[0145] While the Digital Array microfluidic devices are well-suited for carrying out the digital amplification methods described herein, one of ordinary skill in the art would recognize many variations and alternatives to these devices. The microfluidic device which is the 12.765 Dynamic Array commercially available from Fluidigm Corp. (South San Francisco, CA), includes 12 panels, each having 765 reaction chambers with a volume of 6 nL per reaction chamber. However, this geometry is not required for the digital amplification methods described herein. The geometry of a given Digital Array microfluidic device will depend on the particular application. Additional description related to devices suitable for use in the methods described herein is provided in U.S. Patent Application Publication No. 2005/0252773, incorporated herein by reference for its disclosure of Digital Array microfluidic devices.

[0146] In certain embodiments, the methods described herein can be performed using a microfluidic device that provides for recovery of reaction products. Such devices are described in detail in copending U.S. Application No. 61/166,105, filed April 2, 2009, which is hereby incorporated by reference in its entirety and specifically for its description of microfluidic devices that permit reaction product recovery and related methods. For example, the digital PCR method for calibrating DNA samples prior to sequencing can be preformed on such devices, permitting recovery of amplification products, which can then serve as templates for DNA sequencing.

[0147] Fig. 2 is a simplified perspective illustration of a carrier and a microfluidic device according to an embodiment of the present invention. As illustrated in Fig. 2, the carrier 100 supports a microfluidic device 110, which may also be referred to as a Digital Array microfluidic device. The carrier 100 may be made from materials providing suitable mechanical support for the various elements of the carrier. As an example, the carrier is made using a plastic or other suitable material. The outer portion of the carrier has the same footprint as a standard 384-well microplate and enables stand-alone valve operation. Additionally, the carrier 100 is compatible with conventional stand-alone thermal cyclers. As described below, there are 48 sample input ports 120 located on a first side of the carrier 100 and 48 assay input ports 122 located on an opposing side of the carrier. The banks of sample input ports 120 and assay input ports 122 are recessed with respect to the top

IDT Exhibit 1020.2670

of the carrier. Utilizing these recessed features, pressure can be applied concurrently to all of the sample input ports or the assay input ports, driving fluids present in the respective ports through fluid lines 140 connecting the input ports and either vias, fluid input lines, or combinations thereof, present on the microfluidic device 110. The samples may include encoded primers and the assays may also be referred to as amplicon-specific (AS) primers.

[0148] The carrier 100 also includes four sources 130, 132, 134, and 136, which may be used to actuate control lines present in the microfluidic device. In an embodiment, sources 130, 132, and 134 are used to pressurize control lines operable to open and close valves present in the microfluidic device. For example, application of pressure greater than atmospheric pressure to source 132 will result in the liquid present in source 132 flowing into control lines present on the microfluidic device, thereby actuating valves operable to obstruct flow through one or more fluid input lines also present on the microfluidic device. In an embodiment, source 130 is used as a fluid well containing harvesting reagent. Pressure can be applied to source 130, forcing the harvesting reagent to flow through fluid lines provided on the carrier to fluid lines provided on the microfluidic device. Thus, application of pressure to source 130 can result in the flow of a harvesting reagent or other suitable fluid through the microfluidic device. The control lines that are in fluid communication with the sources 130-136 can include control lines for interface valves, containment valves, valves used in dilation pumping, fluid lines for the flow of harvesting reagent, or the like. In a particular embodiment, valve 1 is controlled by source 132, valve 2 is controlled by source 134, harvesting reagent is provided in source 130, and hydration reagent is provided in source 136. In this particular embodiment, the interface valves are controlled by source 150 and containment valves are controlled by source 152. This particular embodiment is not intended to limit the present invention, but merely to provide an example of one configuration. Other configurations can be utilized as appropriate to the particular application.

[0149] As described more fully in relation to Fig. 3, fluid lines 140 present on the carrier 100 are in fluid communication with one or more fluid lines present on the microfluidic device 110. These fluid lines can serve to carry fluids into and out of the microfluidic device or may be used as control lines to actuate valves present on the microfluidic device. Thus, fluids provided in sample input ports 120 or assay input

IDT Exhibit 1020.2671

ports 122 can be loaded into appropriate fluid input lines and chambers of the microfluidic device. Other fluids (e.g., liquids) provided in sources 130-136 can also be loaded into either fluid input lines or control lines of the microfluidic device. Reaction products from the chambers of the microfluidic device can be recovered as they are pumped through fluid lines on the microfluidic device, back into the fluid lines 140 present on the carrier and into the sample or assay input ports 120 or 122.

[0150] Pressure accumulators 150 and 152 may be utilized to pressurize other control lines, provide for hydration of the microfluidic device, or they may not be used in some embodiments. Although 48 sample input ports and 48 assay input ports are shown in the embodiment of the present invention illustrated in Fig. 2, this is not required by the present invention. Other embodiments utilize a different number of samples and assays depending on the particular application. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0151] Fig. 3 is a simplified schematic diagram of a microfluidic device according to an embodiment of the present invention. The microfluidic device 200 includes vias 210 connected to fluid input lines 212 that are used to provide fluid flow paths for 24 different samples. The 24 samples, which can be loaded into a subset of the sample input ports 120 illustrated in Fig. 1, flow through vias 210 and fluid input lines 212 to sample input lines 312 and eventually to sample chambers 310 as illustrated in Fig. 4. The microfluidic device 200 also includes vias 220 connected to fluid input lines 222 that are used to provide fluid flow paths for 21 different assays. The via 250 on the side of the microfluidic device opposing the assay input fluid lines provides for hydration, actuation of a control line, or other suitable operations. The array portion 230 of the microfluidic device is illustrated (in part) in Fig. 4. In the array portion 230, the samples and assays are loaded into sample and assay chambers and then can be mixed to form pairwise combinations.

[0152] As described more fully throughout the present specification, after samples and assays are mixed and reacted, the reaction products can be recovered from the microfluidic device by flowing a recovery fluid through the fluid input lines 212, through the array portion 230 of the microfluidic device as illustrated in Fig. 4, and through the output fluid lines 240 and vias 242. These output fluid lines are in fluid communication with output ports provided on a carrier. Thus, reaction products

IDT Exhibit 1020.2672

pooled from the combination of a sample with each of the various assays are separately provided through each of the independent output fluid lines 240.

[0153] The particular number of sample and assay input lines illustrated in Fig. 3 are provided merely by way of example and particular implementations are not limited to these particular numbers. In other embodiments, additional sample and assay input lines are provided in order to facilitate additional pairwise combinations of samples and assays in the microfluidic device.

[0154] Fig. 4 is a simplified schematic diagram of several unit cells of the microfluidic device illustrated in Fig. 3. In Fig. 4, four unit cells from the array portion 230 are illustrated for purposes of clarity. The sample input lines 316 are in fluid communication with the fluid input lines 212 and the assay input lines 318 are in fluid communication with the assay input lines 222 as illustrated in Fig. 3. The unit cell section of the microfluidic device includes sample chambers 310 and assay chambers 312. Fluid lines 314 provide for fluid communication between the sample chambers and the assay chambers when the interface valves 330 are in the open position. In a specific embodiment, sample input lines 316 are provided in a layer of the microfluidic device underlying the layer containing the sample chