversion, especially double-stranded hairpin regions akin to those described in studies of hairpin-bisulfite PCR.

[00123] In view of the aforementioned considerations, the Yoruban fragment-library that had been reacted with bisulfite as free-solution DNA or PAGE-slice-embedded DNA (Bis-PAGE) for 8-hr or overnight was serially diluted for smPCR,

IDT Exhibit 1020.1086

as discussed above, to provide amplicons for conventional capillary electrophoretic Sanger sequencing. In these initial experiments aimed at comparing the stated reaction conditions, aliquots of optimally diluted sample solutions provided ~20 smPCRs per 96-well PCR plate. This average smPCR success rate of ~20% compares favorably with calculated Poisson-distribution percentages of 36% for an average of 1 molecule/well, and 16% for an average of 0.2 molecule/well (or 1 molecule/5 wells). The presently reported design of a SOLiD fragment-library provides for a single orientation after bisulfite conversion such that the forward primer (P1) led to sequencing the strand depleted of C, and the reverse primer (P2) led to sequencing the complementary strand depleted in G. For all four of the reaction conditions specified above, randomly sampled library-fragments leading to smPRC amplicons and corresponding Sanger-sequencing electropherograms were found to be completely converted, i.e. there were no Cs detected other than those present as CpG dinucleotides and thus indicative of 5mCpG dinucleotides in the starting gDNA sample. Careful visual perusal of all of the Sanger-sequencing electropherograms for this preliminary assessment of four different conditions for reaction library-fragments with bisulfite failed to reveal noticeable differences, despite the aforementioned higher "Ct"-like values for samples incubated overnight. Higher "Ct"-like values have been attributed to loss of DNA by acidic and/or other bisulfite-related degradation mechanisms, which have been discussed in detail elsewhere. Alternatively, or in addition, loss of DNA may occur by diffusion of DNA from the PAGE-slice in the case of Bis-PAGE. Degradation mechanisms may have sequence-dependent aspects, and thus represent a possible source of bias that should be minimized in genome-wide bisulfite-sequencing using SOLiD by limiting the C→T conversion processes for fragment-libraries described herein to an 8-hr incubation time. Reducing this and other sources of loss is especially important when starting out with relatively small amounts of gDNA in order to minimize under-representation of sequences in the bisulfite-converted fragment-library that is ultimately subjected to methylome analysis by SOLiD.

[00124] To further assess the completeness of bisulfite conversion of the 8-hr Bis-PAGE sample discussed above, ten additional 96-well microtiter plates (960 wells total) containing the optimally diluted Yoruban fragment-library were subjected

IDT Exhibit 1020.1087

to smPCR. Instead of applying size-based capillary electrophoretic analysis to select only wells that each contain a single-sequence amplicon, as discussed above, Sanger sequencing reactions were carried out in all 96-wells of each plate (960 wells total) for subsequent capillary electrophoresis. Visual inspection of peak-spacing and peak-color in all of the resultant electropherograms led to identification of ~200 wells that each contained a single-sequence amplicon. Careful perusal of all of the resultant fragment-sequences revealed the following results. There were two of library-fragments giving rise to Sanger sequences having much longer length, i.e. 190 and 147, compared to other library-fragments, which indicated heterogeneity of shearing and PAGE-sizing during preparation of the library. Furthermore, C was present in all of the ~200 Sanger-sequenced library-fragments almost exclusively in CpG dinucleotides that reflect 5mCpG dinucleotides that were present in the original sample of human, Yoruban gDNA. There were only five other instances of C found to be present at non-CpG sites. Three of these five instances were GpC dinucleotides, which may tentatively be attributed to naturally occurring Gp(5mC) dinucleotides in the original sample of human gDNA.

[00125] Common adapter-ends reported herein for ligation to relatively short fragments of gDNA lead to double-stranded SOLiD library-fragments all having the same complementary flanking-sequences. The common complementary flanking sequences represent a significant proportion (up to ~50%) of the total molecular composition of each library-fragment. In principle, this circumstance could "drive" re-annealing and thus lead to inefficient bisulfite conversion, which is known to require single-stranded regions. This concern proved to be a non-issue by finding >99% conversion of C→T by Bis-PAGE using formamide, based on "gold standard" Sanger sequencing of a relatively large number (~200) of randomly sampled library-fragments. In addition to the present use of nick-translation directly in a PAGE-slice to streamline construction of this 5mC-protected fragment-library, Bis-PAGE was shown to be a novel means of simplifying sample handling, and reducing the multiplicity of steps, compared to conventional bisulfite conversion of DNA in free-solution. Bis-PAGE provides a way to bypass potential loss of relatively short (~150-200 base) library-fragments that could likely occur using conventional DNA-binding matrices for

IDT Exhibit 1020.1088

recovery. However, prolonged incubation in Bis-PAGE-slices and/or use of insufficiently (<6%) cross-linked polyacrylamide could lead to inadequate recovery and should therefore be avoided. Comparison of Bis-PAGE using formamide for both pre-denaturing and denaturing after addition of bisulfite in place of conventional pre-denaturing with NaOH indicated slightly higher recovery of PCR-amplifiable bisulfite-converted library-fragments with formamide, although the reasons for this are uncertain at the present time. More importantly, limited results of preliminary experiments indicated that human gDNA, without conventional restriction enzyme-mediated cutting to reduce size, could be simply infused into 6% PAGE-slices for successful Bis-PAGE. This offers the possibility of a more convenient bisulfite-conversion protocol applicable to many types of DNA methylation analyses that are available.

Example 4

[00126] Figures 13-16 depict an exemplary method according to the present teachings wherein each of the strands of circularized DNA comprised a nick. The use of a nick on both strands may allow either of the strands to be converted by a bisulfite reaction.

[00127] In Figure 13, cap adapters 1010 were ligated to a DNA fragment 1001. The cap adapters 1010 were missing a 5′ phosphate from one of the oligonucleotides. The missing 5′ phosphate allowed for the formation of nicks N when the DNA fragment 1001 was circularized. A biotinylated internal adapter 1020 was ligated to the cap adapters 1010 to form the circularized polynucleotide.

[00128] The circularized polynucleotide was nick translated with 5mC dNTP, as shown in Figure 14. The nick translated polynucleotide was then exposed to T7 exonuclease and S1 nuclease to form long mate-pair tags 1002 and 1003. Due to the use of 5mC dNTP in the nick translation, mate-pair tag 1003 was 5mC bisulfite protected and mate-pair tag 1002 retained its native bisulfite sensitivity.

[00129] In the first step of Figure 15, P1 and P2 adapters were ligated to the ends of the DNA. The ligated DNA was then nick translated with DNA polymerase to fill in the non-ligated and non-methyl-C-protected adapter strand.

IDT Exhibit 1020.1089

[00130] Before bisulfite conversion was carried out, the strands were isolated by capturing the biotinylated strand with streptavidin polystyrene beads 1030. See Figure 16. The DNA was denatured and the non-captured strand 1050 was separated and eluted off of the captured strand 1040. Once separated, either one or both of strands 1040 and 1050 were ready for bisulfite conversion and subsequent analysis.

Example 5

[00131] The DNA of Example 5 used 90 μg of MCF-7, DNA from a human cancer cell line.

Shearing the DNA

[00132] The genomic DNA was sheared to yield 600 bp to 6 kb fragments. To shear for a mate-paired library with insert sizes between 600 bp and 1 kb, the Covaris™ S2 system was used. To shear for a mate-paired library with insert sizes between 1 kb and 6 kb, the HydroShear was used. HydroShear used hydrodynamic shearing forces to fragment DNA strands, wherein the DNA in solution flowed through a tube with an abrupt contraction. As it approached the contraction, the fluid accelerated to maintain the volumetric flow rate through the smaller area of the contraction. During this acceleration, drag forces stretched the DNA until it snapped and until the pieces were too short for the shearing forces to break the chemical bonds. The flow rate of the fluid and the size of the contraction determined the final DNA fragment sizes. A calibration run to assess the shearing efficacy of the device prior to starting the first library preparation was performed.

Purification of the DNA with Qiagen QIAquick® Gel Extraction Kit

[00133] Sample purification was performed with Qiagen QIAquick® columns supplied in the QIAquick® Gel Extraction Kit. Qiagen QIAquick® columns have a 10-μg capacity, so multiple columns were used during a purification step. For larger amounts of DNA for library construction, phenol-chloroform-isoamyl alcohol extraction and isopropyl alcohol precipitation can be used.

IDT Exhibit 1020.1090

End-repairing the DNA

[00134] The Epicentre® End-It™ DNA End-Repair Kit was used to convert DNA with damaged or incompatible 5′-protruding and/or 3′-protruding ends to 5′-phosphorylated, blunt-ended DNA for fast and efficient blunt-ended ligation. The conversion to blunt-end DNA was accomplished by exploiting the 5′►3′ polymerase and the 3′►5′ exonuclease activities of T4 DNA Polymerase. T4 polynucleotide kinase and ATP were also included for phosphorylation of the 5′-ends of the blunt-ended DNA for subsequent ligation.

Ligating dsMethyCAP Adapters to the DNA

[00135] The ligation of the dsmethyCAP adapter added the methyCAP adapters to both ends of the sheared, end-repaired DNA. The methyCAP adapter was missing a 5′ phosphate from one of its oligonucleotides, which resulted in a nick on each strand when the DNA is circularized in a later step. The dsmethyCAP adapters were included as a 50 uM solution in double-stranded form in the SOLiD™ Mate-Paired Library Bisulfite-Methylation Kit.

Size-selecting the DNA

[00136] Depending on the desired insert-size range, the ligated, purified DNA was run on a 0.8% or 1% agarose gel. The correctly sized ligation products were excised and purified using the Qiagen QIAquick® Gel Extraction Kit.

Circularization of the DNA

[00137] Sheared DNA ligated to methyCAP Adapters was circularized with a biotinylated internal adapter. To increase the chances that ligation occurred between two ends of one DNA molecule versus two different DNA molecules, a very dilute reaction was used. The circularization reaction products were purified using the QIAquick® Gel Extraction Kit. The biotinylated Internal Adapter dsMethyIA was included as a 2.0 uM solution, double-stranded form in the SOLiD™ Mate-Paired Library Bisulfite Methylation Kit.

IDT Exhibit 1020.1091

Treating the DNA with Plasmid-Safe™ ATP-Dependent DNase

[00138] Epicentre® Plasmid-Safe ATP-Dependent DNase was used to eliminate uncircularized DNA. After the Plasmid-Safe™ DNase-treated DNA was purified using the QIAquick® Gel Extraction Kit, the amount of circularized product was quantified. A minimum of 200 ng of circularized product was needed to proceed with library construction. For more complex genomes, 600 ng to 1 µg circularized DNA is needed for a high-complexity library.

Nick-translating the circularized DNA with 5mC dNTP-containing dNTPs

[00139] Nick translation using *E. coli* DNA polymerase I translated the nick into the genomic DNA region. The size of the mate-paired tags produced was controlled by adjusting the reaction temperature and time. The nick translated portion using 5mC was resistant to bisulfite conversion. Therefore, one end of each strand originating from dsDNA genome had a mate-paired portion that bisulfite converted (except for native 5mC bases) and the other Mate-Pair Tag reference matched to the non-bisulfite genome.

Digesting the DNA with T7 exonuclease and S1 nuclease

[00140] T7 exonuclease recognized the nicks within the circularize DNA and with its 5′►3′ exonuclease activity chewed the unligated strand away from the tags creating a gap in the sequence. This gap created an unexposed single-stranded region that was more easily recognized by S1 nuclease and the library molecule was cleaved from the circularized template.

Capturing on 6.7 Micron Polystyrene Streptavidin Beads Following End-repair

[00141] Regular dNTPs were used for end repair (not 5mC-dNTP) in order to avoid introduction of an inappropriate 5mC in the native strand that would appear to be incomplete bisulfite conversion. The genomic "reference" TAG that was 5mC protected may have occasionally lacked 5mC "protection" because of end-repair, so that a C->T SNP was created. Non-magnetic beads were used to avoid oxidation of the DNA by Fe^{++} during the bisulfite conversion. Capture of the library on polystyrene

IDT Exhibit 1020.1092

beads in place of magnetic beads required pelleting the polystyrene by high speed centrifugation in place of using a magnetic stand. By pelleting in the presence of a small percentage of detergent containing buffer (TEX), the beads packed well and the solution above the beads was efficiently removed without disturbing the bead bed. It was safe to leave traces of supernatant on the beads and carry over small amounts from the previous (wash) steps.

Ligating MethyP1 and MethyP2 Adapters to the DNA

[00142] P1 and P2 adapters were ligated to the ends of the end-repaired DNA. The methyP1 and methyP2 adapters were included in double-stranded form as a 50 uM solution in the SOLiD™ Mate-Paired Library Bisulfite Methylation Kit.

Nick-translating the library with 5mC dNTP-containing dNTPs

[00143] The ligated, purified DNA underwent nick translation with DNA polymerase. The non-ligated and non-methyl-C-protected adapter strand of the adapter pairs was filled in with 5mC dNTP, fully protecting the adapter sequences during the bisulfite conversion.

Bisulfite Conversion

[00144] The polystyrene beads having double stranded library were attached. Bisulfite conversion required single stranded DNA for efficient bisulfite conversion. The beads were treated with 50 uL of 0.1M NaOH just prior to introduction of bisulfite reagent. The NaOH solution was removed, along with the eluted off single stranded library.

[00145] OPTION ONE: It is possible to add the conversion reagent (bisulfite solution) to the beads, incubate at 50°C for 8 hours. Wash steps and desulfonation may be performed on the library still attached to the polystyrene beads. The beads may then used directly in PCR for library amplification. OPTION TWO: The NaOH solution may also be bisulfite treated and purified with Microcon 100 or PureLink micro PCR kit with a desulfonation buffer for the desulfonation step. Recover bisulfite converted library from column with LoTE.

IDT Exhibit 1020.1093

Amplification of the library

[00146] The library was amplified using Library PCR Primers 1 and 2 with SOLiD™ Library PCR Master Mix (Platinum Super Mix) supplemented with additional AmpliTaq Gold DNA Polymerase to improve yields in amplification of uracil (from the deaminated cytosine from the bisulfite conversion). In order to achieve whole genome representation during SOLiD sequencing and obtain quantitative accuracy of a human methylome, library amplification did not exceed 17 cycles. Additional cycles may cause PCR-related biases due to differential amplification of library molecules.

Gel-purified the library

[00147] The library was run on a 3% agarose gel and the library band (~300 bp) was excised and eluted using the Qiagen QIAquick® Gel Extraction Kit. The library was then quantified.

[00148] While the present teachings have been described in terms of these exemplary embodiments, the skilled artisan will readily understand that numerous variations and modifications of these exemplary embodiments are possible without undue experimentation. All such variations and modifications are within the scope of the current teachings.

[00149] Although the disclosed teachings have been described with reference to various applications, methods, kits, and compositions, it will be appreciated that various changes and modifications can be made without departing from the teachings herein and the claimed invention below. The foregoing examples are provided to better illustrate the disclosed teachings and are not intended to limit the scope of the teachings presented herein.

[00150] In this application, the use of the singular can include the plural unless specifically stated otherwise or unless, as will be understood by one of skill in the art in light of the present disclosure, the singular is the only functional embodiment. Thus, for example, “a” can mean more than one, and “one embodiment” can mean that

IDT Exhibit 1020.1094

the description applies to multiple embodiments. Additionally, in this application, "and/or" denotes that both the inclusive meaning of "and" and, alternatively, the exclusive meaning of "or" applies to the list. Thus, the listing should be read to include all possible combinations of the items of the list and to also include each item, exclusively, from the other items. The addition of this term is not meant to denote any particular meaning to the use of the terms "and" or "or" alone. The meaning of such terms will be evident to one of skill in the art upon reading the particular disclosure.

Example 6

[00151] The DNA of Example 6 used 90 μg of MCF-7, DNA from a human cancer cell line.

Sheared the DNA

Prepared for shearing

1. The shearing method used was based on the desired insert size of the mate-paired library (see Table 1).

Table 1. Shearing conditions for desired mate-paired library insert sizes.

Insert Size	Shearing Method	Shearing Conditions
600 to 800 bp	Covaris™ Shearing in 20% glycerol (13 mm × 65 mm borosilicate tube)	• Number of Cycles: 75 • Bath Temperature: 5 °C • Bath Temperature Limit: 12 °C • Mode: Frequency sweeping • Water Quality Testing Function: Off • Duty cycle: 2% • Intensity: 7 • Cycles/burst: 200 • Time: 10 sec
800 to 1000 bp	Covaris™ Shearing in 20%	• Number of Cycles: 30

IDT Exhibit 1020.1095

	glycerol (13 mm × 65 mm borosilicate tube)	• Bath Temperature: 5 °C • Bath Temperature Limit: 12 °C • Mode: Frequency sweeping • Water Quality Testing Function: Off • Duty cycle: 2% • Intensity: 5 • Cycles/burst: 200 • Time: 10 sec
1 to 2 kb	HydroShear® Standard Shearing Assembly	• SC5 • 20 cycles
2 to 3 kb	HydroShear® Standard Shearing Assembly	• SC9 • 20 cycles
3 to 4 kb	HydroShear® Standard Shearing Assembly	• SC13 • 20 cycles
4 to 5 kb	HydroShear® Standard Shearing Assembly	• SC15 • 5 cycles
5 to 6 kb	HydroShear® Standard Shearing Assembly	• SC16 • 25 cycles

2. The shearing conditions were tested to ensure that the shearing conditions resulted in the desired insert sizes. Sheared 5 µg DNA and ran 150 ng sheared DNA on a 0.8% E-gel according to the manufacturer's specifications.

Sheared the DNA using the Covaris™ S2 System

1. In a round bottom 13 mm × 65 mm borosilicate tube, diluted 5 to 20 µg DNA in 500 µL so that the final volume contained 20% glycerol in nuclease-free water.

Component	Amount
99% Glycerol	100 µL
DNA	5 to 20 µg

IDT Exhibit 1020.1096

Nuclease-free water	Variable
Total	500 µL

2. Sheared the DNA using the Covaris™ S2 System shearing program described above.
3. Transfered 500 µL sheared DNA into a clean 1.5-mL LoBind tube.
4. Washed the borosilicate tube with 100 µL nuclease-free water and transferred the wash to the 1.5-mL LoBind tube. Mixed by vortexing and then proceeded to purify the DNA with Qiagen QIAquick® Gel Extraction Kit.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the sheared DNA. If the color of the mixture was orange or violet, added 10 µL 3M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL sheared DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.

IDT Exhibit 1020.1097

11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.
13. Quantitated the purified DNA by using 2 µL of the sample on the NanoDrop™ ND-1000 Spectrophotometer (see Appendix B).

End-repaired the sheared DNA

Repairing the Sheared DNA Ends with Epicentre® End-It™ DNA End-Repair Kit

1. Combined and mixed the following components in a LoBind tube.

Component	Amount
Sheared DNA *X* µg = 15	120 µL
End-Repair 10X Buffer (Epicentre® End-It™)	20 µL
ATP (10 mM) (Epicentre® End-It™)	20 µL
dNTPs (2.5 mM each) (Epicentre® End-It™)	20 µL
End-Repair Enzyme Mix (Epicentre® End-It™)	6.7 µL
Nuclease-free water (Variable)	13.3 µL
Total	200 µL

2. Incubated the mixture at room temperature for 30 minutes.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the end-repaired DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.

IDT Exhibit 1020.1098

2. Applied 750 µL end-repaired DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.
13. Quantitated the purified DNA by using 2 µL of the sample on the NanoDrop™ ND-1000 Spectrophotometer (see Appendix C).
14. For structural variation studies where tighter size selection of fragments was required, performed one of two size selections (see "Size-select the DNA") at this point and then proceeded to "Ligate LMP CAP Adapters to the DNA." If tight insert size distribution were not as critical, proceeded directly to "Ligate LMP CAP Adapters to the DNA." This optional size-selection was not used if the starting DNA input was less than 10 µg.

IDT Exhibit 1020.1099

Ligated dsMethyCAP Adapters to the DNA

CapBnoPhos	ACAGCAG	(SEQUENCE ID NO: 13)
EcoP151 Cap-A (5mC)	5' PHOS-CTGCTGTAC	(SEQUENCE ID NO: 14)

Ligate thed adapters to the DNA

1. Calculated the amount of adapter needed for the reaction based on the amount of DNA from the last purification step.

For 12 µg of purified end-repaired DNA with an average insert size of 1.5 kb

$$\textbf{X}\ \text{pmol/µg DNA} = 1\ \text{µg DNA} \times \frac{10^6\ \text{pg}}{1\ \text{µg}} \times \frac{1\ \text{pmol}}{660\ \text{pg}} \times \frac{1}{1500} = 1.0\ \text{pmol/µg DNA}$$

$$\textbf{Y}\ \text{µL adaptor needed} = 12\ \text{µg DNA} \times \frac{1.0\ \text{pmol}}{1\ \text{µg DNA}} \times 100 \times \frac{1\ \text{µL adaptor needed}}{50\ \text{pmol}}$$

$$= 24\ \text{µL adaptor needed}$$

2. Combined and mixed the components below. If a larger reaction volume was required to incorporate all of the DNA, scaled up the Quick Ligase and Quick Ligase Buffer. Added 1 µL Quick Ligase per 40 µL of reaction volume. Added 1 µL 2× Quick Ligase Buffer per 2 µL of reaction volume.
*From NEB

Component	Volume (µL)
dsMethy-CAP Adapter (ds) (50 pmol/µL)	22.5 (varied slightly)
2× Quick Ligase Buffer*	150
Quick Ligase Enzyme*	7.5
DNA 15 ug	120
Nuclease-free water	NONE

IDT Exhibit 1020.1100

Total	302

3. Incubated the reaction mixture at room temperature for 10 minutes.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the ligated DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL ligated DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. If necessary, pooled the eluted DNA.

IDT Exhibit 1020.1101

Size-select the DNA

Size-selected the DNA fragments with an agarose gel

1. Determined the appropriate percentage of agarose gel needed to size-select DNA.

Desired Insert Size	Agarose gel needed (%)
600 to 3000 bp	1.0
3 to 6 kb	0.8

2. Prepared the appropriate percentage agarose gel in 1× TAE buffer with 10 µL of 10 mg/mL ethidium bromide per 100 to150 mL gel volume. To prepare the 1% gels, used either Agarose-LE (Applied Biosystems, AM9040) or 1% Mini ReadyAgarose Gel (Bio-Rad, 161-3016).
3. Added 10× Gel Loading Solution to the purified ligated DNA (1 µL 10× Gel Loading Solution for every 10 µL DNA).
4. Loaded 1 µL 1 kb DNA ladder. Loaded up to 20µL dye-mixed sample per well. At least one lane in between the ladder well and the sample wells was used to avoid contamination of the sample with ladder.
5. Ran the gel at 120 V until the marker was close to the edge of the gel.
6. Destained the gel in nuclease-free water twice for 2 minutes each time and visualized the gel on a UV transilluminator with a ruler lying on top.
7. Using the ladder bands and the ruler for reference, excised the band of the gel corresponding to the insert size range of interest with a clean razor blade. If desired, a tighter size selection could be carried out at this stage by taking a tighter cut. If the gel piece was large, it was sliced it up.

Eluted the DNA Using Qiagen QIAquick® Gel Extraction Kit

1. Weighed the gel slice(s) in a 15-mL polypropylene conical colorless tube.
2. Added 3 volumes Buffer QG to 1 volume of gel.
3. Dissolved the gel slice by vortexing at room temperature until the gel slice was dissolved completely (~5 minutes).
4. If the color of the mixture was yellow, proceeded to step 5. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.

5. Added one gel volume of isopropyl alcohol to the sample and mixed by inverting the tube several times.
6. Applied about 700 µL sample to the column(s). The maximum amount of gel that could be applied to a QIAquick® column was 400 mg. Used more columns as necessary.
7. Let the column(s) stand for 2 minutes at room temperature.
8. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
9. Repeated steps 6 and 8 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
10. Added 750 µL Buffer PE to wash the column(s).
11. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
12. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
13. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
14. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
15. Repeated steps 13 and 14.
16. If necessary, pooled the eluted DNA in a 1.5-mL LoBind tube.
17. Quantitated the purified DNA by using 2 µL of the sample on the NanoDrop™ ND-1000 Spectrophotometer.

Circularize the DNA with dsMethylInternal adapter

dsMethyIA 5' (PHOS) CGTACA(BIO-dT)CCGCCTTGGCCGT
3' TGGCATGT A GGCGGAACCGG-PHOS5'
(SEQUENCE ID NO: 15)

IDT Exhibit 1020.1103

Circularized the DNA

1. Prepared a circularization reaction by mixing the components listed below (in order) based on the desired insert size where *X* was the number of micrograms of DNA to be circularized (see table). If a larger reaction volume was required, scaled up the Quick Ligase and Quick Ligase Buffer. Added 1 µL Quick Ligase per 20 µL of reaction volume.

Components	Amount						
	600 to 800 bp	800 to 1000 bp	1 to 2 kb	2 to 3 kb	3 to 4 kb	4 to 5 kb	5 to 6 kb
Nuclease-free water	Variable	Variable	Variable	Variable	Variable	Variable	Variable
DNA	*X* µg	*X* µg	*X* µg	*X* µg	*X* µg	*X* µg	*X* µg
2× Quick Ligase Buffer	(*X* × 117.5) µL	(*X* × 135) µL	(*X* × 182.5) µL	(*X* × 250) µL	(*X* × 280) µL	(*X* × 312.5) µL	(*X* × 360) µL
Internal Adapter (ds) (2 µM)	(*X* × 3.75) µL	(*X* × 2.84) µL	(*X* × 1.5) µL	(*X* × 0.9) µL	(*X* × 0.65) µL	(*X* × 0.5) µL	(*X* × 0.4) µL
Quick Ligase (Use double)	(*X* × 6) µL	(*X* × 6.75) µL	(*X* × 9) µL	(*X* × 12.5) µL	(*X* × 14) µL	(*X* × 15.6) µL	(*X* × 18) µL
Total	(*X* × 235) µL	(*X* × 270) µL	(*X* × 365) µL	(*X* × 500) µL	(*X* × 560) µL	(*X* × 625) µL	(*X* × 720) µL

For DNA in 2 to 3 kb range circularized

Components	Amount
Nuclease-free water	552.3 µL

IDT Exhibit 1020.1104

Variable	
DNA 3.0 µg	120 µL
2× Quick Ligase Buffer	750 µL
dsMethyIA Internal Adapter (ds) (2 µM)	2.7 µL (varied slightly with the measured amount of DNA is the 7 samples.)
Quick Ligase (2X)	75 µL
Total	1500 µL

2. Incubated at room temperature for 10 minutes.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the circularized DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL circularized DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.

IDT Exhibit 1020.1105

8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.

Isolate the circularized DNA

Treated the DNA with Plasmid-Safe™ ATP-Dependent DNase

1. Combined and mixed the components below.
   For 3.46µg X 6 of DNA used in the circularization reaction.

Components	Volume (µL)
ATP (25 mM)	5
10× Plasmid-Safe™ Buffer	10
Plasmid-Safe™ DNase (10 U/µL)	1.15 µL
DNA (3.46 µg)	60 µL
Nuclease-free water	24 µL
Total	100 µL

2. Incubated the reaction mixture at 37 ºC for 40 minutes.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the Plasmid-Safe™ DNase-treated DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL Plasmid-Safe™ DNase-treated DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.

IDT Exhibit 1020.1106

3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.
13. Quantitated the purified DNA by using 2 µL of the sample on the NanoDrop™ ND-1000 Spectrophotometer (see Appendix C).

Nick-translate the circularized DNA with 5mC contining dNTPs (25 mM each)

Nick-translated the circularized DNA

1. This step created the 5mC bisulfite protected tags. Combined and mixed the components listed below on ice. First, mixed all of the components except the enzyme and chilled on ice. Added the enzyme, quickly vortexed and immediately proceeded to the next step.

For 1 µg of circularized DNA

Components	Amount
dNTP Mix (100 mM, 25 mM each)	5 µL
10× NEBuffer 2	50 µL
DNA Polymerase I (10 U/µL)	10 µL

IDT Exhibit 1020.1107

DNA 1000 ng	60 µL
Nuclease-free water	375
VARIABLE	
Total	500 µL

2. Incubated the reaction at 0 ºC in an ice-water bath for 12 to 14 minutes.
3. Stopped the reaction immediately by proceeding to "Purify the DNA with Qiagen QIAquick® Gel Extraction Kit."

Purify the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the nick-translated DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL nick-translated DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.

12. If necessary, pooled the eluted DNA.

Digest the DNA with T7 exonuclease and S1 nuclease

Digested the DNA with T7 exonuclease

1. Combined:

For 1.26 µg of circularized DNA in each of the 4 samples:

Component	Amount
DNA 1260 ng Always 60 µL	60 µL
NEBuffer 4, 10×	63.2 µL
T7 exonuclease (10 U/µL)	25.3 µL
Nuclease-free water Variable	483.5 µL
Total	632 µL

2. Incubated the reaction mixture at 37 °C for 30 minutes. Immediately proceeded to the next step.

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the T7 exonuclease digested DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL T7 exonuclease digested DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).

IDT Exhibit 1020.1109

7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.

Digested the DNA with S1 nuclease

1. Freshly diluted Invitrogen S1 nuclease to 1 U/µL with S1 dilution buffer.
2. Combined:

   For T7 exonuclease digested DNA from 1260 ng circularized DNA for each of the 4 tubes in the previous step (The total amount of DNA prior to linearization was 5.056µg divided into the 4 tubes based on the circularized DNA present. The actual µg of DNA was much less after it has been linearized) :

Component	Amount
T7 exonuclease digested DNA 1260 ng	60 µL
S1 nuclease buffer, 10×	63.2 µL
3 M sodium chloride	31.6 µL
100 mM magnesium chloride	63.2 µL
S1 nuclease, diluted to 1 U/µL	25.3µL
Nuclease-free water Variable	388.7 µL
Total	632 µL

3. Incubated the reaction mixture at 37 °C for 30 minutes. Immediately proceeded to the next step.

IDT Exhibit 1020.1110

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the digested DNA. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL digested DNA in Buffer QG to the column(s). The maximum amount of DNA that could be applied to a QIAquick® column was 10 µg. Used more columns as necessary.
3. Let the column(s) stand for 2 minutes at room temperature.
4. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the column(s). Placed the QIAquick® column(s) back into the same collection tube.
6. Added 750 µL Buffer PE to wash the column(s).
7. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
8. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Repeated steps 9 and 10.
12. If necessary, pooled the eluted DNA.

End-repair the digested DNA

The end-repaired DNA was repaired with a regular dNTP mix comprising no 5mCdNTP. During SOLiD sequencing, the 5mC preserved sequence may have had a T where there was an end-repaired C. Because most Cs are not methylated, use of "regular" dNTPs erred on the side of an occasional missed 5mC.

IDT Exhibit 1020.1111

Repaired the digested DNA ends with the Epicentre® End-It™ DNA End-Repair Kit

1. Prepared Streptavidin Binding Buffer:

Components	Volume (µL)
500 mM Tris-HCl (pH 7.5)	10
5 M Sodium chloride	200
500 mM EDTA	1
Nuclease-free water	289
Total	500

2. Combined:

Component	Amount
S1 digested DNA *X* ng	60 µL
End-repair buffer, 10X	10 µL
10 mM ATP	10 µL
Regular dNTPs (2.5 mM each)	10 µL
End-Repair Enzyme Mix *	2 µL
Nuclease-free water Variable	8
Total	100 µL

*From the Epicentre® End-It™ DNA End-Repair Kit

3. Incubated the reaction mixture at room temperature for 30 minutes.
4. Stopped the reaction by combining and mixing the components below:

Components	Volume (µL)
First End-repaired DNA	100
500 mM EDTA	5
Streptavidin Binding Buffer	200
Second End-repaired DNA	100
Total	405

IDT Exhibit 1020.1112

Bind the library molecules to POLYSTYRENE-streptavidin beads

Pre-washed the beads

1. Prepared 1× BSA:

Components	Volume (µL)
100× BSA	5
Nuclease-free water	495
Total	500

2. Vortexed a 5 mL bottle of Spherotech streptavidin beads (6.7 micron beads supplied as a 5% w/vol slurry in water) to thoroughly suspend the polystyrene beads in solution. Transferred 200 µL per library sample (1 mg of beads/200µL) into a 1.5-mL LoBind Tube using a 1 mL pipette tip with a suitable pipettor.
3. Centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute. Discarded the supernatant without disturbing the polystyrene bed.
4. Added 400 µL 1× Bead Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
5. Discarded the supernatant without disturbing the polystyrene bead bed.
6. Added 400 µL 1× BSA and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
7. Discarded the supernatant without disturbing the polystyrene bead bed.
8. Added 400 µL 1× Bind & Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
9. Discarded the supernatant without disturbing the polystyrene bead bed.

Bound the library DNA molecules to the beads

1. Added the entire 405 µL solution of library DNA in Streptavidin Binding Buffer to the pre-washed beads and vortexed.

IDT Exhibit 1020.1113

2. Mixed by rotation at room temperature for 30 minutes. Afterwards, pulse-spun.

Washed the Bead-DNA Complex

1. Prepared 1× Quick Ligase Buffer:

Components	Volume (µL)
Quick Ligase Buffer, 2×	300
Nuclease-free water	300
Total	600

2. Added 100 µL 1X TEX Buffer to the library-bead attachment reaction, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
3. Discarded the supernatant without disturbing the polystyrene bead bed.
4. Added 400 µL 1× Bead Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
5. Discarded the supernatant without disturbing the polystyrene bead bed.
6. Added 400 µL 1× Bind & Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
7. Discarded the supernatant without disturbing the polystyrene bead bed.
8. Added 400 µL 1× Bind & Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
9. Discarded the supernatant without disturbing the polystyrene bead bed.
10. Resuspended the beads in 500 µL 1× Quick Ligase Buffer. Vortexed for 15 seconds and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Discarded the supernatant without disturbing the polystyrene bead bed. .
12. Resuspended the beads in 97.5 µL 1× Quick Ligase Buffer

IDT Exhibit 1020.1114

Ligate 5mCP1A/B and 5mC-P2A/B Adapters to DNA

5mC-P1-A CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT
(SEQUENCE ID NO: 16)

5mC-P2-A CTGCCCCGGGTTCCTCATTCTCT (SEQUENCE ID NO: 17)

The top strand adapters P1-A and P2-A were synthesized with 5mC. The Nick translation step filled in bottom strand (P1-B and P2-B) with 5mC so that both the top and bottom strands of the adapters were fully 5mC protected (from bisulfite).

Used the bisulfite-SOLiD dsAdapters: dsMethyP1 adapter = 5mCP1A/"regular"B *and dsMethyP2 adapter* = 5mC-P2A/"regular"B

Ligated the P1 and P2 Adapters to the end-repaired DNA

1. Calculated the amount of P1 and P2 Adapters needed for the ligation reaction based on the amount of circularized DNA from "Treat the DNA with Plasmid-Safe™ ATP-Dependent DNase".

For 1 µg of purified circularized DNA with an average size of 1536 (1500 bp insert + 36 bp internal adaptor)

$$\textbf{X} \text{ pmol/µg DNA} = 1 \text{ µg DNA} \times \frac{10^6 \text{ pg}}{1 \text{ µg}} \times \frac{1 \text{ pmol}}{660 \text{ pg}} \times \frac{1}{1536} = 1 \text{ pmol/µg DNA}$$

$$\textbf{Y} \text{ µL adaptor needed} = 1 \text{ µg DNA} \times \frac{1 \text{ pmol}}{1 \text{ µg DNA}} \times 30 \times \frac{1 \text{ µL adaptor needed}}{50 \text{ pmol}}$$

$$= 0.6 \text{ µL adaptor needed}$$

2. Combined:

Components	Volume (µL)
DNA-Bead Complex	97.5
P1 Adapter (ds) (50 µM)	*0.916*
P2 Adapter (ds) (50 µM)	*0.916*
Quick Ligase	2.5
Total	Variable

IDT Exhibit 1020.1115

3. Incubated the reaction mixture at room temperature for 15 minutes.

Wash the DNA-bound streptavidin beads

Washed the bead-DNA complex

Prepared 1× NEBuffer 2 (see table):

Components	Volume (µL)
NEBuffer 2, 10×	60
Nuclease-free water	540
Total	600

1. Centrifuged the adapter ligation reaction at ≥10,000 × *g* (13,000 rpm) for 1 minute.
2. Discarded the supernatant without disturbing the polystyrene bead bed.
3. Resuspended the beads in 400 µL 1× Bead Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
4. Discarded the supernatant without disturbing the polystyrene bead bed.
5. Resuspended the beads in 400 µL 1× Bind & Wash Buffer. Vortexed for 15 seconds and pulse-spun. Added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
6. Discarded the supernatant without disturbing the polystyrene bead bed
7. Resuspended the beads in 400 µL 1× Bind & Wash Buffer and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
8. Discarded the supernatant without disturbing the polystyrene bead bed
9. Resuspended the beads in 500 µL 1× NEBuffer 2. Vortexed for 15 seconds and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
10. Discarded the supernatant without disturbing the polystyrene bead bed
11. Resuspended the beads in 96 µL 1× NEBuffer 2.

IDT Exhibit 1020.1116

Nick-translate the DNA with 5mC-containing dNTPs

Nick-translated the DNA

This step filled-in the 5mC-protected bottom strand adapter sequence.

1. Combined:

Components	Volume (µL)
DNA-Bead Complex	96
5mC-dNTP Mix (100 mM, 25 mM each)	2
DNA Polymerase I (10 U/µL)	2
Total	100

2. Incubated the reaction mixture at 16 °C for 30 minutes.
3. Centrifuged the nick-translation reaction at ≥10,000 × *g* (13,000 rpm) for 1 minute.
4. Discarded the supernatant without disturbing the polystyrene bead bed.
5. Resuspended the beads in 400 µL Buffer EB (Qiagen). Vortexed for 15 seconds and pulse-spun. Add 100 µL 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
6. Discarded the supernatant without disturbing the polystyrene bead bed.
7. Suspended the beads in 500 µL of Lo-TE.
8. Optional saved 50 µL of the 500 uL nick-translated library DNA, in case a library QC needed to be run for troubleshooting purposes.

Bisulfite conversion

One strand of the double stranded library was eluted off the polystyrene beads with dilute NaOH. The biotinylated strand of the library was left attached to the beads. Either or both of these single stranded libraries could be bisulfite converted.

IDT Exhibit 1020.1117

1. Freshly prepared the bisulfite conversion reagent:

Components	Volume (μL)
Zymo CT conversion reagent (1 tube)	
Nuclease free water	750
M Dilution Buffer	210
Total	~1000

Vortexed intermittently over 10 minutes to completely dissolve the sodium bisulfite.

Prepared 0.1 M NaOH (user supplied)

Co-processing of Bisulfite-in-Solution and Bisulfite-on-Bead

1. Centrifuged the nick translated DNA on polystyrene beads (in 500 μL of Lo-TE) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
2. Removed as much of the Lo-TE as possible, minimizing disruption of the polystyrene bead bed.
3. Added 50 μL of 0.1 M NaOH, vortexed for 15 sec and pulse-spun. Incubated for 10 minutes at room temperature to elute the non-biotinylated ssDNA library into the NaOH solution.
4. Centrifuged the beads at ≥10,000 × *g* (13,000 rpm) for 1 minute.
5. Carefully transferred the NaOH solution (supernatant) into a MicroAmp tube.
6. Added 100 μL of the freshly prepared CT bisulfite reagent to the NaOH supernatant, mixed by pipeting up and down a couple of times, capped and incubated at 50°C for 8 hrs in a thermalcycler.
7. Resuspended the beads in 500 μL of Lo-TE to keep as a reserve OR proceeded to step 8 to co-process (bisulfite-convert) the other strand of the library on the polystyrene beads
8. If co-processing the bisulfite-on-beads, added 50 μL of the freshly prepared CT bisulfite reagent to the polystyrene beads. Mixed the bead and bisulfite mixture by pipetting up and down a couple of times and transferred the slurry to a

IDT Exhibit 1020.1118

MicroAmp tube. Added another 50 µL of the freshly prepared CT bisulfite reagent to the original 1.5 mL Lo-Bind Tube(s) to rinse any remaining beads and transferred to the MicroAmp tube (total volume was now ~100 µL). Capped and incubated at 50°C for 8 hrs in a thermalcycler.

Post Bisulfite reaction Processing—Bisulfite-in-Solution

Required an Invitrogen cat# K310050 Purelink PCR Micro kit supplied with a desulfonation solution.

Desalting and Desulfonation

A. Captured Bisulfite converted Library on a PureLink Column

1. Added 600ul Purelink binding buffer (B2) to the PureLink column, and transferred the sample(s) (150 µL bisulfite reaction) into the column containing the binding buffer. Closed the cap and mixed by inverting the column several times.
2. Centrifuged at 10,000 rpm for 1 minute. Discarded the flow-through.
3. Added 600 µL wash buffer to the column, centrifuged at 10,000 rpm for 1 minute or until all the wash buffer was through the filter.

B. Desulfonation

1. Added 200 µL desulfonation buffer to the column and let stand at room temperature (20-30 °C) for 15 minutes. After the incubation, centrifuged 1 minute at 10,000 rpm. Discarded the flow-through.
2. Added 400 µL wash buffer again and centrifuged for 2 minutes at 10,000 rpm to make sure there was no trace amount of wash buffer left on the column. If it was necessary, discarded the flow-through and spun for another 1 minute at 10,000 rpm.
3. Transferred the column to a new elution tube. Added 30 µL of Lo- TE directly to the column matrix. Left at room temperature for 2 minutes and then centrifuged for 1 minute at 10,000 rpm.

IDT Exhibit 1020.1119

Post Bisulfite reaction Processing—Bisulfite-on-Bead

Desalting and Desulfonation

1. Transferred the 100 µL of the bisulfite-on-bead slurry from the microamp tube(s) into a 1.5 mL Lo-bind tube using a total of 600-800 µL of nuclease free water in portions in order to use as rinses during the transfer.
2. Centrifuged the diluted bisulfite reaction at ≥10,000 × *g* (13,000 rpm) for 1 minute.
3. Removed as much of the supernatant without bead loss.
4. Replenished the removed supernatant with nuclease free water (up to ~ 600 µL), vortexed for 15 sec, pulse-spun, added 100 µL TEX buffer, vortexed briefly and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
5. Removed as much of the supernatant without bead loss.
6. Repeated steps 4 and 5 two times.
7. Added 500 µL of 0.1 M NaOH, vortexed for 15 sec, pulse-spun, and allowed to sit at room temperature for 15 minutes. Briefly vortexed and pulse spun a couple of times during the 15 minute wait.
8. Added 100 µL of 1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
9. Discarded the supernatant without disturbing the polystyrene bead bed.
10. Added 500 µL of nuclease free water and vortexed for 15 seconds. Afterwards, pulse-spun, and added 100 µL1X TEX Buffer, briefly vortexed, and centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. Discarded the supernatant without disturbing the polystyrene bead bed.
12. Added 500 µL of Lo-TE and vortexed for 15 seconds. Do Not add TEX. Centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
13. Discarded the supernatant without disturbing the polystyrene bead bed.
14. Resuspended in 30 µL of Lo-TE per sample and proceeded to "Amplify the Library".

IDT Exhibit 1020.1120

Amplify the library

Both the Bisulfite-in-Solution and Bisulfite-on Beads could be processed similiarly (same volume) but the user must have ensured that the beads were suspended in solution before removing the two 2 µL aliquots. The correct number of cycles of PCR needed for optimal amplification.of the bulk of the library was determined during a trial PCR.

1. Prepared a serial dilution of the bisulfite converted library as follows across one row of a PCR plate:

1	2	3	4	5	6	7	8	9	10	11	12
UnDiluted	1/2	1/4	1/8	1/16	1/32	1/64	1/128	1/256	1/512	1/1024	H_2O

The serially diluted bisulfite DNA library volume was 2 µL per well. Well #1 was 2 µL of the undiluted bisulfite DNA library. Introduced 2 µL of H_2O into wells #2-12. Added a second 2 µL aliquot of the bisulfite-DNA library to the 2 µL of H_2O in well#2. Pipetted up and down to mix, and transferred 2 µL into well #3. Mixed by pipetting and transferred 2 µL into the adjacent well. Repeated this procedure until well #11, where the final 2 µL of the serial dilution was discarded. Well #12 served as the blank.

2. Prepared the master mix with Platinum PCR mix:

Component	Volume (µL) 1X	Volume (µL) 14X (12 wells)
Platinum PCR Master Mix	22	308
Library PCR Primer 1, 50 µM	0.5	7
Library PCR Primer 2, 50 µM	0.5	7
AmpliTaq LD 5.0 U/µL	0.5	7

IDT Exhibit 1020.1121

Total	23.5	329

3. Added 23.5 µL of the master mix to each well, bringing the total volume per well to 25.5 µL.
4. Performed 20 cycles of PCR as shown in the following table:

Stage	Step	Temp	Time
Holding	Denature	95 °C	5 min
Cycling (20 cycles)	Denature	95 °C	15 sec
	Anneal	62 °C	15 sec
	Extend	70 °C	1 min
Holding	-	4 °C	∞

5. If library amplification was not detected in any of the wells, SOLiD sequencing was not performed.

Confirmed Library Amplification Using Lonza FlashGel®

1. Added 1 µL 5X FlashGel® Loading Dye to 4 µL from the 100 µL PCR reaction and loaded on a 2.2% Lonza FlashGel®. Loaded FlashGel® DNA Marker (50 bp-1.5 kb or 100 bp-4 kb) in an adjacent well for reference.
2. Ran the FlashGel® for 6 minutes at 275 V.
3. Calculated the optimum number of PCR cycles that provided detectable product.

Amplified Library

1. Performed PCR on the remaining bisulfite-converted library based on using 4 µL of library solution per each 51-µL volume PCR reaction. Dividing 56 µL by 4 µL required 14 X 51 µL PCR reactions. Therefore, 16 X master mix was prepared (for filling the 14 wells), and 47 µL of the master mix was aliquoted into the 14 wells. The 4 µL of template solution was added last and mixed by pipetting up and down a few times.

IDT Exhibit 1020.1122

Component	Volume (μL) 1X	Volume (μL) 16 X
Platinum PCR Master Mix	44	1408
Library PCR Primer 1, 50 μM	1	32
Library PCR Primer 2, 50 μM	1	32
AmpliTaq LD 5.0 U/μL	1	32
Bisulfite Library	4	
Total	51	

2. Prepared the PCR components as shown above. Vortexed to mix and then divided evenly among the required number of PCR wells.
3. Ran the PCR according to the following settings:

Stage	Step	Temp	Time
Holding	Denature	95 °C	5min
Cycling (TBD during trial PCR)	Denature	95 °C	15 sec
	Anneal	62 °C	15 sec
	Extend	70 °C	1 min
Holding	-	4 °C	∞

4. Pooled all of the PCR samples (from like-source, i.e. kept the bisulfite-in-solution together and the bisulfite-on-beads together when processing both) into a 1.5-mL LoBind tube.
5. If the pooled reactions were library amplification from the polystyrene beads, centrifuged at ≥10,000 × *g* (13,000 rpm) for 1 minute.
6. Transferred the pooled supernatant off the beads into a fresh 1.5-mL LoBind tube. Re-suspended the beads in 500 μL of Lo-TE and set aside until successful Bisulfite-SOLiD sequencing was performed.

IDT Exhibit 1020.1123

Purified the DNA with Qiagen QIAquick® Gel Extraction Kit

1. Added 3 volumes Buffer QG and 1 volume isopropyl alcohol to the pooled PCR product. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The color turned yellow. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
2. Applied 750 µL PCR product in Buffer QG two columns.
3. Let the columns stand for 2 minutes at room temperature.
4. Centrifuged the columns at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
5. Repeated steps 2 and 4 until the entire sample had been loaded onto the columns. Placed the MinElute® columns back into the same collection tube.
6. Added 750 µL Buffer PE to wash the columns.
7. Centrifuged the columns at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeat to remove residual wash buffer.
8. Air-dried the columns for 2 minutes to evaporate any residual alcohol. Transferred the columns to clean 1.5-mL LoBind tube(s).
9. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the columns stand for 2 minutes.
10. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
11. If necessary, pooled the eluted DNA.

Gel-purify the library

Size-selecdt the DNA fragments with an agarose gel

1. To the 30 µL of QiaQuick purified library was added 3 µL of 10× PCR buffer and 6 µL of "5X Gel Pilot Loading Dye" resulting in a total volume of 39 µL. This volume required 2 wells of the BioRad precast gel.
2. Loaded 2 µL TrackIt™ 25 bp Ladder. The brightest band for this size ladder was 125 bp. Loaded ~20 µL dye-mixed sample per well. At least one lane was present between the ladder well and the sample wells to avoid contamination of the sample with ladder.
3. Ran the gel at 120 V until the marker was close to the edge of the gel.

IDT Exhibit 1020.1124

4. If needed, stained the gel in 50 to 100 mL 1× TAE or 1× TBE Buffer with 8 µL ethidium bromide (10 mg/mL) for 5 minutes.
5. Destained the gel in nuclease-free water twice for 2 minutes each time and visualized the gel on a UV transilluminator.
6. Excised the entire band which had an average size ranging from 200 to 300 bp using a clean razor blade.

Eluted the DNA Using Qiagen QIAquick® Gel Extraction Kit

1. Weighed the gel slice(s) in a 15-mL polypropylene conical colorless tube.
2. Added 6 volumes Buffer QG to 1 volume of gel.
3. Dissolved the gel slice by vortexing at room temperature until the gel slice had dissolved completely (~5 minutes).
4. If the color of the mixture was yellow, proceeded to step 5. If the color of the mixture was orange or violet, added 10 µL 3 M sodium acetate, pH 5.5 and mixed. The pH required for efficient adsorption of the DNA to the membrane was ≤7.5.
5. Added one gel volume of isopropyl alcohol to the sample and mixed by inverting the tube several times.
6. Applied about 700 µL sample to the column(s). The maximum amount of gel that could be applied to a MinElute® column was 400 mg. Used more columns as necessary.
7. Let the column(s) stand for 2 minutes at room temperature.
8. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute and discarded the flow-through.
9. Repeated steps 6 and 8 until the entire sample had been loaded onto the column(s). Placed the MinElute® column(s) back into the same collection tube.
10. Added 750 µL Buffer PE to wash the column(s).
11. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 2 minutes. Discarded the flow-through. Repeated to remove residual wash buffer.
12. Air-dried the column(s) for 2 minutes to evaporate any residual alcohol. Transferred the column(s) to clean 1.5-mL LoBind tube(s).

IDT Exhibit 1020.1125

13. Added 30 µL Buffer EB to the column(s) to elute the DNA and let the column(s) stand for 2 minutes.
14. Centrifuged the column(s) at ≥10,000 × *g* (13,000 rpm) for 1 minute.
15. If necessary, pooled the eluted DNA in a 1.5-mL LoBind tube.

Quantitate the library by Qbit and BioAnalyzer

Qbit quantitation of the bisulfite-on-bead library was 1.2 ng/µL

Qbit quantitation of the bisulfite-in-solution library was 2.4 ng/µL

Quantitated the library by performing quantitative PCR (qPCR)

Quantitation method	Sensitivity
Lonza 2.2% FlashGel® with FlashGel® QuantLadder	3 ng/µL
Invitrogen Qubit™	200 pg/µL
Agilent Bioanalyzer DNA 1000 Assay	100 pg/µL

Incorporation by Reference

[00152] All references cited herein, including patents, patent applications, papers, text books, and the like, and the references cited therein, to the extent that they are not already, are hereby incorporated by reference in their entirety. In the event that one or more of the incorporated literature and similar materials differs from or contradicts this application, including but not limited to defined terms, term usage, described techniques, or the like, this application controls.

Equivalents

[00153] The foregoing description and Examples detail certain specific embodiments of the present teachings and describes the best mode contemplated by the inventors. It will be appreciated, however, that no matter how detailed the foregoing can appear in text, the present teachings can be practiced in many ways and the invention should be construed in accordance with the appended claims and any equivalents thereof.

IDT Exhibit 1020.1126

What is claimed is:

1. A method of analyzing the methylation state of genomic DNA, comprising:

fragmenting a genomic DNA sample, whereby genomic DNA fragments are produced;

circularizing a genomic DNA fragment to produce a double-stranded circular DNA comprising a nick on at least one strand of the double-stranded circular DNA;

linearizing the circular DNA;

adding a nick translation enzyme in the presence of methylation conversion agent resistant nucleotide triphosphates, whereby a partially methylation conversion agent resistant polynucleotide is generated, wherein the partially methylation conversion agent resistant polynucleotide has a tag region that is methylation conversion agent resistant and a tag region that is not methylation conversion agent resistant.

2. The method of claim 1, further comprising exposing the partially methylation conversion agent resistant polynucleotide to a methylation conversion agent, whereby a conversion agent treated polynucleotide is produced.

3. The method of claim 2, wherein the polynucleotide exposed to the methylation conversion agent is amplified to produce an amplicon.

4. The method of claim 3, further comprising sequencing a region of the amplicon that is derived from the tag region that is methylation conversion agent resistant, and sequencing a region of the amplicon that is derived from the tag region that is not methylation conversion agent resistant .

5. The method of claim 1, wherein the circular DNA comprises a specific binding pair member.

6. The method of claim 5, wherein the specific binding pair member is biotin.

IDT Exhibit 1020.1127

7. The method of claim 5, further comprising the step of attaching adapters to the ends of the partially methylation conversion agent resistant polynucleotide to produce an adapter modified polynucleotide.

8. The method of claims 7, wherein the adapters are double-stranded, and wherein at least one of the stands contains methylation conversion resistant nucleotides and at least one of the strands comprises a first primer binding site sequence.

9. The method of 8, further comprising exposing the adapter modified polynucleotide to a nick translation enzyme and a set of dNTPS.

10. The method of claim 9, further comprising:

exposing the adapter modified polynucleotide to a cognate receptor of the specific binding pair member;

denaturing the adapter modified polynucleotide; and

exposing strands of the adapter modified polynucleotide that are not bound to the cognate receptor to a methylation conversion agent, whereby converted stands are produced.

11. The method of claim 10, further comprising preferentially amplifying the converted strands.

12. The method of claim 11, wherein the preferential amplification introduces a second primer binding site on one end, but not the other end of the preferential amplification product.

13. The method of claim 9, further comprising:

exposing the adapter modified polynucleotide to a cognate receptor of the specific binding pair member;

denaturing the adapter modified polynucleotide; and

IDT Exhibit 1020.1128

separating strands of the adapter modified polynucleotide that are not bound to the cognate receptor from strands of the adapter modified polynucleotide that are bound to the cognate receptor.

14. The method of claim 13, wherein the cognate receptor is bound to a solid support.

15. The method of claim 14, wherein the cognate receptor comprises streptavidin bound to non-magnetic polystyrene beads.

16. The method of claim 13, further comprising:

exposing at least one of the strands of the adapter modified polynucleotide that are not bound to the cognate receptor and strands of the adapter modified polynucleotide that are bound to the cognate receptor to a methylation conversion agent, whereby converted stands are produced.

17. The method of claim 16, further comprising preferentially amplifying the converted strands.

18. The method of claim 17, wherein the preferential amplification introduces a second primer binding site on one end of the preferential amplification product but not on the other end of the preferential amplification product.

19. A method of analyzing the methylation state of genomic DNA, comprising:

fragmenting a genomic DNA sample, whereby genomic DNA fragments are produced;

forming a first tag sequence and a second tag sequence, wherein the first tag sequence and the second tag sequence are derived from a single genomic DNA fragment;

wherein the first tag sequence has been converted by a methylation conversion agent and the second tag sequence has not been converted by a methylation conversion agent.

IDT Exhibit 1020.1129

20. The method of claim 19, wherein the first tag sequence and the second tag sequence are present on a single polynucleotide molecule.

21. The method of claim 20, further comprising amplifying the single polynucleotide molecule to produce an amplicon.

22. The method of claim 21, wherein the amplification is clonal amplification.

23. The method of claim 21, wherein the clonal amplification is solid phase amplification.

24. The method of claim 22, wherein the clonal amplification is emulsion PCR.

25. A polynucleotide construction comprising a first tag sequence and a second tag sequence, wherein the first tag sequence and the second tag sequence are derived from a single fragment of genomic DNA, wherein the first tag comprises methylation conversion resistant nucleotide that have been incorporated into the construction by an in vitro reaction and the second tag does not comprise methylation conversion resistant nucleotide that have been incorporated into the construction by an in vitro reaction.

26. The polynucleotide construction of claim 25, further comprising an internal adapter located between the first tag and the second tag.

27. The polynucleotide construction of claim 26, wherein the internal adapter comprises a specific binding pair member.

28. The polynucleotide construction of claim 27, further comprising primer binding sequences located in functional proximity to the first tag sequence and the second tag sequence, wherein amplification primers binding to the priming sites can amplify both the first and the second tag sequences.

IDT Exhibit 1020.1130

29. An adapter comprising a first strand having methylation conversion resistant nucleotides and a second strand complementary to the first strand, wherein the second strand optionally contains methylation conversion resistant nucleotides.

30. A kit comprising an adapter of claim 29 and oligonucleotide primers specific for a strand of the adapter.

31. A method of matching a DNA sequence to a genomic sequence database, said method comprising:

comparing a data record comprising (1) a first tag sequence that corresponds to a DNA sequence that has not been modified by a methylation conversion agent and (2) a second tag sequence that corresponds to a DNA sequence that may have been modified by a methylation conversion agent, with DNA sequence information in the genomic database.

32. The method of claim 31, wherein comparing the data record uses a value indicative of the approximate distance in the genome between the first tag sequence and the second tag sequence.

33. The method of claim 32, further comprising detecting a first nucleic acid sequence in the genomic sequence database that corresponds to the first tag sequence and detecting a second nucleic acid sequence in the genomic sequence database that corresponds to the second tag sequence.

34. The method of claim 33, further comprising comparing the second tag sequence with the corresponding genomic reference sequence to detect sequence differences indicative of methylation of a region of genomic DNA from which the second tag sequence was derived.

35. The method of claim 33, further comprising:

IDT Exhibit 1020.1131

comparing a plurality of second tag sequence with a corresponding genomic reference sequence;

determining a value or set of values indicative of the degree of methylation of a base or bases in the second tag sequence; and

displaying the value or set of values indicative of the degree of methylation of a base or bases in the second tag sequence.

36. A method of amplifying polynucleotides converted by a methylation conversion agent, comprising;

providing a polynucleotide fragment having two termini;

ligating a primer-adapter to both of the termini, wherein the primer-adapter is a double-stranded polynucleotide having a first stand and second strand complementary to the first strand, wherein the first strand comprises methylation conversion resistant nucleotides and the second strand optionally comprises methylation conversion resistant nucleotides, whereby an adapter modified polynucleotide is produced;

exposing the adapter-modified polynucleotide to a methylation conversion reagent, whereby a converted adapter modified polynucleotide is produced; and

amplifying the converted adapter modified polynucleotide, wherein amplifying the converted adapter modified polynucleotide uses primers specific for sequences in the second strand of the adapter.

37. The method of claim 36, further comprising:

denaturing the adapter modified polynucleotide to produce separated strands;

enriching one of the separated strands; and

performing the amplification step on the enriched strand.

38. A method of analyzing the methylation state of a genomic DNA sample, said method comprising:

mixing a DNA sample with formamide, whereby a sample mixture is formed;

heating the sample mixture at temperature sufficient to denature the genomic DNA; and

IDT Exhibit 1020.1132

adding a bisulfite salt to the sample mixture.

39. The method of claim 38, wherein the formamide concentration in the sample mixture prior to the addition of the bisulfite salt is a least 50%.

40. The method of claim 39, wherein the formamide concentration in the sample mixture prior to the addition of the bisulfite salt is a least 75%.

41. The method of claim 40, wherein the formamide concentration in the sample mixture prior to the addition of the bisulfite salt is a least 90%.

42. The method of claim 41, wherein the formamide concentration in the sample mixture prior to the addition of the bisulfite salt is a least 95%.

43. The method of claim 38, wherein the DNA sample is present in a gel matrix.

44. The method of claim 43, wherein the gel matrix comprise polyacrylamide.

45. The method of claim 38, wherein the DNA sample is derived from a paraffin embedded sample.

46. The method of claim 43, further comprising the step of amplifying the DNA sample in the gel matrix, wherein the amplification occurs within the matrix and the amplification occurs after the bisulfite has been added.

47. A method of analyzing the methylation state of a polynucleotide, comprising:

providing a polynucleotide fragment having two termini;

ligating a primer-adapter to both of the termini;

circularizing the adapter-modified polynucleotide with an internal adapter to produce a double-stranded circular polynucleotide comprising a nick on one strand of the circular polynucleotide, wherein the internal adapter comprises a specific binding moiety;

IDT Exhibit 1020.1133

nick-translating the circular polynucleotide;

capturing the strand comprising the specific binding moiety with a cognate specific binding moiety on a solid support;

separating the captured strand and the non-captured strand; and

exposing at least one of the captured strand and the non-captured strand to a methylation conversion reagent, whereby at least one converted strand is produced; and

sequencing the at least one converted strand.

48. The method of claim 47, wherein the specific binding moiety comprises biotin.

49. The method of claim 48, wherein the cognate specific binding moiety is chosen from avidin and streptavidin.

50. The method of claim 49, wherein the solid support comprises a non-magnetic polystyrene bead.

IDT Exhibit 1020.1134

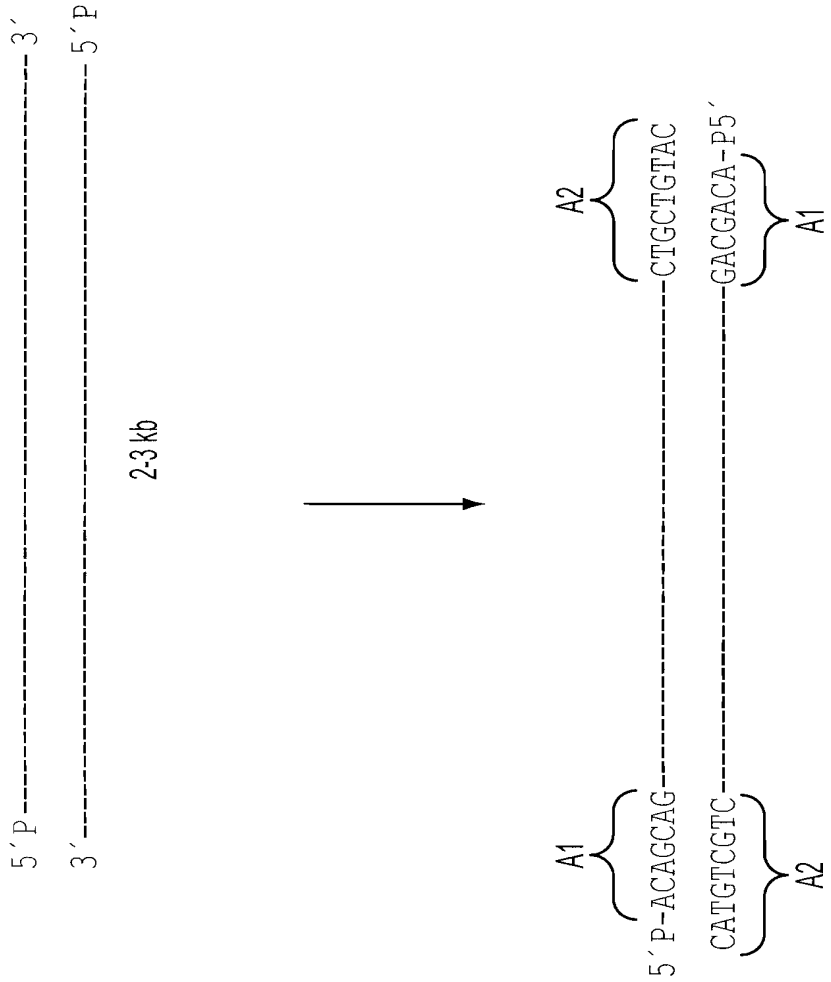

5′P
3′
3′
5′P
2-3 kb
A1
5′P-ACAGCAG
CATGTCGTC
A2
A2
CTGCTGTAC
GACGACA-P5′
A1

FIG. 1

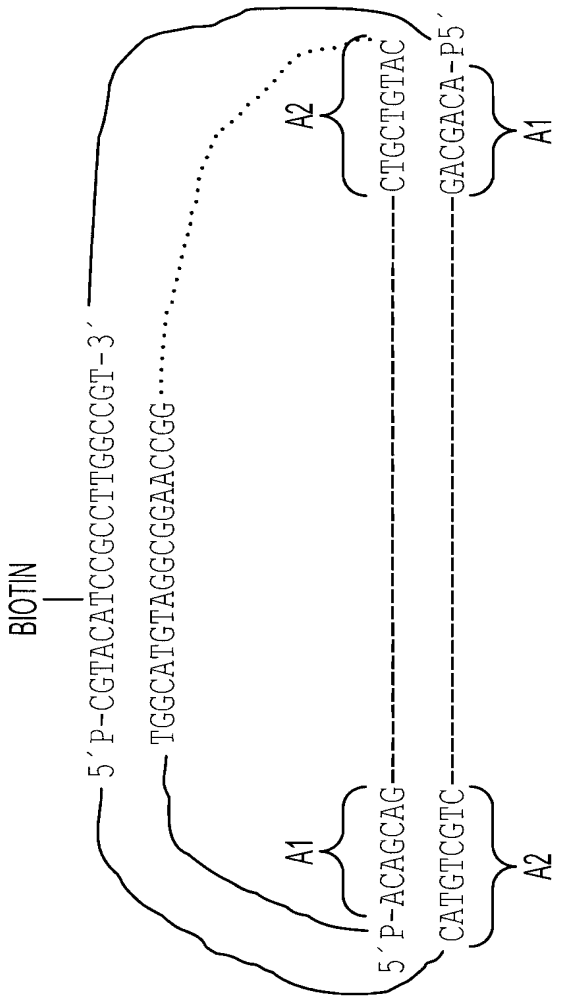

BIOTIN
5´P-CGTACATCCGCCTTGGCCGT-3´
TGGCATGTAGGCGGAACCGG
A2
CTGCTGTAC
GACGACA-P5´
A1
A1
5´P-ACAGCAG
CATGTCGTC
A2

FIG. 2

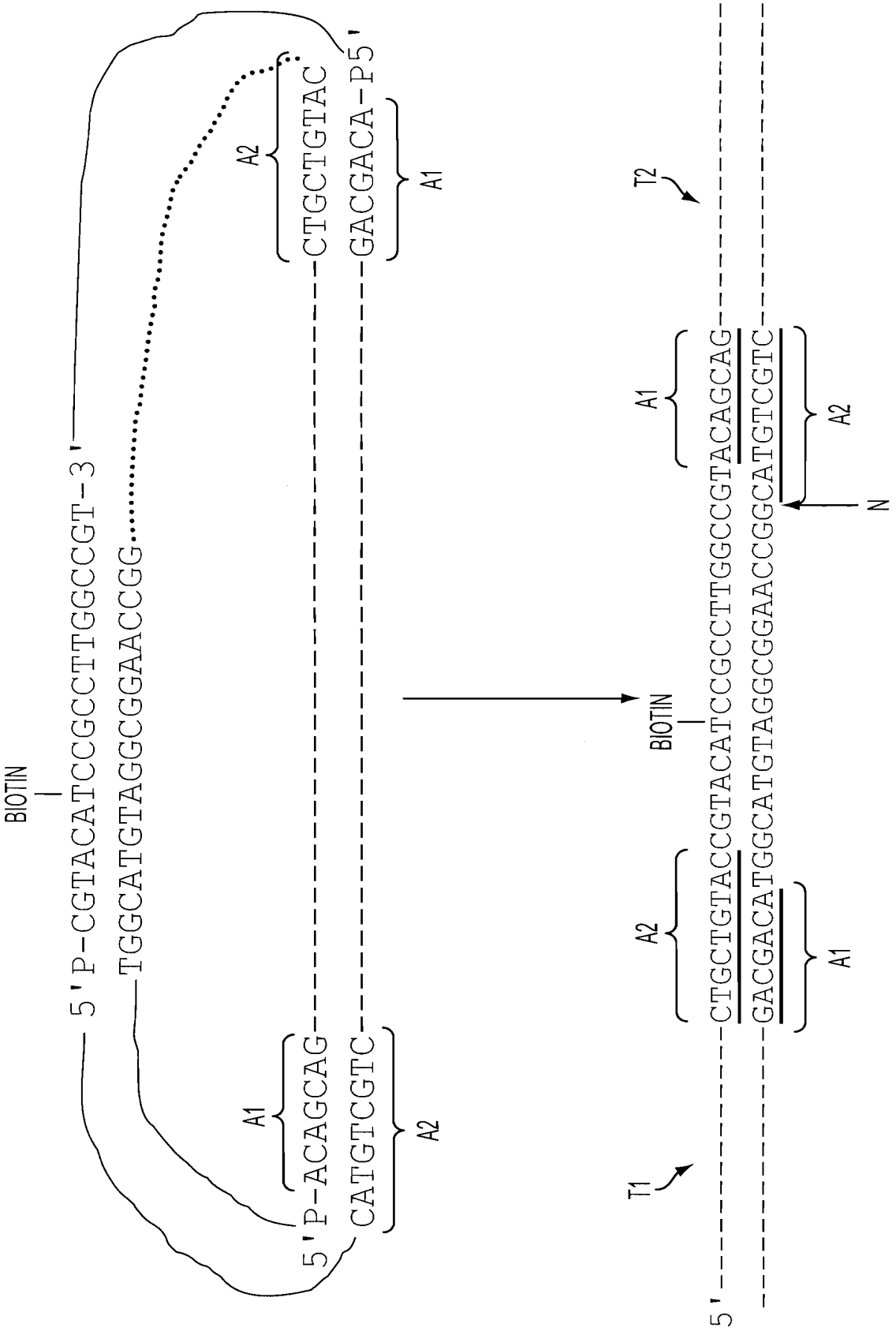

BIOTIN
5'P-CGTACATCCGCCTTGGCCGT-3'
TGGCATGTAGGCGGAACCGG
A2
CTGCTGTAC
GACGACA-P5'
A1
A1
5'P-ACAGCAG
CATGTCGTC
A2
T1
T2
BIOTIN
A2
A1
CTGCTGTACCGTACATCCGCCTTGGCCGTACAGCAG
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
5'
A1
A2
N

FIG. 3

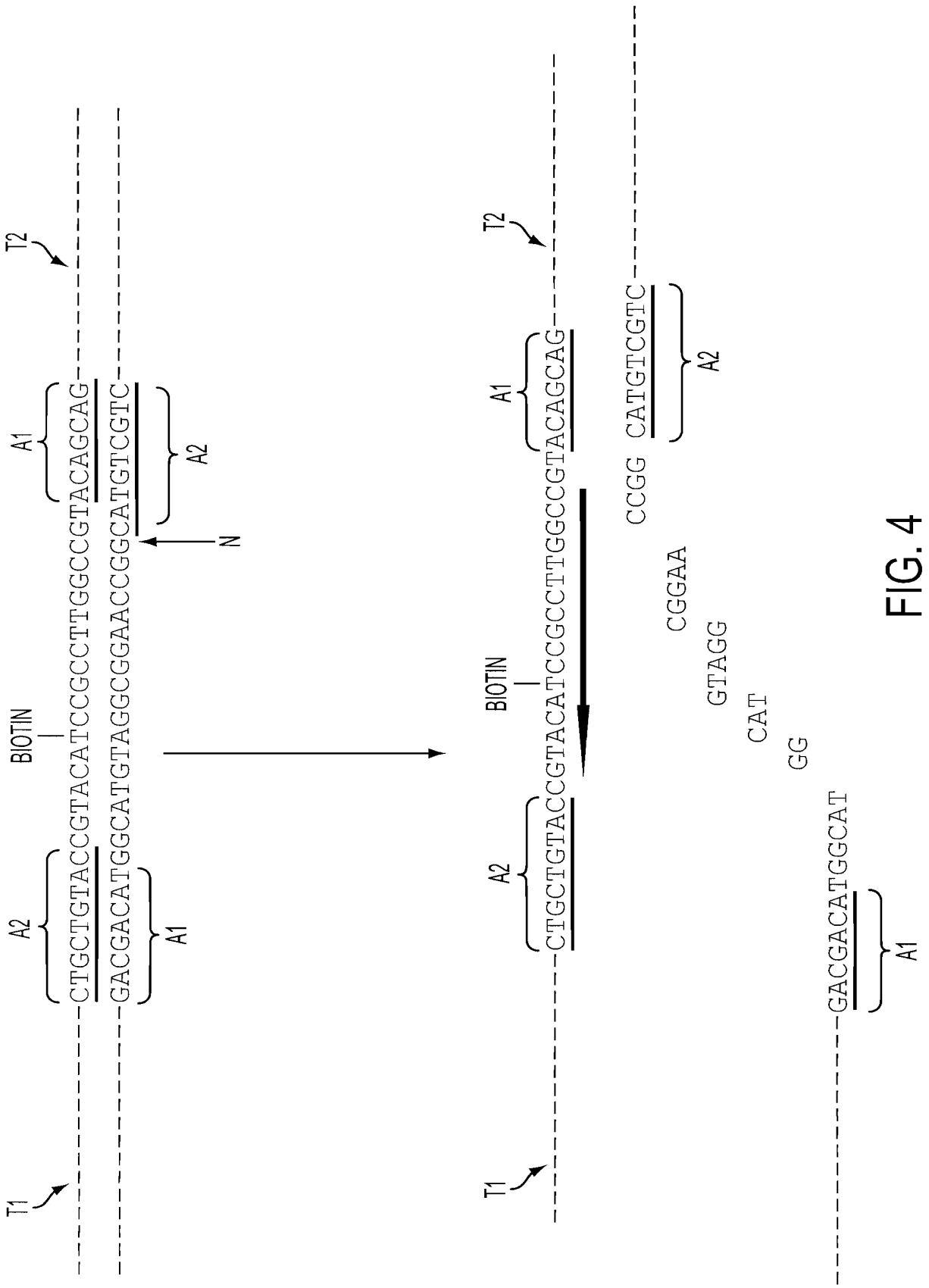

T2
A1
A2
BIOTIN
CTGCTGTACCGTACATCCGCCTTGGCCGTACAGCAG
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
A2
A1
N
T1
T2
A1
A2
BIOTIN
CTGCTGTACCGTACATCCGCCTTGGCCGTACAGCAG
CCGG CATGTCGTC
A2
CGGAA
GTAGG
CAT
GG
GACGACATGGCAT
A1
T1

FIG. 4

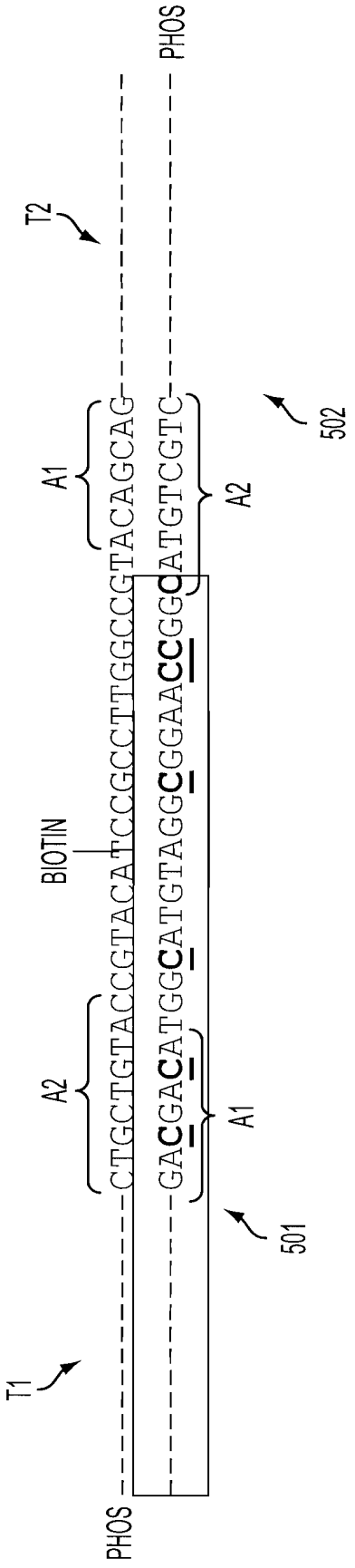

T2
PHOS
A1
A2
502
BIOTIN
CTGCTGTACCGTACATCCGCCTTGGCCGTACAGCAG
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
A2
A1
501
T1
PHOS

FIG. 5

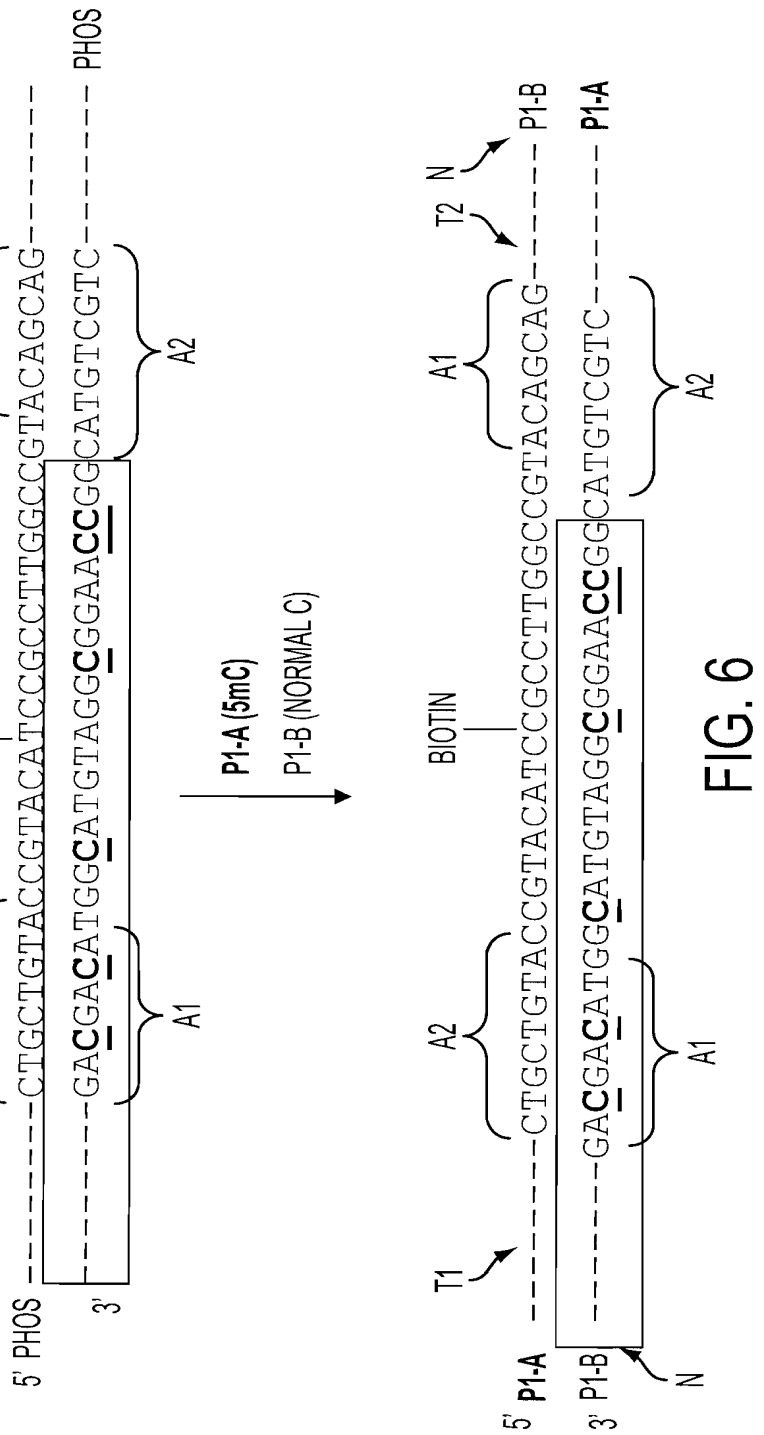


FIG. 6

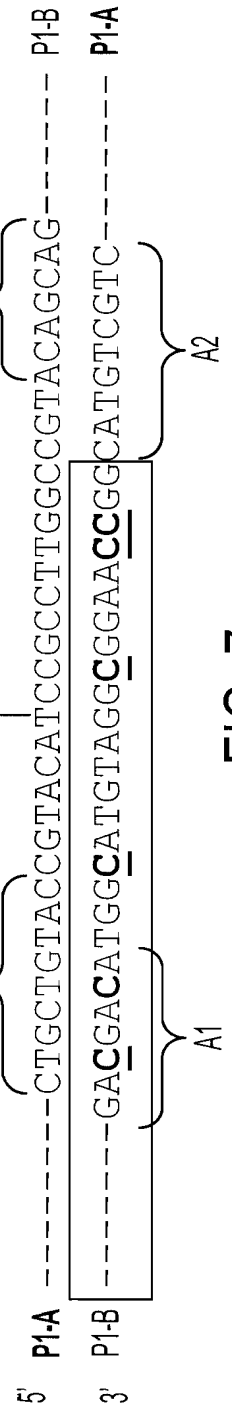

5'
P1-A
CTGCTGTACCGTACATCCGCCTTGGCCGTACAGCAG
P1-B
3'
P1-B
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
P1-A
A2
A1
BIOTIN
A1
A2

FIG. 7


T2
P1-A
A2
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
A1
T1
P1-B
3'

FIG. 8

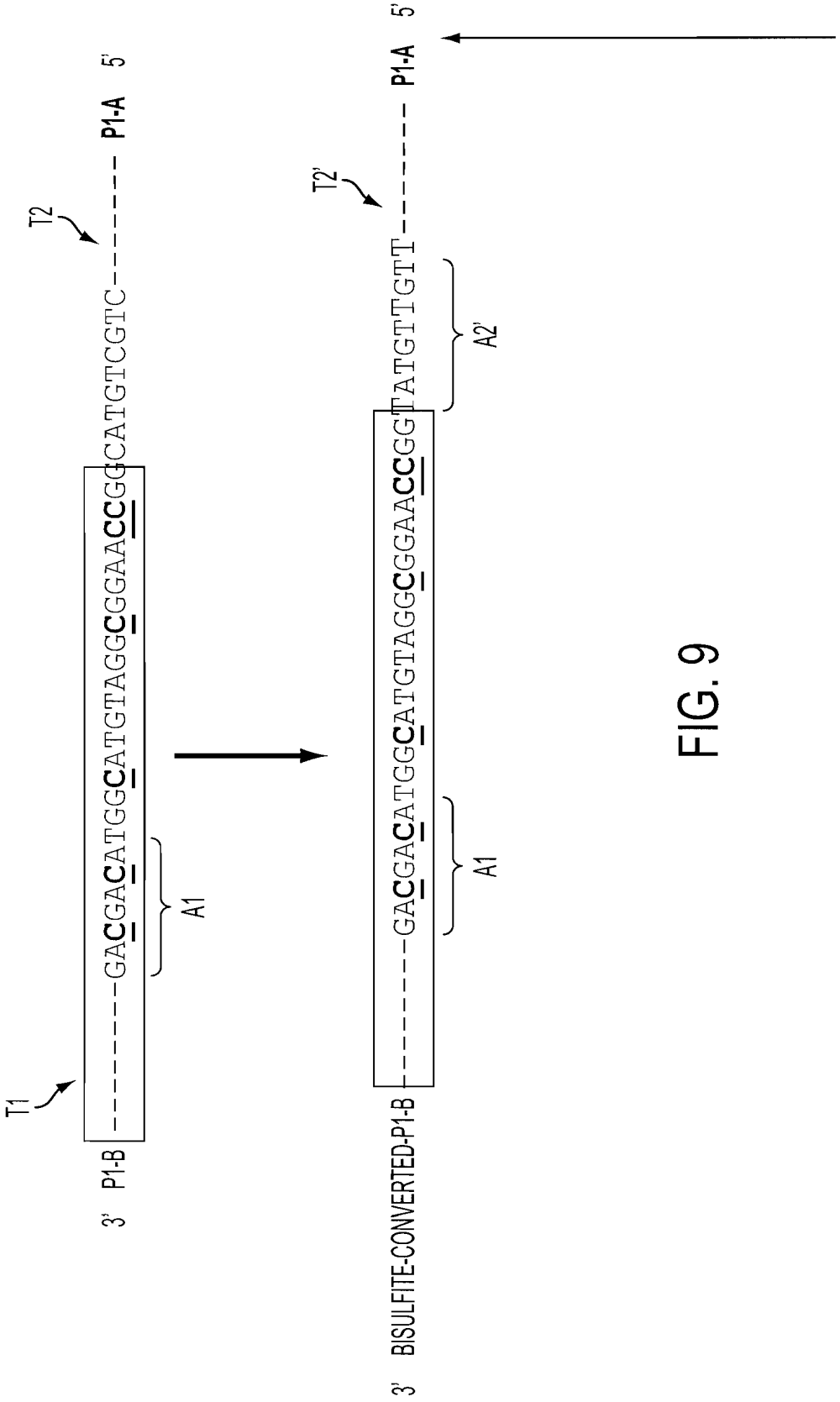

T2
P1-A 5'
GACGACATGGCATGTAGGCGGAACCGGCATGTCGTC
A1
T1
3' P1-B
T2'
5' P1-A
TATGTTGTGTT
A2'
GACGACATGGCATGTAGGCGGAACCGG
A1
3' BISULFITE-CONVERTED-P1-B

FIG. 9

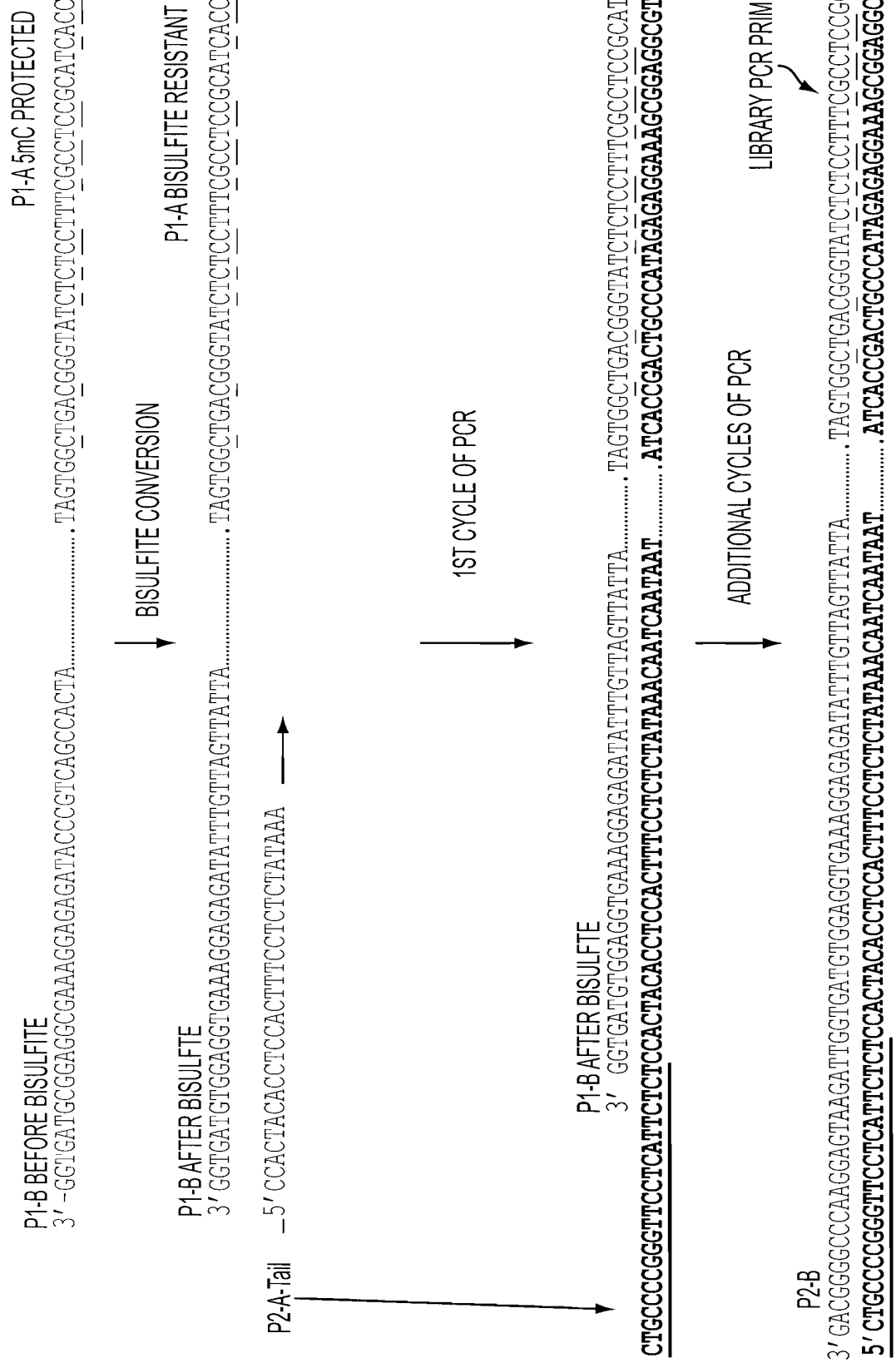

P1-A 5mC PROTECTED
P1-B BEFORE BISULFITE
3′ -GGTGATGCGGAGGCGAAAGGAGAGATACCCGTCAGCCACTA..... TAGTGGCTGACGGGTATCTCTCCTTTCGCCTCCGCATCACC-5′
BISULFITE CONVERSION
P1-A BISULFITE RESISTANT
P1-B AFTER BISULFTE
3′ GGTGATGTGGAGGTGAAAGGAGAGATATTTGTTAGTTATTA..... TAGTGGCTGACGGGTATCTCTCCTTTCGCCTCCGCATCACC-5′
1ST CYCLE OF PCR
P2-A-Tail _5′ CCACTACACCTCCACTTTCCTCTCTATAAA
P1-B AFTER BISULFTE
3′ GGTGATGTGGAGGTGAAAGGAGAGATATTTGTTAGTTATTA..... TAGTGGCTGACGGGTATCTCTCCTTTCGCCTCCGCATCACC-5′
CTGCCCCCGGGTTCCTCATTCTCTCCACTACACCTCCACTTTCCTCTCTATAAACAATCAATAAT..... ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG
ADDITIONAL CYCLES OF PCR
LIBRARY PCR PRIMER 1
P2-B
3′ GACGGGGCCCAAGGAGTAAGATTGGTGATGTGGAGGTGAAAGGAGAGATATTTGTTAGTTATTA..... TAGTGGCTGACGGGTATCTCTCCTTTCGCCTCCGCATCACC-5′
5′ CTGCCCCCGGGTTCCTCATTCTCTCCACTACACCTCCACTTTCCTCTCTATAAACAATCAATAAT..... ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG
LIBRARY PCR PRIMER 2

FIG. 10

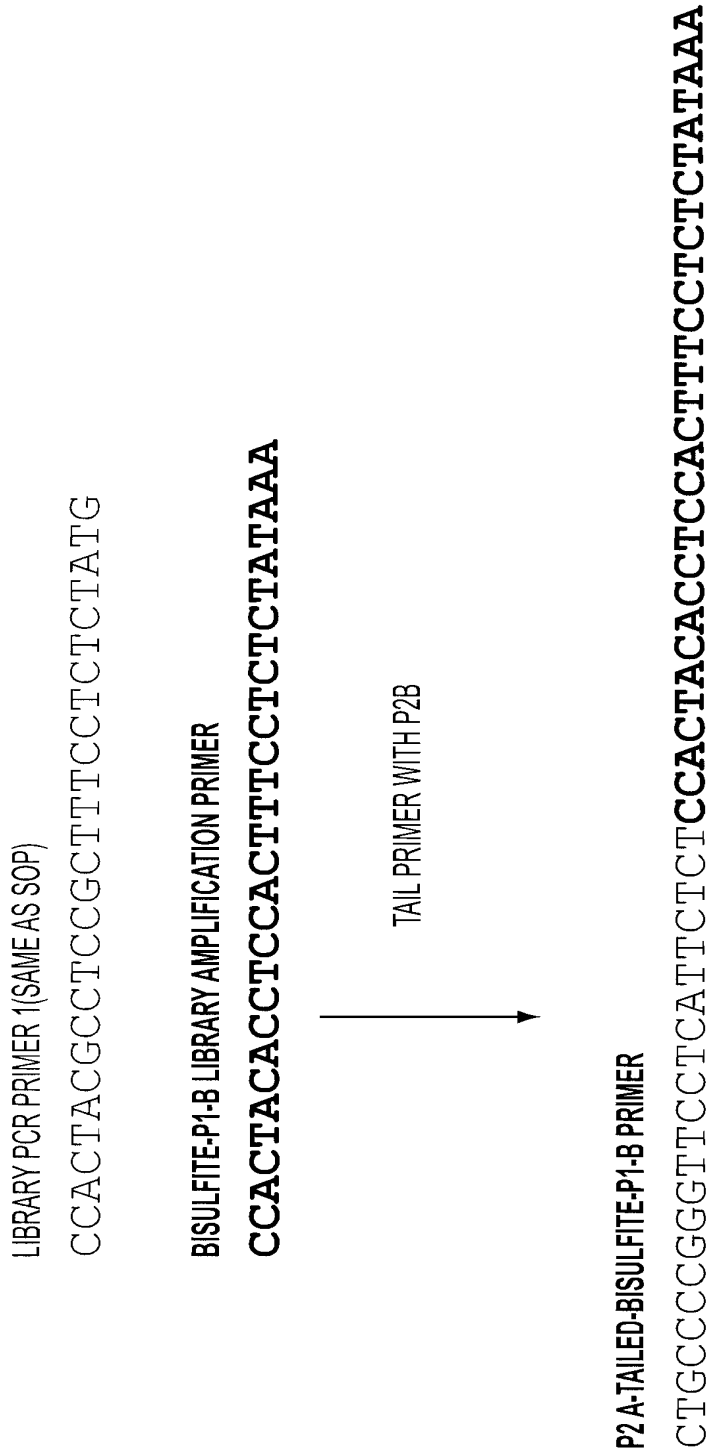

LIBRARY PCR PRIMER 1(SAME AS SOP)
CCACTACGCCTCCGCTTTCCTCTCTATG
BISULFITE-P1-B LIBRARY AMPLIFICATION PRIMER
CCACTACACCTCCACTTTCCTCTCTATAAA
TAIL PRIMER WITH P2B
P2 A-TAILED-BISULFITE-P1-B PRIMER
CTGCCCCGGGTTCCTCATTCTCTCCACTACACCTCCACTTTCCTCTCTATAAA

FIG. 11

1. INTERNAL ADAPTER

BIOTIN

5′ P-CGTACATCCGCCTTGGCCGT-3′

3′ -TGGCATGTAGGCGGAACCGG-5′

2. **5mC P1-A**/P1-B ADAPTER

5′ **CC**AC**C**TAC**G**C**C**TC**C**GC**T**TT**CC**T**C**T**C**TATGGG**C**AGT**C**GGTGAT

3′ TTGGTGATGCGGAGGCGAAAGGAGAGATACCCGTCAGCCACTA

3. P2-A-TAILED LIBRARY AMPLIFICATION PRIMER

CTGCCCCGGGTTCCTCATTCTCTCCACTACACCTCCACTTTCCTCTCTATAAA

FIG. 12

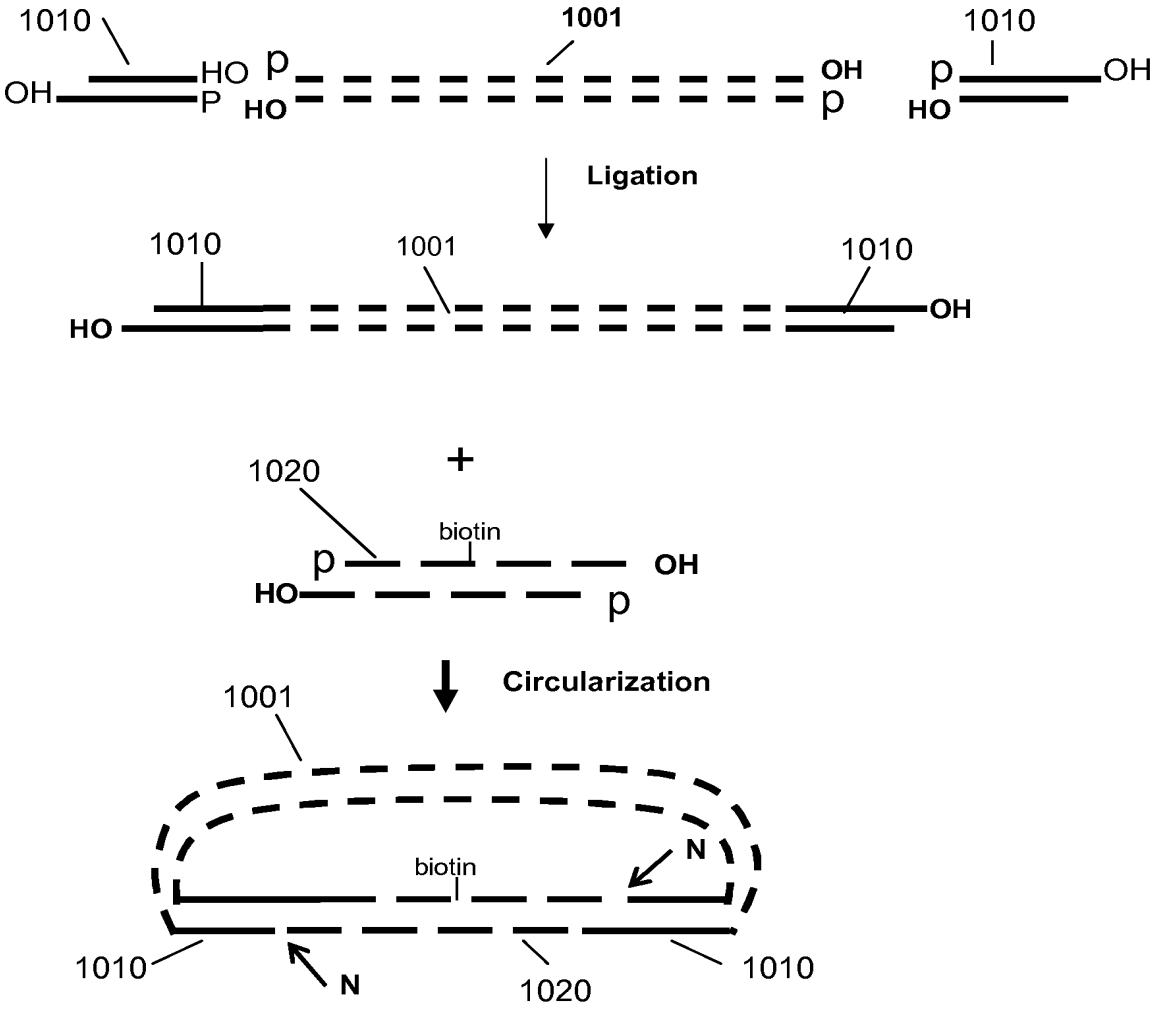

1010
1001
1010
OH
HO
P
p
HO
OH
p
p
HO
OH
Ligation
1010
1001
1010
HO
OH
+
1020
biotin
p
OH
HO
p
Circularization
1001
N
biotin
1010
N
1020
1010

FIG. 13

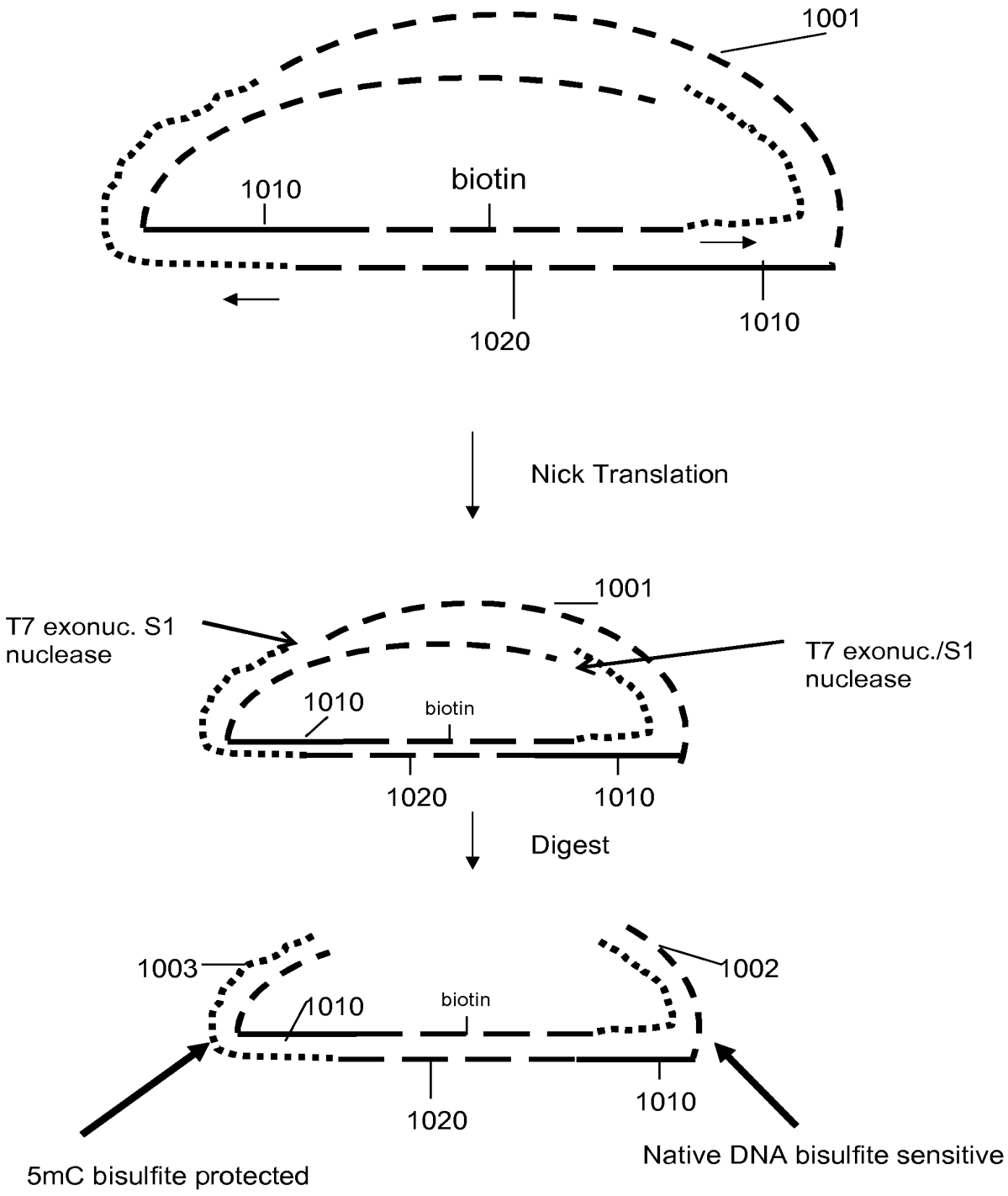

1001
1010
biotin
1020
1010
Nick Translation
1001
T7 exonuc. S1 nuclease
T7 exonuc./S1 nuclease
1010
biotin
1020
1010
Digest
1003
1002
1010
biotin
1020
1010
5mC bisulfite protected
Native DNA bisulfite sensitive

FIG. 14

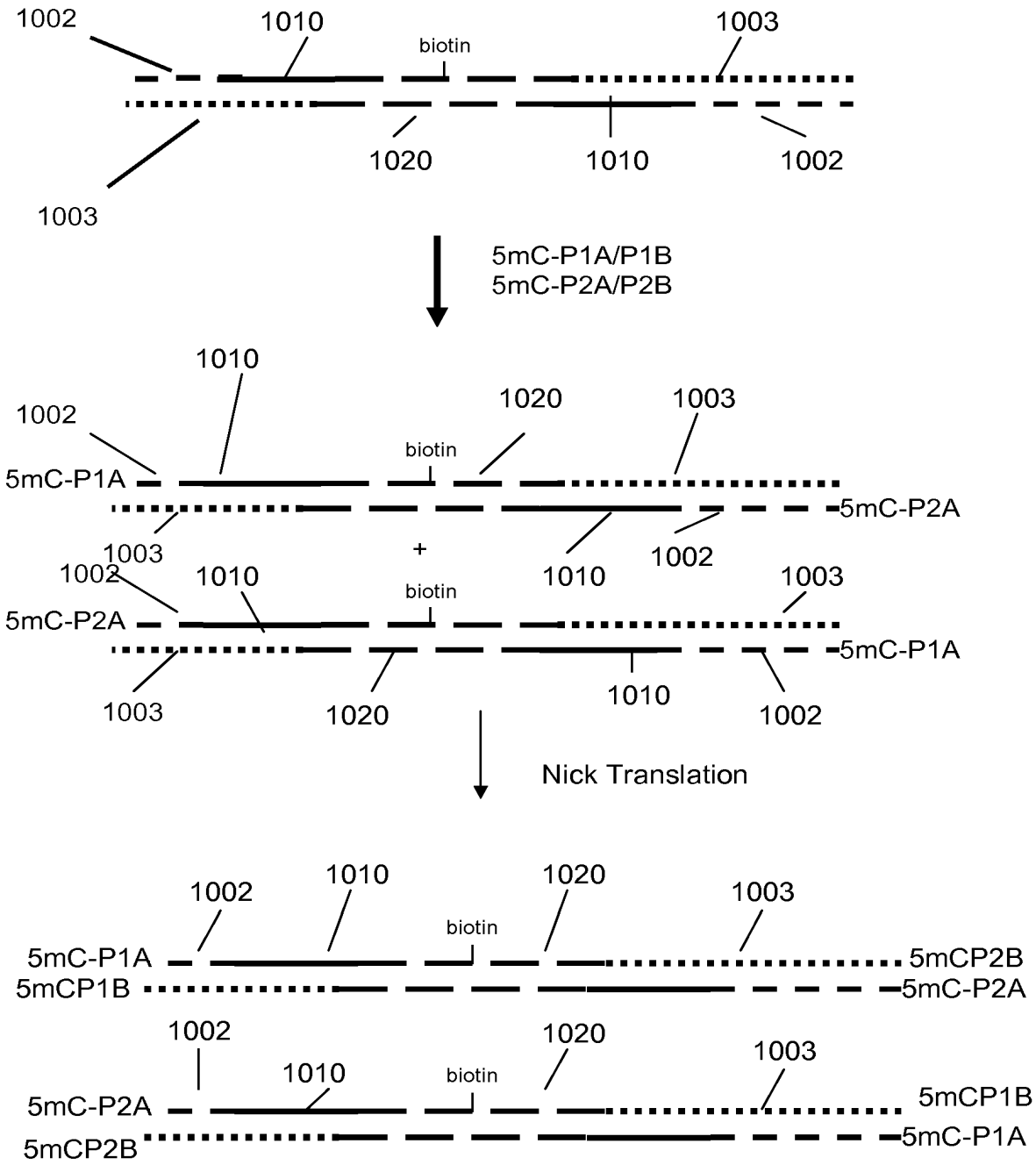

1002
1010
biotin
1003
1020
1010
1002
1003
5mC-P1A/P1B
5mC-P2A/P2B
1010
1002
1020
1003
biotin
5mC-P1A
5mC-P2A
1003
+
1010
1002
1002
1010
biotin
1003
5mC-P2A
5mC-P1A
1003
1020
1010
1002
Nick Translation
1002
1010
1020
1003
biotin
5mC-P1A
5mCP2B
5mCP1B
5mC-P2A
1002
1020
1010
biotin
1003
5mC-P2A
5mCP1B
5mCP2B
5mC-P1A

FIG. 15

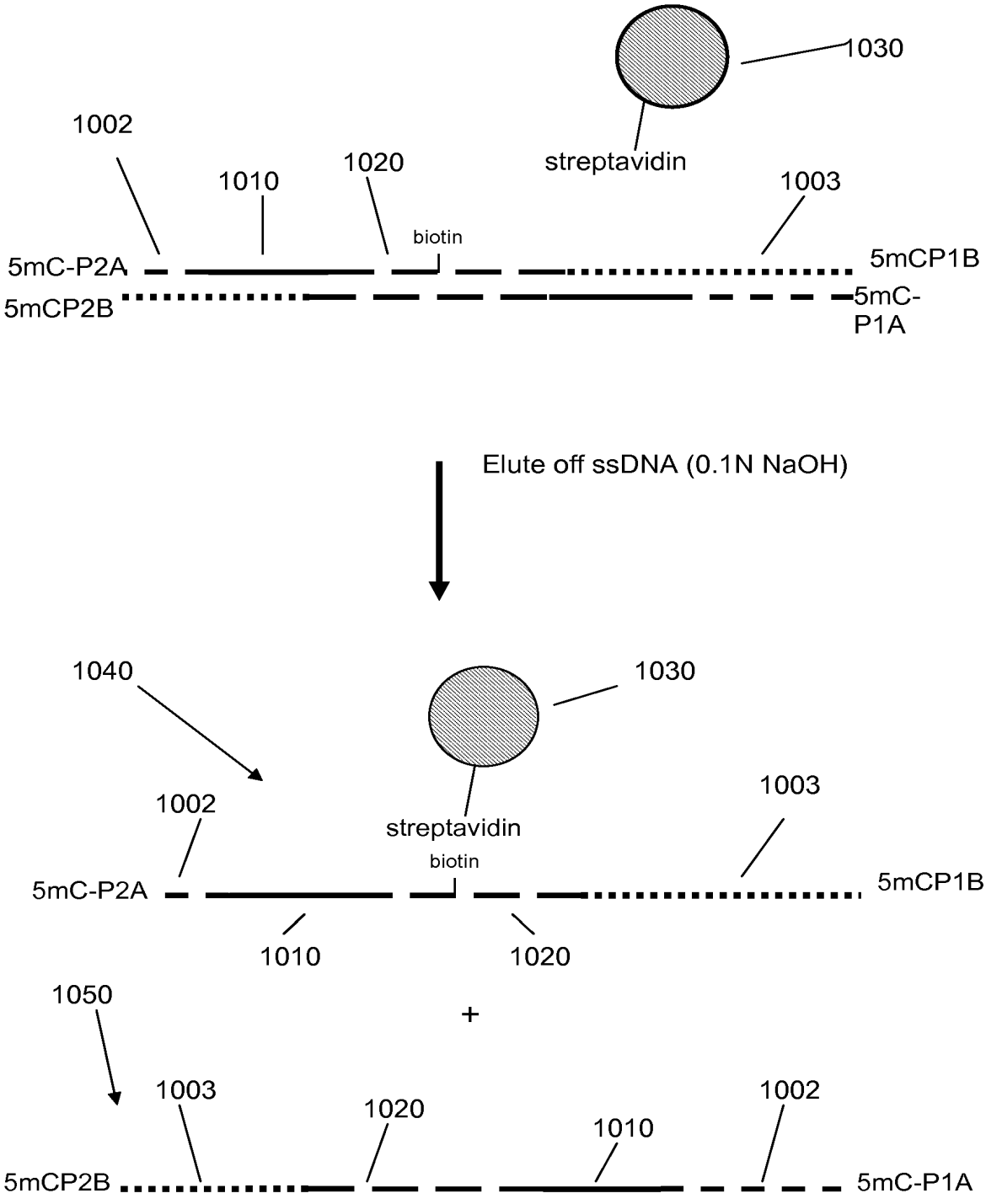

1030
1002
1010
1020
streptavidin
1003
biotin
5mC-P2A
5mCP1B
5mCP2B
5mC-
P1A
Elute off ssDNA (0.1N NaOH)
1040
1030
1002
streptavidin
1003
biotin
5mC-P2A
5mCP1B
1010
1020
1050
+
1003
1020
1010
1002
5mCP2B
5mC-P1A

FIG. 16

Espacenet

Bibliographic data: WO2009120372 (A2) — 2009-10-01

COMPOSITIONS AND METHODS FOR NUCLEIC ACID SEQUENCING

Inventor(s): TRAVERS KEVIN [US]; OTTO GEOFF [US]; TURNER STEPHEN [US]; HEINER CHERYL [US]; MA CONGCONG [US] ± (TRAVERS, KEVIN, ; OTTO, GEOFF, ; TURNER, STEPHEN, ; HEINER, CHERYL, ; MA, CONGCONG)

Applicant(s): PACIFIC BIOSCIENCES CALIFORNIA [US]; TRAVERS KEVIN [US]; OTTO GEOFF [US]; TURNER STEPHEN [US]; HEINER CHERYL [US]; MA CONGCONG [US] ± (PACIFIC BIOSCIENCES OF CALIFORNIA, INC, ; TRAVERS, KEVIN, ; OTTO, GEOFF, ; TURNER, STEPHEN, ; HEINER, CHERYL, ; MA, CONGCONG)

Classification: **- international:** *C12Q1/68*
- cooperative: C12Q1/6806; C12Q1/6855; C12Q1/6869; C12Q1/689; G06F19/22 more

Application number: WO2009US01926 20090327 Global Dossier

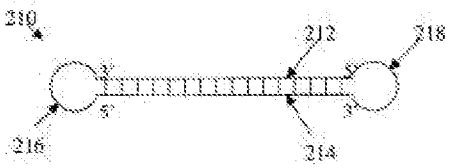
Priority number(s): US20080072160P 20080328 ; US20080099696P 20080924

Also published as: WO2009120372 (A3) AU2009229157 (A1) AU2009229157 (B2) BRPI0909212 (A2) CA2719747 (A1) more

Abstract of WO2009120372 (A2)

Compositions and methods for nucleic acid sequencing include template constructs that comprise double stranded portions in a partially or completely contiguous constructs, to provide for redundant sequence determination through one or both of sequencing sense and antisense strands, and iteratively sequencing the entire construct multiple times. Additional sequence components are also optionally included within such template constructs. Methods are also provided for the use and preparation of these constructs as well as sequencing compositions for their application.

Fig. 2B



IDT Exhibit 1020.1151

IDT Exhibit 1020.1152

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 October 2009 (01.10.2009)

PCT

(10) International Publication Number
WO 2009/120372 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2009/001926

(22) International Filing Date:
27 March 2009 (27.03.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/072,160 28 March 2008 (28.03.2008) US
61/099,696 24 September 2008 (24.09.2008) US

(71) Applicant *(for all designated States except US)*: **PACIFIC BIOSCIENCES OF CALIFORNIA, INC.** [US/US]; 1505 Adams Drive, Menlo Park, CA 94025 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TRAVERS, Kevin** [US/US]; 3385 Gonzaga Place, Santa Clara, CA 95051 (US). **OTTO, Geoff** [US/US]; 790 Knoll Drive, San Carlos, CA 94070 (US). **TURNER, Stephen** [US/US]; 356 Stanford Avenue, Menlo Park, CA 94025 (US). **HEINER, Cheryl** [CN/US]; 1952 Ticonderoga Drive, San Mateo, CA 94402 (US). **MA, Congcong** [US/US]; 687 Portofino Lane, Foster City, CA 94404 (US).

(74) Agents: **LITTLEPAGE, Paul** et al.; QUINE INTELLECTUAL PROPERTY LAW GROUP, P.C., P.O. Box 458, Alameda, CA 94501 (US).

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

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

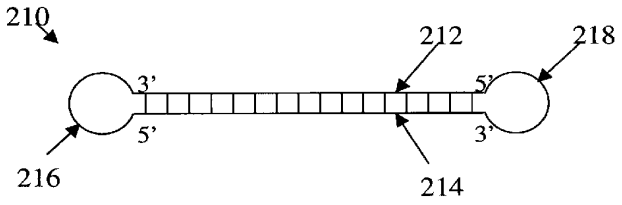
Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2009/120372 A2

(54) **Title**: COMPOSITIONS AND METHODS FOR NUCLEIC ACID SEQUENCING

Fig. 2B



(57) **Abstract**: Compositions and methods for nucleic acid sequencing include template constructs that comprise double stranded portions in a partially or completely contiguous constructs, to provide for redundant sequence determination through one or both of sequencing sense and antisense strands, and iteratively sequencing the entire construct multiple times. Additional sequence components are also optionally included within such template constructs. Methods are also provided for the use and preparation of these constructs as well as sequencing compositions for their application.

IDT Exhibit 1020.1153

COMPOSITIONS AND METHODS FOR NUCLEIC ACID SEQUENCING

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of Provisional U.S. Patent Application No. 61/072,160, filed March 28, 2008, and Provisional U.S. Patent Application No. 61/099,696, filed September 24, 2008, the full disclosures of which are incorporated herein by reference in their entirety for all purposes.

[0002] This application is also related to Provisional U.S. Patent Application No. 61/139,402, filed December 19, 2008, U.S. Patent Application No. [unassigned], filed March 27, 2009 (Attorney Docket No. 01-007701) and U.S. Patent Application No. [unassigned], filed March 27, 2009 (Attorney Docket No. 01-005902), the full disclosures of which are hereby incorporated herein by reference in their entirety for all purposes.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0003] Not Applicable.

BACKGROUND OF THE INVENTION

[0004] The ability to understand the genetic code that serves as the blueprint for the framework of all life has yielded countless advances in countless areas. From the ability to diagnose disease to the ability to identify evolutionary connections and/or diversity, to the ability to manipulate the genetic framework in the development of new materials and compositions, this understanding has opened doors to innumerable advances that have benefitted and will continue to benefit the human race.

[0005] Integral to these advances have been the advances in technology directed to the reading and/or characterization of the genetic code. For example, development of nucleic acid sequencing technologies has allowed for the base by base identification of the nucleic acid sequences that make up the genetic code to the point that entire human genomes have been elucidated. Other advances include rapid array based technologies that allow reasonably facile identification of genetic patterns from patients or other biological samples.

[0006] With each technological advance, there exist opportunities to further improve the state of the art through advances in related or ancillary technologies associated with those advanced

IDT Exhibit 1020.1154

areas. For example, advances in fluorescent dye chemistries have fueled many advances in genetic technologies by permitting simple optical analyses of biological reactions and their products. Likewise, development of microfluidic technologies have provided for advances in fluid and reagent handling to yield a reproducibility that had not been previously achievable through more conventional means.

[0007] The present invention is directed to improved processes, systems and compositions used in genetic analysis that can yield enhanced accuracy and ease of use in such analyses.

BRIEF SUMMARY OF THE INVENTION

[0008] The present invention provides improved nucleic acid template constructs, compositions, kits and systems incorporating such constructs, and methods for preparing and using such constructs.

[0009] In a first aspect, the invention provides a method of determining a consensus sequence of nucleotides in a template nucleic acid segment. The method comprises providing a sense and an antisense strand of the template nucleic acid segment in a contiguous nucleic acid molecule. Both the sense and antisense strands are then sequenced in a polymerase mediated, template dependent sequencing process. The consensus sequence of the target nucleic acid segment is then determined from a sequence of the sense and antisense strands.

[0010] In a related aspect, the invention provides methods of sequencing a nucleic acid sequence, that comprise providing a template nucleic acid that comprises a double stranded nucleic acid segment having a first and second end. A first hairpin oligonucleotide connects each strand of the template nucleic acid at the first end, and a second hairpin oligonucleotide connects each strand of the template nucleic acid at the second end. The nucleotide sequence of at least one strand of the template nucleic acid is then determined using a template directed, polymerase mediated nucleic acid sequencing process.

[0011] In still another aspect, the invention provides methods of sequencing a nucleic acid sequence, that comprise providing a template nucleic acid that comprises a double stranded segment having a first and second end, where a first hairpin oligonucleotide connects each strand of the template nucleic acid at the first end. The nucleotide sequence of the template nucleic acid is then determined using a template directed, polymerase mediated nucleic acid sequencing process.

IDT Exhibit 1020.1155

[0012] In further aspects are provided methods of sequencing a nucleic acid segment, that comprise providing a template nucleic acid segment. The segment comprises first and second complementary nucleic acid strands, a first connecting nucleic acid linking a 3' end of the first nucleic acid segment with a 5' end of the second nucleic acid segment, and a second connecting nucleic acid linking the 5' end of the first nucleic acid strand to the 3' end of the second nucleic acid strand. The template nucleic acid sequence is then contacted with a primer sequence complementary to at least a portion of the template nucleic acid. A template dependent primer extension reaction is then monitored to determine a sequence of nucleotides in the template nucleic acid.

[0013] In still a further aspect, the invention provides a method of sequencing a nucleic acid, comprising providing a template nucleic acid comprising a double stranded segment comprising first and second complementary strands and at least a first single stranded oligonucleotide segment linking a 3' end of the first strand to the 5' end of the second strand; and monitoring nucleotides incorporated in a template dependent synthesis reaction to identify a sequence of nucleotides in the first strand, the linking oligonucleotide segment and the second strand.

[0014] In other preferred aspects, control sequences are provided within one or more of the liking oligonucleotides or hairpin oligonucleotides described above, such as primer recognition sequences and the like.

[0015] In related aspects, the invention provides methods of determining a consensus nucleic acid sequence for a template nucleic acid segment where the sense and an antisense strand of the template nucleic acid segment are provided within a contiguous nucleic acid molecule. The contiguous nucleic acid molecule is contacted and/or complexed with a primer sequence complementary to at least a portion of the contiguous nucleic acid molecule and a polymerase enzyme to provide a nucleic acid synthesis complex, in the presence of a plurality of types of nucleotides or nucleotide analogs, wherein individual nucleic acid synthesis complexes are disposed upon a substrate such that individual complexes are optically resolvable. The sequence of nucleotides or nucleotide analogs incorporated into a nucleic acid synthesis reaction by the individual complexes in a template dependent manner is then monitored and/or detected to determine a nucleic acid sequence of the sense and antisense strands. The nucleic acid sequence of the sense and antisense strands are then used or compared to determine a consensus nucleic acid sequence of the template nucleic acid.

IDT Exhibit 1020.1156

[0016] Pooled sample methods are also envisioned by the present invention. For example, in some aspects, the methods of the invention comprise preparing template nucleic acid segments from each of a plurality of discrete nucleic acid samples, wherein the template nucleic acid segments comprise double stranded segments of the nucleic acid samples, a first strand of the double stranded segment being linked to a second strand of the double stranded segment by a linking oligonucleotide, wherein the linking oligonucleotide in each discrete nucleic acid sample comprises a unique, identifiable sequence characteristic. The template nucleic acid segments from the plurality of discrete nucleic acid samples are then pooled, and the pooled template nucleic acid segments are then sequenced to identify the identifiable sequence characteristic, and the nucleic acid sequences deriving from the discrete nucleic acid samples are identified based at least in part on the unique identifiable sequence characteristic identified in the sequencing step.

[0017] The invention is also directed to compositions useful in the foregoing methods. In one aspect, the invention provides compositions that comprise a template nucleic acid having a double stranded nucleic acid segment comprising a first strand segment and a second strand segment substantially complementary to the first strand segment, and at least a first linking oligonucleotide segment joining a 3' end of the first strand segment to the 5' end of the second strand segment. The compositions also typically include one or more of a primer sequence capable of hybridizing to at least a portion of the template nucleic acid and initiating polymerase mediated nucleic acid synthesis, and a polymerase enzyme.

[0018] The invention also provides template preparation kits useful in practicing the methods of the invention. Such kits typically comprise a first linking oligonucleotide, a primer sequence complementary to at least a portion of the first linking oligonucleotide, and one or more ligation reagents for coupling the first linking oligonucleotide to a 3' end of a first strand of a double stranded template and a 5' end of a second strand of the double stranded template nucleic acid. In addition, such kits will also typically include instruction protocols, and optionally reagents, for coupling the linking oligonucleotide(s) to double stranded target nucleic acid segments derived from samples for analysis.

[0019] The invention also includes systems for sequencing nucleic acid templates. The systems typically comprise a reaction mixture comprising a complex comprising a template nucleic acid, a polymerase enzyme and a primer sequence complementary to at least a portion of the template sequence, wherein the template sequence comprises a double stranded segment having a sense strand and an antisense strand, and at least a first linking oligonucleotide linking a 5' end of

IDT Exhibit 1020.1157

the sense strand to the 3' end of the antisense strand, at least a first nucleotide or nucleotide analog bearing a detectable labeling group, and a detection system configured to detect incorporation of the first nucleotide or nucleotide analog into a primer extension product by the polymerase.

[0020] Although described in significant detail for purposes of illustration, it will be appreciated that a variety of variations to the invention may be practiced within the scope of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] Figure 1A schematically illustrates a single molecule real time sequencing process. Figure 1B illustrates an exemplary short segment of a sequence read using the process schematically illustrated in Figure 1A.

[0022] Figure 2 illustrates two exemplary embodiments of template constructs used in the present invention.

[0023] Figure 3 schematically illustrates redundant or consensus sequencing using the constructs shown in Figure 2.

[0024] Figure 4 illustrates alternate template constructs and their application in sequencing processes.

[0025] Figure 5 schematically illustrates an overall consensus sequencing process using a template construct of the invention.

[0026] Figure 6 schematically illustrates an alternative approach to a sequencing process using template constructs of the invention.

[0027] Figure 7 schematically illustrates an exemplary assembly process of the template constructs used in the invention.

[0028] Figure 8 schematically illustrates an alternate assembly process for template constructs.

[0029] Figure 9 schematically illustrates another template preparation process of the invention.

[0030] Figure 10 schematically illustrates an alternate template preparation process according to the invention.

[0031] Figure 11 provides sequence readout from a template construct of the invention.

[0032] Figure 12 shows a schematic illustration of single nucleotide variant template constructs.

IDT Exhibit 1020.1158

[0033] Figure 13 shows a plot of the percentage of variant template in a reaction mixture vs. the percent of variants called experimentally.

[0034] Figure 14 shows a plot of sequence coverage over the entire E. coli genome, using the template constructs of the invention

[0035] Figure 15 shows a plot of the number of bases vs. depth of coverage from E.coli genomic sequencing both uncorrected (Panel A) and corrected (Panel B) for the reduced replication of E. coli genome away from the origin of replication.

[0036] Figure 16 shows a schematic illustration of a large repeat insert from E. coli, sequenced in a single contiguous template construct.

DETAILED DESCRIPTION OF THE INVENTION

I. General

[0037] The present invention is generally directed to improved methods, systems and compositions for carrying out nucleic acid sequence analysis, and particularly sequence analysis that employs template dependent synthesis in identifying the nucleotide sequence of target nucleic acids. Nucleic acid sequence analysis that employs template dependent synthesis identifies individual bases, or groups of bases as they are added during a template mediated synthesis reaction, such as a primer extension reaction, where the identity of the base is required to be complementary to the template sequence to which the primer sequence is hybridized during synthesis. Other such processes include ligation driven processes, where oligo- or polynucleotides are complexed with an underlying template sequence, in order to identify the sequence of nucleotides in that sequence. Typically, such processes are enzymatically mediated using nucleic acid polymerases, such as DNA polymerases, RNA polymerases, reverse transcriptases, and the like, or other enzymes such as in the case of ligation driven processes, e.g., ligases.

[0038] Sequence analysis using template dependent synthesis can include a number of different processes. For example, in the ubiquitously practiced four-color Sanger sequencing methods, a population of template molecules is used to create a population of complementary fragment sequences. Primer extension is carried out in the presence of the four naturally occurring nucleotides, and with a sub-population of dye labeled terminator nucleotides, e.g., dideoxyribonucleotides, where each type of terminator (ddATP, ddGTP, ddTTP, ddCTP) includes a different detectable label. As a result, a nested set of fragments is created where the fragments terminate at each nucleotide in the sequence beyond the primer, and are labeled in a manner that

IDT Exhibit 1020.1159

permits identification of the terminating nucleotide. The nested fragment population is then subjected to size based separation, e.g., using capillary electrophoresis, and the labels associated with each different sized fragment are identified to identify the terminating nucleotide. As a result, the sequence of labels moving past a detector in the separation system provides a direct readout of the sequence information of the synthesized fragments, and by complementarity, the underlying template (See, e.g., U.S. Patent No. 5,171,534, incorporated herein by reference in its entirety for all purposes).

[0039] Other examples of template dependent sequencing methods include sequence by synthesis processes, where individual nucleotides are identified iteratively, as they are added to the growing primer extension product.

[0040] Pyrosequencing is a sequence by synthesis process that identifies the incorporation of a nucleotide by assaying the resulting synthesis mixture for the presence of by-products of the sequencing reaction, namely pyrophosphate. In particular, a primer//template/polymerase complex is contacted with a single type of nucleotide. If that nucleotide is incorporated, the polymerization reaction cleaves the nucleoside triphosphate between the α and β phosphates of the triphosphate chain, releasing pyrophosphate. The presence of released pyrophosphate is then identified using a chemiluminescent enzyme reporter system that converts the pyrophosphate, with AMP, into ATP, then measures ATP using a luciferase enzyme to produce measurable light signals. Where light is detected, the base is incorporated, where no light is detected, the base is not incorporated. Following appropriate washing steps, the various bases are cyclically contacted with the complex to sequentially identify subsequent bases in the template sequence. See, e.g., U.S. Patent No. 6,210,891, incorporated herein by reference in its entirety for all purposes).

[0041] In related processes, the primer/template/polymerase complex is immobilized upon a substrate and the complex is contacted with labeled nucleotides. The immobilization of the complex may be through the primer sequence, the template sequence and/or the polymerase enzyme, and may be covalent or noncovalent. In general, preferred aspects, particularly in accordance with the invention provide for immobilization of the complex via a linkage between the polymerase or the primer and the substrate surface. A variety of types of linkages are useful for this attachment, including, e.g., provision of biotinylated surface components, using e.g., biotin-PEG-silane linkage chemistries, followed by biotinylation of the molecule to be immobilized, and subsequent linkage through, e.g., a streptavidin bridge. Other synthetic coupling chemistries, as well as non-specific protein adsorption can also be employed for immobilization. In alternate

IDT Exhibit 1020.1160

configurations, the nucleotides are provided with and without removable terminator groups. Upon incorporation, the label is coupled with the complex and is thus detectable. In the case of terminator bearing nucleotides, all four different nucleotides, bearing individually identifiable labels, are contacted with the complex. Incorporation of the labeled nucleotide arrests extension, by virtue of the presence of the terminator, and adds the label to the complex. The label and terminator are then removed from the incorporated nucleotide, and following appropriate washing steps, the process is repeated. In the case of non-terminated nucleotides, a single type of labeled nucleotide is added to the complex to determine whether it will be incorporated, as with pyrosequencing. Following removal of the label group on the nucleotide and appropriate washing steps, the various different nucleotides are cycled through the reaction mixture in the same process. See, e.g., U.S. Patent No. 6,833,246, incorporated herein by reference in its entirety for all purposes).

[0042] In yet a further sequence by synthesis process, the incorporation of differently labeled nucleotides is observed in real time as template dependent synthesis is carried out. In particular, an individual immobilized primer/template/polymerase complex is observed as fluorescently labeled nucleotides are incorporated, permitting real time identification of each added base as it is added. In this process, label groups are attached to a portion of the nucleotide that is cleaved during incorporation. For example, by attaching the label group to a portion of the phosphate chain removed during incorporation, i.e., a β, γ, or other terminal phosphate group on a nucleoside polyphosphate, the label is not incorporated into the nascent strand, and instead, natural DNA is produced. Observation of individual molecules typically involves the optical confinement of the complex within a very small illumination volume. By optically confining the complex, one creates a monitored region in which randomly diffusing nucleotides are present for a very short period of time, while incorporated nucleotides are retained within the observation volume for longer as they are being incorporated. This results in a characteristic signal associated with the incorporation event, which is also characterized by a signal profile that is characteristic of the base being added. In related aspects, interacting label components, such as fluorescent resonant energy transfer (FRET) dye pairs, are provided upon the polymerase or other portion of the complex and the incorporating nucleotide, such that the incorporation event puts the labeling components in interactive proximity, and a characteristic signal results, that is again, also characteristic of the base being incorporated (See, e.g., U.S. Patent Nos. 6,056,661, 6,917,726, 7,033,764, 7,052,847, 7,056,676, 7,170,050, 7,361,466, 7,416,844 and Published U.S. Patent Application No. 2007-

0134128, the full disclosures of which are hereby incorporated herein by reference in their entirety for all purposes).

[0043] One exemplary sequencing process based upon such Single Molecule Real Time (SMRT™) processes is schematically illustrated in Figure 1. As shown in Panel I of Figure 1A, a nucleic acid synthesis complex comprising a polymerase enzyme 102, a template sequence 104 and a primer sequence 106 complementary to a portion of the template sequence 104, is provided immobilized within a confined illumination volume (indicated by the dashed line 108), e.g., resulting from the evanescent optical field resulting from illumination of a zero mode waveguide 100, or in a total internal reflectance fluorescence microscope system or other optical confinement system, as described above.

[0044] The reaction mixture surrounding the complex contains the four different nucleotides (A, G, T and C) each labeled with a spectrally distinguishable fluorescent label attached through its terminal phosphate group. Because the illumination volume is so small, nucleotides and their associated fluorescent labels diffuse in and out of the illumination volume very quickly, and thus provide only very short fluorescent signals (shown as short pulses 110, in the underlying schematic data plot). When a particular nucleotide is incorporated by the polymerase in a primer extension reaction, the fluorescent label associated with the nucleotide is retained within the illumination volume for a longer time (shown as longer pulses 112). Once incorporated, the fluorescent label is cleaved from the base through the action of the polymerase, and the label diffuses away.

[0045] By identifying longer pulses of different spectral characteristics, one can detect, in real time, the identity of each incorporated base as it is being incorporated. An example plot of a short segment of a sequence read from such process is provided in Figure 1B, with the incorporated nucleotides identified based upon their spectral characteristics, and listed above their respective pulses. Shorter pulses not associated with incorporation tend to be so short that they are not detected by the camera, while pulses from incorporation provide more pronounced and detectable pulses, e.g., as shown in Figure 1B.

[0046] Although described in terms of the specific SMRT™ sequencing process, it will be appreciated that in accordance with the sequencing compositions of the invention, the nucleotides or nucleotide analogs may be detectable by any of a variety of different mechanisms including the presence of fluorescent dye labels coupled to the nucleotide through a β, γ or other more distal phosphate group. For example, as alluded to previously, the nucleotides may bear interacting components, such as one or both members of FRET pairs (dyes, semiconductor nanocrystals, or the

like) that interact with their complements elsewhere in the system e.g., on the polymerase, primer, the nucleotide itself, or the substrate. Similarly, these nucleotide analogs may bear other interactive components, such as energy donors or quenchers that alter signal capability of other proximal components. Likewise, non-optical labels may be employed, such as highly charged moieties, magnetic particles or the like, that may be detected by electrochemical systems, e.g., ChemFET sensors, nanopore sensors (see, e.g., Clarke et al., Nature Nanotechnology, Published online: 22 February 2009 | doi:10.1038/nnano.2009.12), and the like. In addition, the nucleoside polyphosphates described herein may generally include tri, tetra, penta, hexa or other phosphate chain lengths incorporatable by the polymerases used. Such compounds, including those bearing detectable labeling groups are described in, e.g., U.S. Patent No. 7,041,812, the full disclosure of which is incorporated herein by reference in its entirety for all purposes.

[0047] For a number of approaches, e.g., single molecule methods as described above, it may be desirable to provide the nucleic acid synthesis complexes in individually optically resolvable configurations, such that the synthesis reactions of a single complex can be monitored. Providing such complexes in individually resolvable configuration can be accomplished through a number of mechanisms. For example, by providing a dilute solution of complexes on a substrate surface suited for immobilization, one will be able to provide individually optically resolvable complexes (See, e.g., European Patent No.1105529 to Balasubramanian, et al., which is incorporated herein by reference in its entirety for all purposes. Alternatively, one may provide a low density activated surface to which complexes are coupled (See, e.g., Published International Patent Application No. WO 2007/041394, the full disclosure of which is incorporated herein by reference in its entirety for all purposes). Such individual complexes may be provided on planar substrates or otherwise incorporated into other structures, e.g., zero mode waveguides or waveguide arrays, to facilitate their observation.

II. Contiguous Double Stranded Templates

[0048] The present invention provides novel template configurations and methods for exploiting these compositions in template directed sequencing processes. While these compositions and methods have utility across all of the various template directed processes described herein, for ease of discussion, they are being primarily discussed in terms of preferred single molecule, real-time sequencing processes, in which they provide myriad benefits. In particular, the present invention is generally directed to nucleic acid sequences that employ improved template sequences

IDT Exhibit 1020.1163

to improve the accuracy of sequencing processes. For example, in at least one aspect, the template compositions of the invention are generally characterized by the presence of a double stranded segment or a pair of sub-segments that are internally complementary, i.e., complementary to each other. In particular contexts, the target nucleic acid segment that is included within a template construct will typically be substantially comprised of a double stranded segment, e.g., greater than 75%, or even greater than 90% of the target segment will be double stranded or otherwise internally complementary. For ease of discussion, these double stranded target segments, whether entirely complementary or predominantly complementary, e.g., having overhang regions, or other non-complementary portions such as secondary loop structures or the like, are referred to herein as complementary or substantially complementary. Where complete complementarity between two strands is intended and required from the context or explicitly, the phrase 'completely complementary' or 'entirely complementary' will be used.

[0049] The strands that make up the double stranded segment, and/or the internally complementary strands are, in the context of the invention, at least partially contiguous, and in preferred aspects are completely contiguous. As used herein, two strands are partially contiguous if they are joined at at least one end of each strand, and are completely contiguous if they are joined at both ends, resulting in an overall circular strand configuration, where such joining may be direct coupling of the ends of the sense and antisense strands, or through a linking oligonucleotide. As will be appreciated, the term circular, when referring to the strand configuration, merely denotes a strand of a nucleic acid that includes no terminal nucleotides, e.g., a loop, and does not necessarily denote any geometric configuration.

[0050] Examples of template configurations of the invention that are partially and completely contiguous are schematically illustrated in Figure 2A and 2B, respectively. In particular, as shown in Figure 2A, a partially contiguous template sequence 200 is shown which includes a double stranded portion, comprised of two complementary segments 202 and 204, which, for example, represent a target sequence or portion thereof. As shown, the 3' end of segment 202 is linked to the 5' end of segment 204 by linking oligonucleotide 206, providing a single stranded portion of the template, and yielding a partially contiguous sequence. By comparison, as shown in Figure 2B, a completely contiguous template sequence 210 is shown. Sequence 210 includes a double stranded portion again comprised of two complementary segments 212 and 214. As with the partially contiguous sequence of Figure 2A, the 3' end of segment 212 is joined to the 5' end of segment 214 via oligonucleotide 216 in a first single stranded portion. In addition, the 5' end of

IDT Exhibit 1020.1164

segment 212 is joined to the 3' end of segment 214 via linking oligonucleotide 218, providing a second single stranded portion, and yielding a completely contiguous or circular template sequence.

[0051] In a typical sequencing process, a double stranded target nucleic acid, e.g., a nucleic acid for which sequence information is desired, is converted to a template configuration of the invention. Typically, this involves fragmentation of a population of a larger double stranded nucleic acid, e.g., a genome, plasmid, or portion thereof, or the like, into pools of smaller, overlapping double stranded target nucleic acids or nucleic acid fragments. The double stranded fragments are then processed to provide the linking oligonucleotides described above and in greater detail, below. A number of methods, as set forth herein, may be employed to provide the linking oligonucleotides connecting the sense and antisense strand. In one simple example, exogenous hairpin adapter sequences are used and simply ligated to the end(s) of the double stranded fragments to form such linking oligonucleotides. Such exogenous hairpin adapters will typically include two complementary nuclei acid sequence segments at opposing ends of a longer sequence, where they are separated by another stretch of non-complementary nucleotides. The result is a structure that includes a double stranded stem structure coupled to the single stranded loop. Examples of such hairpin structures have been described previously (See, e.g., U.S. Patent Nos. 6,498,023, 6,451,563, and 7,368,265, the full disclosures of which are incorporated herein by reference in their entirety for all purposes). Generally, such structures are readily synthesized using conventional nucleic acid synthesis techniques. A number of other methods employ alternate approaches to generating the linking oligonucleotide structures, and are described below.

[0052] Following sequence determination of the various target nucleic acid fragments, the sequence of the starting target nucleic acid is then determined by identifying and aligning overlap between sequence data from the various different overlapping fragments.

[0053] The templates of the invention provide numerous advantages over simple linear template sequences, and even other circular template sequences (See, e.g., U.S. Patent No. 7,302,146 for a discussion of circular templates for sequencing applications, the full disclosure of which is incorporated herein by reference in its entirety for all purposes). In particular, as with circular templates, the template configurations of the invention allow for single molecular consensus sequence determination, where sequencing a given template provides duplicative or replicate data of the sequence information obtained, and thereby improves accuracy over linear templates by providing multiple reads for a given template sequence or sequence portion, that can be used to derive consensus sequence data from a given template sequence and/or for specific base

IDT Exhibit 1020.1165

locations within such sequence. In these templates, the potential for consensus sequence determination is provided, in one respect, by virtue of the circular nature of the overall template structure, for a completely contiguous template, allowing repeated processing of the same molecule to obtain consensus base calls and/or a consensus sequence. In addition or alternatively, the templates of the invention, by virtue of their inclusion of double stranded segments, provide consensus sequence determination through the sequencing of both the sense and antisense strand of such sequences (in both the partially and completely contiguous configurations). Although described herein primarily as consensus sequence determination, it will be appreciated that the level of consensus determination extends to the individual base level for a given base position within a sequence, which, when placed in sequence context yields consensus sequence data for a particular sequence. Thus, the invention also encompasses consensus base calling (nucleotide identification) for individual positions within a template sequence.

[0054] By way of example, with respect to a partially contiguous template shown in Figure 2A, obtaining the entire sequence, e.g., that of segments 202, 204 and 206 provides a measure of consensus sequence determination by virtue of having sequenced both the sense strand, e.g., segment 202, and the antisense strand, e.g., segment 204. In addition to providing sense and antisense sequence reads from a single template molecule that can be sequenced in one integrated process, the presence of linking segment 206 also provides an opportunity to provide a registration sequence that permits the identification of when one segment, e.g., 202, is completed and the other begins, e.g., 204. Such registration sequences provide a basis for alignment sequence data from multiple sequence reads from the same template sequences, e.g., the same molecule, or identical molecules in a template population. The progress of sequencing processes is schematically illustrated in Figure 3A. In particular, as shown, a sequencing process that begins, e.g., is primed, at the open end of the partially contiguous template, proceeds along the first or sense strand, providing the nucleotide sequence (A) of that strand, as represented in the schematic sequence readout provided. The process then proceeds around the linking oligonucleotide of the template, providing the nucleotide sequence (B) of that segment. The process then continues along the antisense strand to the A sequence, and provides the nucleotide sequence (A'), which sequence can be used to derive or determine a consensus sequence for the sense strand, as its antisense counterpart. As noted, because the B sequence may be exogenously provided, and thus known, it may also provide a registration sequence indicating a point in the sequence determination at which the sequencing

IDT Exhibit 1020.1166

reaction, and thus, the sequence data being obtained from the overall template construct, transitions from the sense to the antisense strands.

[0055] With respect to completely contiguous or circular template sequences configured in accordance with the invention, the potential for obtaining replicate sequence read data from which one may assemble consensus sequence information, is further increased. In particular, as with the partially contiguous sequences shown in Figure 2A, the completely contiguous sequences also provide sense and antisense sequence data. In addition, such templates provide for the potential for iterative sequencing of the same molecule multiple times, by virtue of the circular configuration of the template. Restated, a sequence process may progress around the completely contiguous sequence repeatedly obtaining sequence data for each segment from the complementary sequences, as well as sequence data within each segment, by repeatedly sequencing that segment. All or portions of such sequence data is then useful in deriving a consensus sequence for the template and its various segments. This is schematically illustrated in Figure 3B, again with a representative illustration of a sequence readout provided. As shown, a sequencing process that is primed at one end, e.g., primed within one linking oligonucleotide sequence, e.g., linking oligonucleotide 218 of Figure 2, proceeds along the first or sense strand 214, again providing the nucleotide sequence A of that strand. The sequence process then proceeds around the first linking oligonucleotide, e.g., linking oligonucleotide 216 from Figure 2, to provide the nucleotide sequence B of that segment of the template. Proceeding along the antisense strand, e.g., segment 212 of Figure 2B, provides the nucleotide sequence A', which is again, complementary to sequence A. The sequencing process then continues around the template providing the nucleotide sequence for the other linking oligonucleotide, e.g., linking oligonucleotide 218 of Figure 2B, where the illustrated sequencing process began, providing nucleotide sequence C. Because the template is circular, this process can continue to provided multiple repeated sequence reads from the one template, e.g., shown as providing a second round of the sequence data (A-B-A'-C-A-B-A'). Thus, sequence redundancy comes from both the determination of complementary sequences A and A', and the repeated sequencing of each segment.

[0056] As will be appreciated, in iteratively sequencing circular templates, strand displacing polymerases, as discussed elsewhere herein, are particularly preferred, as they will displace the nascent strand with each cycle around the template, allowing continuous sequencing. Other approaches will similarly allow such iterative sequencing including, e.g., use of an enzyme having 5'-3' exonuclease activity in the reaction mixture to digest the nascent strand post synthesis.

IDT Exhibit 1020.1167

[0057] As will be appreciated, and as noted elsewhere herein, the repeated sequencing of a given sequence of nucleotides, whether as the identical sequence or as a complement to that sequence, and whether as an individual sequence or as a repeating copy of that sequence, e.g., in a concatameric orientation, provides the ability to assemble a consensus sequence for that sequence segment. In particular, one can employ iterative sequence reads of a given sequence, and/or reads of a sequence and its complement, to establish a consensus sequence of nucleotides for that sequence segment. Rephrased, each base called in a sequence can be based upon a consensus base call for that particular position based upon multiple reads at that position and/or reads at that position and it's complement in an antisense strand. These multiple reads are then assembled or compared to provide a consensus determination of a given base at a given position, and as a result, a consensus sequence for the particular sequence segment.

[0058] Although referred to herein as comparing or assembling the sequence data from multiple reads of a given sequence, and/or from the sense and antisense strands of the sequence, it will be appreciated that any method of assigning a consensus determination to a particular base call from multiple reads of that position of sequence, and/or to a provide an overall consensus sequence for that segment, will be envisioned and encompassed by the term "compare". Such methods include actual side-by-side comparisons, scoring methods whereby calls of iterative reads or from complementary sequences are scored by the number of occurrences, and optionally or alternatively additional signal/base calling metrics, to complex algorithms for assigning a base call from multiple indications of the base at a given position.

[0059] By way of example, some such methods generally encompass algorithms that combine data derived from a base with data derived from other determinations of that sequence position, whether in a duplicate sequence read (from the same or a duplicate sequence segment), or from its complement in the opposite strand.

[0060] One category for combination of such data includes a first step of determining the association between signals originating from a base and signals originating from other reads at that position or from its complement, followed by a second step of combining those data. Methods for determining such association are known in the art and include heuristic methods for multiple-sequence alignment, optimal methods for multiple sequence alignment, and hidden Markov models, and in the broadest sense all of these algorithms have in common that they seek to align corresponding bases in the input reads. Such associations are not limited to basecalls alone and can be derived from inherent features of the signal such as channel, amplitude, width, or other time-

IDT Exhibit 1020.1168

dependent patterns associated with the signals. Algorithms known in the art for the combination of data from pre-associated bases include Plurality, Quality-weighted Plurality, Bayesian methods, and machine learning approaches (neural networks, self-organizing maps), and generally these algorithms have in common that they classify the presented evidence into a consensus basecall, often with an associated probability of error. In this category of algorithm, there are numerous variations that include additional steps before, after and between these steps that have beneficial impacts such as improvement the quality of results reduction in computation time. For example, in one such method data associated with other bases in said first step are referenced during or after said second step in order to evaluate the possibility that errors were made in said association. The association step can be made between reads from a single molecule, or between reads from different molecules.

[0061] In another method for combination of data, a consensus is first established for one sequence read, e.g., a sense strand, and another consensus sequence is established for another read, e.g., the antisense strand. These sequences can be established by methods known in the art including heuristic methods for multiple-sequence alignment, optimal methods for multiple sequence alignment, or hidden Markov models in combination with a consensus determination algorithm such as Plurality, Quality-weighted Plurality, Bayesian methods, or machine learning approaches (neural networks, self-organizing maps). These consensus sequences can then be associated and combined via algorithms such as heuristic methods for multiple-sequence alignment, optimal methods for multiple sequence alignment, or hidden Markov models in combination with a consensus determination algorithm such as Plurality, Quality-weighted Plurality, Bayesian methods, or machine learning approaches (neural networks, self-organizing maps).

[0062] Another class of algorithms performs the association and combination in a single step. These methods proceed from a probabilistic model which seeks to find the most likely template given the observed reads. Such models belong in the category of probabilistic graphical models and include Bayesian networks, hidden Markov models, neural networks, and conditional random fields. The input to such a model is not limited to basecalls alone, but can include local or global measures of the individual sequence qualities, sequence context, and characteristic features of the raw signals. Consensus determination by these models proceeds by finding the template which maximizes the likelihood of the observed reads under the constraints of the probabilistic model. It is usual that such methods can produce one or more likely consensus sequences and provide scoring information for ranking the likelihood of each sequence. For purposes of discussion

herein, the methods by which a consensus sequence is called from repeated sequence reads of a given segment (including reads of the sense and antisense strands), are referred to generally herein as "comparing."

[0063] In addition to the aspects related to consensus sequence determination, set forth above, there are a number of other significant advantages of the template structures of the invention. For example, with respect to completely contiguous constructs, a number of advantages of these templates derive from their basic circular structure. For example, because the completely contiguous templates of the invention are circular in structure, one can initiate a polymerase mediated synthesis process and sequencing at any point in the template, with the expectation that the entire template will be sequenced or at least sequenceable. In contrast, with linear template sequences, one faces the risk of missing upstream portions of the template for sequencing. As a result, sequence initiation processes, also referred to as "hot start" processes, are typically used when employing linear templates, e.g., withholding a key reagent until one is ready to initiate synthesis and sequence data collection.

[0064] In addition to obviating the need for sequencing initiation processes, the use of such circular templates also provides the ability to obtain sequence data from disparate portions of the same template molecule without necessarily requiring that one obtain the entire sequence of the template. For example, in single molecule sequencing processes, or in sequencing processes that employ populations of identical template molecules, one can obtain multiple subsets of sequence information on the template molecule with the understanding that each subset is contextually related to each other subset, e.g., contained within the same template sequence. This aspect of the invention is discussed in greater detail below with reference to paired end sequencing processes.

[0065] For example, because of the structure of these templates, one can effectively make a circular template out of any length of target nucleic acid. In particular, because one can use exogenous linking sequences, which may include their own self complementary segments, one can effectively circularize very short double stranded oligonucleotide fragments. In particular, circularization of a single stranded nucleic acid segment typically requires a sequence that is at least 30 nucleotides in length, with probable additions of adapter sequences and the like, in order to circularize the segment. The templates of the invention, on the other hand, can be used to convert a target nucleic acid of much shorter length, into a circular template. In particular, target sequences equal to or shorter than 50 bases, 20 bases, or even 10 bases, can readily be converted to a circular template. In addition, the converse is also true in exploiting the template constructs of the

IDT Exhibit 1020.1170

invention. In particular, using these configurations, one can effectively provide completely contiguous or "circular" templates from very large double stranded target sequences, without many of the difficulties attendant to the production of larger circular nucleic acids. In particular, one can effectively employ very large double stranded inserts, e.g., greater than 100 base duplexes (total nucleotides), greater than 500 base duplexes, greater than 1000 base duplexes, greater than 5000 base duplexes, and even greater than 10,000 base duplexes in producing a template construct.

[0066] The template configurations of the invention are also advantageous in that they do not require working with single stranded nucleic acid sample segments. For example, in cases where one is presented with target nucleic acids that are double stranded, e.g., PCR products, genomic fragments, or the like, one can readily convert these to template molecules of the invention, by providing linking segments that couple the 3' end of one strand to the 5' end of the other strand, and vice versa. As noted above, relatively small double stranded segments, e.g., having 50 or fewer base pairs, 20 or fewer base pairs, or even 10 or fewer base pairs, may be readily incorporated into circular templates in accordance with the invention. Typically, circularization of double stranded segments, because of their structural rigidity, requires segments on the order of several hundred base pairs in length. However, the methods of the present invention can provide circular templates from smaller segments.

[0067] The template configurations of the invention, also provide advantages by virtue of the presence of both single stranded and double stranded portions within the same template. By way of example, the single stranded linking oligonucleotides, not only can be selected to present a specific primer recognition sequence, they also provide the primer and polymerase binding site in a required single stranded conformation, thus allowing primer annealing and polymerase complexing, without the need for any denaturing steps. While advantages over other double stranded templates are readily apparent, priming advantages also exist with respect to other single stranded templates. In particular, in cases where one is employing a polymerase that binds single stranded segments with high affinity, one can better control polymerase binding when only a relatively small portion of the template is single stranded, e.g., directing binding to the desired priming sites. A variety of examples of such polymerases are available, including, e.g., phi29 type polymerases (See, e.g., U.S. Patent Nos. 5,001,050, 5,576,204, the full disclosures of which are incorporated herein by reference in their entirety for all purposes) and derivatives thereof, as well as other strand displacing polymerases (See, e.g., International Patent Application Nos. WO 2007/075987, WO 2007/075873, WO 2007/076057, incorporated herein by reference in their entirety for all purposes).

IDT Exhibit 1020.1171

[0068] Notwithstanding the foregoing, however, while the single stranded portion of the template is often and advantageously employed as the priming site, in some cases, priming can also be carried out within the double stranded portions of the template, e.g., allowing the use of primers specific for the target sequence.

[0069] In addition, they provide an element of controllability in presenting primer binding sites that other single stranded templates may lack. In particular, selection of primer sequences for large target nucleic acids can be difficult, as one typically wants to avoid random and multiple priming of the template material. Accordingly, primer sequences typically must be carefully selected to be highly specific, requiring relatively long sequences. However, by providing only a minimal amount of single stranded material within a template population, e.g., those portions within the linking oligonucleotides, and further enhanced by the knowledge of those sequences, one can utilize substantially shorter primers while retaining specificity to the desired priming location, as the larger target fragments, as double stranded nucleic acids, will be unavailable for primer binding. Further, one can utilize primers that are engineered to provide substantially higher affinity for the template, without the concurrent concern that such higher affinity to the template will yield a higher rate of random or non-specific priming. In particular, one can select tighter binding primer sequences, e.g., GC rich sequences, as well as employ primers that include within their structure non-natural nucleotides or nucleotide analogs, e.g., peptide nucleic acids (PNAs) or locked nucleic acids (LNAs), that can demonstrate higher affinity pairing with the template.

[0070] In addition to the foregoing advantages, partially contiguous templates of the invention provide a number of specific advantages, as well. By way of example, in some cases, a partially contiguous template may be provided in which, in addition to the linking oligonucleotide segment, a portion of the target segment may exist as a single stranded segment. Depending upon which strand is provided as a single stranded segment, e.g., the 5' or 3' end of the template, different alternatives for sequencing, and particularly, priming, are presented. For example, in the configuration where the single stranded portion is that which includes the 3' terminus, the single stranded portion provides a primer binding location in order to sequence the remainder of the template. In contrast, by providing the 5' end of the template as the single stranded portion, the template itself provides the primer for polymerase mediated synthesis and thus, sequencing. In other configurations, a gap or nick (as described elsewhere herein) may be provided within a double stranded segment of an otherwise completely contiguous template structure, to provide a priming site for polymerase attachment.

IDT Exhibit 1020.1172

[0071] These configurations and the alternative synthesis/sequencing approaches are schematically illustrated in Figure 4A, 4B and 4C. As shown in Figure 4A, a partially contiguous template 400 is provided that includes a double stranded portion 402, a single stranded linking oligonucleotide portion 404. A portion of the target sequence is also provided in single stranded form, e.g., segment 406. Figure 4B illustrates the configuration where the single stranded portion is the 3' portion of the template. In particular, one can use the single stranded segment 406 as a binding location for a primer for synthesis and sequencing, e.g., primer 408. Sequencing then proceeds along the remainder of the template, as indicated by the arrow.

[0072] In contrast, as indicated in Figure 4C, the 5' portion of the template is provided as the single stranded segment 406. As such, the double stranded portion of the template 402, functions as a self primer to allow synthesis and sequencing of the single stranded portion 406, as indicated by the arrow. Additional components could be added to the template structure, e.g., to protect the 3' end from exonuclease digestion, e.g., a phosphorothioate, etc.

[0073] In addition to all of the foregoing advantages, and relative to the sense/antisense consensus sequence determination aspects of the template configurations of the invention, the templates described herein, may be applied in sequencing processes that utilize fewer than four differentially detectable events for incorporation of the four different nucleotides. In particular, many sequencing processes employ the four different nucleotides each labeled with a distinct and separately detectable label group, e.g., a fluorescent label. In the past, there have been proposals for performing sequencing operations using fewer than four distinct labeling groups. By way of example, a proposed process employed nucleotides labeled with only two distinct labels, where one label is associated with a particular nucleotide, while the other is associated with all other nucleotides. The label on three of the nucleotides is merely provided to provide relative context or spacing between each of the specifically labeled nucleotides. Complete sequence information, however, would require passing through the template at least three, and most likely four times, to provide a specific label for each base.

[0074] In the context of the templates of the invention, because of the presence of both the sense and antisense sequences within the template construct, a single pass through a given template could provide complete sequence information using fewer than four distinct labeled events. In particular, by providing two, non-complementary nucleotides, e.g., A and G, with the same label, and providing a second and third label on the T and C nucleotides, one could obtain the complete sequence by comparing the sense and antisense strands to identify which of the identically labeled

IDT Exhibit 1020.1173

bases would be an A and which would be a G, by virtue of its complementary to a specifically identified T or C from the other strand. Similarly, one could effectively obtain complete target sequence information from sense and antisense strands using only two distinguishable labels. In particular, one could associate a first label with one set of the four nucleotides, e.g., A and G, while providing a second label on the other set, T and C for sequencing the sense strand. The labeling configuration could then be switched for sequencing the antisense strand, e.g., the first label linked to the A and C bases and the second label linked to the G and T bases. Comparison of the sequence data would then yield the precise identity of each base. In operation, sequencing could be carried out in the presence of the first labeling configuration and then new reagents would be introduced bearing the second labeling configuration, e.g., in a wash step. As will be appreciated, repeated sequencing would, again, provide for the ability to obtain consensus sequence data, as described above.

III. Additional Sequences

[0075] In addition to advantages of consensus potential within each template molecule, and the other advantages described above, the template configurations have a number of different advantages for many or all of the different template dependent sequencing processes associated with the potential for the addition of other sequences into the template molecule.

[0076] For example, in some cases, connecting or linking sequences may be selected and/or observed as registration sequences to provide landmarks within the overall template sequence, e.g., to provide alignment of iterative sequence data, to identify the level of coverage in a consensus sequence read, to identify points in a sequencing process where one is progressing into a consensus sequence, e.g., an antisense strand or repeated sequence of the entire template, and the like.

[0077] In addition, such sequences may provide control opportunities for the sequencing process using such templates. For example, and preferably in the case of completely contiguous sequences, as discussed previously, one may incorporate primer recognition sequences within the connecting oligonucleotides to initiate polymerization. As noted previously, the flexibility as to the types and configuration of the primer sequences is increased by virtue of immunity from binding to the target portion of the sequence, which exists as a double stranded segment.

[0078] Additional control sequences may also be provided, e.g., sequences that allow control over the initiation of synthesis, e.g., through a hybridized probe or reversibly modified nucleotide, or the like (See, e.g., U.S. Patent Application No. 2008-0009007, the full disclosure of

IDT Exhibit 1020.1174

which is incorporated herein by reference in its entirety for all purposes.). Other control sequences may include binding sites for transcription factors. For example, repressor binding regions may be provided as control sequences within the linking oligonucleotides, such as the lac repressor recognition sequence, which when bound by the lac repressor protein, has been shown to block replication both in vivo and in vitro. Reinitiation of replication is accomplished through the addition of appropriate initiators, such as isophenylthiogalactoside (IPTG) or allolactose. Other DNA binding protein recognition sites may also be included within the linking oligonucleotide to allow control over the progress of synthesis using the templates of the invention. Other controllable elements may include the use of non-natural bases (also termed 5th bases) within the linking region which are not paired with any of the four basic nucleoside polyphosphates in the synthesis reaction. Upon encountering such a base, the polymerase would pause until its own particular complement was added to the reaction mixture. Likewise, an engineered pause point within the linking oliogonucleotide region could include a "damaged" base that causes a stop in replication until repair enzymes are added to the mixture. For example within the linking oligonucleotide could be included a base position having a pyrimidine dimer. Such compounds would cause the replication complex to pause. Addition of the photolyase DNA repair enzyme would repair the problem location and allow replication, and sequencing to continue.

[0079] Recognition sites for a variety of other oligonucleotide probes are also optionally incorporated into these linking sequences, e.g., hybridization sites for labeled probes, molecular beacons, TaqMan® probes, Invader® probes (Third Wave Technologies, Inc.), or the like, that can be used to provide other indications of the commencement of synthesis. Additionally, non-native bases that interact/complement other non-native bases may be used to provide an initiation point for synthesis and sequencing.

[0080] In some cases, it may be desirable to provide endonuclease recognition sites within the linking oligonucleotide, which can allow for a mechanism to release a given template sequence from a synthesis reaction, i.e., by linearizing it, and allowing the polymerase to run off the linear template, and/or to expose the template to exonuclease activity, and thus terminate synthesis through removal of the template. Such sites could additionally be exploited as control sequences by providing specific binding locations for endonucleases engineered to lack cleavage activity, but retain sequence specific binding.

[0081] In some cases, nicking sites, e.g., sites recognized by nicking endonucleases, may be included within a portion of the template molecule, and particularly within the double stranded

portion of the template, e.g., in the double stranded fragment portion or in the stem portion of an exogenous hairpin structure. Such nicking sites provide a break in one strand of a double stranded sequence, to present a priming location for, e.g., a strand displacing polymerase enzyme. In the context of the templates of the invention, the nicking site may be provided for example, within a hairpin adapter that is annealed and ligated to a double stranded target fragment. Other methods as described below may similarly introduce nicking sites. Alternatively, nicking endonucleases may be applied randomly against the target fragment to initiate priming. A variety of nicking enzymes and their recognition sequences are known in the art, with such enzymes being generally commercially available, e.g., from New England Biolabs. Alternatively, one may employ pre-nicked double stranded segments in the hairpin adapters used in preparing the template construct. Such nicks could include gaps in the double stranded segments of from 0 to 20 nucleotides, depending upon the need of the application.

IV. Template Structures

[0082] In addition to the basic configuration of the templates of the invention, as described and illustrated previously, a number of other structural considerations may also be incorporated into the templates of the invention. For example, the relative size or length of the overall template molecule, as well as the sizes of the various segments that make up the template, may be selected for optimal benefit in given applications.

[0083] In general, the overall size of the template will be dictated by the application in which the template will be used. By way of example, where a given template is being subjected to a polymerase mediated sequencing process, limitations on the readlength for the particular system may be factored into the selection of the overall template size, e.g., to ensure complete, and preferably redundant sequencing of the entire template. For example, where a given polymerase mediated sequencing process has readlength of 1000 bases, a requirement for at least 2X redundant sequencing would dictate a template of 500 bases, including both the linking oligonucleotides and the target segment. Of course, because the sequence of the start/finish linking oligonucleotide may be known and is not relevant to determination of the target sequence, it may not be necessary to obtain 2X redundancy of that segment, and thus a consequent increase in template size could be tolerated. For purposes of certain redundant sequencing applications, a template that is between about 50 and about 500 bases may be desired. In other applications, where longer readlengths are obtained, or in non-redundant applications, templates that are from about 200 to about 50,000 bases

IDT Exhibit 1020.1176

in length may be used. Although described in terms of specific lengths, it will be appreciated that a variety of different template sizes may be employed for a variety of different specific applications.

[0084] In addition to readlength considerations, an overall template may be subject to application-specific structural requirements. For example, where a sequencing process employs nanostructured reaction regions, it may be desirable to provide smaller template molecules to ensure rapid diffusion into and out of the reaction region.

[0085] The size of the target portion may also be varied depending upon the application in which the template is being used. For example, in genomic sequencing applications, e.g., de novo or resequencing processes, longer target segments will be desired in order to reduce the level of duplicate coverage that is required among different fragments. In particular, the ability to sequence template fragments that are in excess of 100, preferably in excess of 200, still more preferably, in excess of 500, in excess of 1000, and even in excess of 10,000 nucleotides in length, provides substantial benefits in genomic assembly from overlapping fragments. In particular, the level of required duplicate coverage for identical sequence portions is substantially reduced by increases in the size of any individual sequence read.

[0086] In addition to advantages for long read length sequencing applications, larger target segments also provide advantages in the ability to provide paired end sequence data using single molecule sequencing processes. Briefly, in many sequencing processes, one can obtain sequence context of relatively short sequence reads, by reading the sequence that is disposed at opposing ends of a large target fragment. Typically, this involves the sequencing of a relatively short stretch of bases at either end of a double stranded target segment. From the knowledge that these two sequences are derived from the same target molecule, and optionally, from a general understanding of the size of the fragment, one obtains contextual data for the short sequences. While paired end sequencing has distinct advantages in short readlength sequence processes in providing two pieces of sequence information from a given target, it also is useful in longer read sequence technologies as it provides the ability to obtain contextual "waypoints" for very large nucleic acid sequences, which can be used in aligning sequence data.

[0087] In the context of the template sequences of the invention, one can readily obtain sequence data from opposing ends of a single template by first obtaining sequence data from a first end of the target portion. One may then wait an appropriate amount of time for a given sequencing system, for the process to reach the opposing end of the target, and begin obtaining sequence data again. As a result, one has obtained sequence data from paired ends of the same target. As will be

IDT Exhibit 1020.1177

appreciated, the foregoing process has particular use where an overall readlength of a sequencing system is impacted by the data collection process, e.g., through the continuous illumination of the complex (See, e.g., U.S. Patent Application No. 2007-0161017, the full disclosure of which is incorporated herein by reference in its entirety for all purposes). Alternatively, one may employ a reaction stop point within the template sequence, such as a reversibly bound blocking group at one location on the template, e.g., on the single stranded portion that was not used in priming. By way of example, and with reference to figure 2B, following initial sequencing from the original priming location, e.g., at single stranded linking oligonucleotide portion 216, through one end of the sense strand 214, the data acquisition may be switched off, allowing the polymerase to proceed around the template, e.g., through sense strand 214, to the other previously single stranded portion, e.g., linking oligonucleotide portion 218. The incorporation of a synthesis blocking group coupled to the linking oligonucleotide will allow control of initiation of the sequencing of the opposing end of the antisense strand, e.g., strand 212. One would thereby obtain paired end sequence data for the overall double stranded segment. A variety of synthesis controlling groups may be employed, including, e.g., large photolabile groups coupled to the nucleobase portion of one or more bases in the single stranded portion, which inhibit polymerase mediated replication, strand binding moieties that prevent processive synthesis, non-native nucleotides included within the primer (as described in greater detail elsewhere herein), and the like.

[0088] Alternatively, one may employ primer recognition sites on each of the two linking oligonucleotide sequences employed in a population of like template molecules, e.g., PCR products. By then separately sequencing from each end, one can obtain sequence data from different ends of the same double stranded fragment, and thus obtain the paired end data desired.

[0089] In contrast, for diagnostic sequencing applications, it may be necessary only to provide sequence data for a small fragment of DNA, but do so in an extremely accurate sequencing process. For such applications, shorter target segments may be employed, thus permitting a higher level of redundancy by sequencing multiple times around a small circular template, where such redundancy provides the desired accuracy. Thus, in some cases, the double stranded target segment may be much shorter, e.g., from 10 to 200, from 20 to 100 or from 20 to 50 or from 20 to 75 bases in length. For purposes of the foregoing, the length of the target segment in terms of bases denotes the length of one strand of the double stranded segment.

[0090] While different applications will have different impacts on the length of the target sequence portion that is included in the template molecule, the length and structure of the linking

IDT Exhibit 1020.1178

oligonucleotide or single stranded portions of the template may be dictated, at least in part, by structural considerations in addition to application specific criteria. In particular, at a minimum, the linking oligonucleotides are required to be able to form a connecting loop between the 3' end of one strand of a double stranded nucleic acid segment and the 5' end of the other strand. As such, where employed primarily as a linking oligonucleotides, e.g., without accommodating larger functional elements, the linking oligonucleotide typically will be from about 4 nucleotides to about 100 nucleotides or more, while linking oligonucleotides of from 4 nucleotides to about 20 nucleotides will be generally preferred. For example, where short linkages are desired, linking oligoucleotides may be from 4 to about 8 nucleotides in length.

[0091] In addition to the foregoing structural requirements, where a given linking oligonucleotide portion provides a primer and/or polymerase binding site, that segment must be of sufficient length to accommodate the desired primer length, as well as a complexed polymerase. Accordingly, linking oligonucleotides that include primer recognition sites will typically be greater than about 20 bases in length, and preferably at least about 36 bases in length. In some cases, it may be desirable to provide sufficient space on one or both sides of the primer within the single stranded portion, e.g., to accommodate polymerase binding, etc. As such, in some cases, the single stranded portion will be substantially greater than as set forth above, e.g., 50 bases, 80 bases, 100 bases or larger.

[0092] Notwithstanding the foregoing, in some cases, shorter linking oligonucleotides may be desirable, as templates with smaller hairpin loops show increased efficiency as templates in that less of the overall template construct, and thus, less of the sequencing capability of the system, is taken by the "overhead" of the linking oligonucleotides. Accordingly, linking oligonucleotides in some cases will be smaller than 20 bases in length, preferably smaller than 12 bases in length. As will be appreciated, where one desires to provide optimal primer binding, but enhanced efficiency, the linking oligonucleotides will generally be in the range of from about 20 to about 100 bases in length, preferably, from about 20 to about 80 bases in length. In addition, asymmetric linking oligonucleotides, e.g., having different numbers of nucleotides joining the sense and antisense strands, may be used within a single template construct. Such constructs could be generated through, e.g., iterative processes of cleavage of a sample segment with a first type of restriction endonuclase, followed by annealing/ligation of a first adapter/linking hairpin sequence that is complementary to the cleavage site/overhang sequence, followed by treatment with a second

IDT Exhibit 1020.1179

restriction endonuclease, followed by annealing/ligation with a second differently sized hairpin adapter, complementary to the second cleavage site/overhang.

V. Strand Displacement

[0093] As noted previously, the complementary segments of the template may be provided in double stranded form, e.g., as shown in Figure 2B. As will be appreciated, in such cases, it will be preferable to affect strand separation either prior to or during the template dependent sequencing process. In the case of, for example, sequence by incorporation processes, strand separation is preferably carried out through the selection and use of a strand displacing polymerase enzyme. A variety of strand displacing polymerase enzymes are readily available, including, for example, Φ29 polymerase and Φ29 type polymerases (See, e.g., U.S. Patent Nos. 5,001,050, 5,576,204, the full disclosures of which are incorporated herein by reference in their entirety for all purposes), Bst Polymerase (available from New England Biolabs), as well as those polymerases described in commonly owned International Patent Application Nos. WO 2007/075987, WO 2007/075873, WO 2007/076057 the full disclosures of which are incorporated herein by reference in their entirety for all purposes.

[0094] The synthesis process for such templates and strand displacing enzymes is schematically illustrated in Figure 5. As shown, a completely contiguous template 500 is complexed with a primer sequence 502 and a strand displacing polymerase 504, and contacted with the four nucleotides 506, or in the case of certain preferred aspects, fluorescently labeled nucleotide analogs. As synthesis progresses, the polymerase's own activity displaces one complementary strand 508 from the other 510 and synthesis of the nascent strand 512 continues. Upon complete synthesis, e.g., one full cycle around the template, a double stranded circular sequence results, made up of the original template 500 and the newly synthesized or nascent strand 512. Because the strand displacing enzyme can continue to displace the hybridized strand, e.g., the newly synthesized nascent strand 512, the synthesis, and by implication, the sequencing process can continue through the template multiple times to provide multiple sequences for use in consensus sequence determination, typically generating a long, concatamer molecule containing repeated regions complementary to the contiguous template 500.

[0095] Alternatively, other mechanisms may be employed to affect strand separation prior to or during synthesis. For example, elevation of the temperature of the reaction mixture may be used to melt the double stranded portion of the template, and permit primer extension through that

IDT Exhibit 1020.1180

region. As will be appreciated, for such applications, it may be desirable to employ thermally stable polymerase enzymes that are better suited to the temperatures required for melting, and continued synthesis. A wide variety of thermostable polymerases are known in the art and are applicable to this type of application, including, for example Taq polymerase and its variants.

[0096] A schematic of synthesis using a thermally regulated initiation process is illustrated in Figure 6. As shown, a primer 602 is tethered to the template structure 600 and contacted with the non-strand displacing polymerase enzyme 604. Because the template exists in a double stranded configuration and the polymerase is unable to displace the complementary strand, the synthesis does not readily proceed. At a desired point, the double stranded segment is separated to allow synthesis of the nascent strand 606 through the previously double stranded portion of the template 600, e.g., through heating sufficient to melt the double stranded segment without removing the primer (indicated as ΔC). As will be appreciated, primer sequences, as well as additional portions of the linking oligo-nucleotides, may be employed that have relatively higher melting temperatures, e.g., GC rich sequences, that have higher melting temperatures than an average naturally occurring nucleic acid sequence. Once the double stranded segment is duplicated sufficient to prevent re-hybridization of the original template, by virtue of the presence of the nascent strand, there is no longer a need for denaturation steps or additives.

[0097] As will also be appreciated, in the case of the use of non-strand displacing enzymes, additional strand separation steps will typically be needed following one complete cycle around the template, as the nascent strand would then be in position to block continued synthesis. As with initiation of the primer extension, the requirement for another triggering event can provide advantages of synchronizing different template sequence steps. Alternatively, following an initial triggering event, the synthesis reaction may be maintained at elevated temperatures to ensure continuous, uninterrupted synthesis and sequencing.

VI. Sequence Alignment

[0098] As alluded to above, and as an additional advantage, the template configurations of the invention have an inherent alignment potential for consensus sequence determination of the same or identical templates. In particular, because the connecting oligonucleotide is known or knowable, one can readily exploit this pre-known sequence in aligning long strings of sequence data from such template or templates, e.g., as a landmark or registration sequence. Further, even without knowledge of the sequence of the connecting oligonucleotide, one can derive the sequence by

IDT Exhibit 1020.1181

observing that portion of the overall sequence that does not possess a complementary portion elsewhere in the sequence data. In particular, the target sequence portion of the template, as a double stranded segment, will have, within the same sequence data, an internal complement, i.e., will have both a sense and an antisense strand. However, depending upon the length of the target segment and the connecting oligonucleotide, the probability of an exact complement to the connecting oligonucleotide existing within the target segment will be low to zero. As such, one can scan sequence data derived from a given template construct for a portion that has no internal complement elsewhere within the sequence, and assume this is the connecting oligonucleotide, and consequently, exploit it as an alignment marker.

[0099] One exemplary algorithm for accomplishing this would proceed as follows. First the entire sequence would be subjected to Smith-Waterman alignment with its own reverse complement. An alignment quality threshold would be applied so that the sequence would be annotated according to those regions that have an alignment with the reverse complement, and those that do not. The repeat unit of the sequence would be identified, e.g., using Fourier transform methods known in the art, or by application of the same Smith-Waterman algorithm of the sequence with itself rather than the reverse complement. The annotations from the first step are collapsed into a single repeat unit and summed. Then a more certain identification of the sequence insert can be made using the statistics of all of the repeats. For example, in 10 repeats, regions that show 1 or fewer hits against the reverse complement can be called "marker sequence" and regions that have 2 or more hits can be declared "genomic sequence". The precise thresholds can be determined by the needs of the application.

[00100] As alluded to above, however, the connecting or linking oligonucleotide may be selected or generated to include sequence with identifiable sequence characteristics to facilitate its identification, both with respect to the contiguous template sequence with which it is associated, and also with respect to other template sequences that may exist in the same sample mixture. In particular, one can utilize connecting oligonucleotides in the template constructs that include sequence markers like a barcode, to indicate the origin of a given template sample. Different template samples including distinguishable connecting oligonucleotide tags could then be pooled for analysis in a single sequencing process. Sequence data derived from individual templates is then attributable to the originating sample preparation process through the identification of the connecting oligonucleotide tag.

IDT Exhibit 1020.1182

[00101] In particular, a number of discrete sample preparation processes are performed for different nucleic acid containing samples. These discrete sample preparation processes may be performed on different starting materials, e.g., different samples (cells, cell cultures, patients, etc.), different portions of the same original material, i.e., different size-selected portions or the like, or different portions of the same population, cell culture, etc. The template deriving from each discrete process is bar-coded through the use of a unique, identifiable and discrete linking nucleotide sequence in the template construct.

[00102] The different samples are then pooled for sequencing in a unified sequencing reaction. As the sequencing output from each run is based upon individual molecules, the sequencing data is then parsed out according to its origin by virtue of the barcode sequence that is integrated into the template molecules. In particular, because each sequence read is derived from a single molecule, the sequence of the insert portion of the template can be unambiguously linked to the sequence of its attached adapter sequence that includes the bar code sequence within that adapter sequence. Accordingly, each template sequence is then traced back to its origin, e.g., a particular sample, patient, etc.

VII. Contiguous Template Preparation

[00103] The template structures of the invention may be prepared in a number of different ways. In a first exemplary process, double stranded target fragments are coupled at one or both ends with separate hairpin loops to provide the template structures of the invention. Such methods provide simplified template preparation processes that reduce undesired concatamerization of fragments, and permit facile cleanup of random nucleic acid fragments from the preparation. For example, completely complementary double stranded nucleic acid fragments may be joined via blunt end ligation with hairpin adapter sequences. In such cases, given the reduced ability to control which adapter ligates to the different ends of the double stranded fragments, it may be desirable to employ a single type of adapter. Such completely complementary double stranded segments may be generated using blunt end cutting enzymes, or through the use of restriction enzymes that yield overhangs, followed by filling in of the overhanging single strands, e.g., using Klenow fragment or the like.

[00104] In other methods, described herein, ligation processes may be employed that controllably ligate a given adapter type to a given end of the double stranded fragment, thus

IDT Exhibit 1020.1183

allowing the use of identifiable sequences within the linking oligonucleotides, to facilitate identification of one end of the template from the other.

[00105] One of such methods is schematically illustrated in Figure 7. As shown, a double-stranded nucleic acid fragment or fragments, e.g., double stranded fragment 700, is provided. The double stranded fragment may be derived from the fragmentation of larger target nucleic acids, e.g., genomic DNA, cDNA, DNA concatamers, and/or amplification products, e.g., from PCR or LCR amplification, or the like. Hairpin adapters 710 are then appended to each end of the double stranded fragment 700. As shown, the attachment of the hairpin adapters 710 relies upon the presence of unique overhang sequences 720 at the 3' end of each strand of the double stranded fragment 700. Complementary overhang sequences 722 are provided on the hairpin adapters, to provide for specific annealing and ligation of the hairpin adapters 710 to the double stranded fragments 700. As shown, the overhang sequences are the product of A-tailing of the double stranded fragment which appends a series of adenosine nucleotides to the 3' end of each strand. A complementary set of thymidines at the 3' end of each of the hairpin adapters provides for specific annealing. However, a number of different specific overhang sequences may be provided at the ends of the double stranded fragments. For example, restriction endonucleoases may be used to fragment a larger segment of double stranded DNA leaving a characteristic overhang sequence at every cleavage point. The complementary sequence to this characteristic sequence is then provided upon the hairpin adapter to provide specific annealing and ligation. Alternatively, specific overhang sequences may be separately ligated to the 3' or 5' ends of each strand for use as the overhang sequence. In addition to providing specificity for hairpin annealing, the overhang sequences also serve to prevent concatamerization of the fragments prior to annealing and ligation of the hairpin adapters. Following annealing, the hairpin adapters are ligated to the fragments, using standard ligation processes.

[00106] As noted above, while less preferred for its lack of additional specificity and other advantages, blunt end ligation also may be employed in ligating hairpin adapters to the ends of the double stranded fragments. In such cases, concatamerization of the template fragments or other non-specific associations may be avoided through the use of excess amounts of hairpin adapters. Alternatively, emulsion based reactions, where individual droplets within the emulsion are sized substantially to provide individual molecules, may provide protections against concatamerization.

[00107] In an alterative process, a template sequence may be formed using an alternate ligation process to form the template configuration provided herein. In some cases, this alternate

IDT Exhibit 1020.1184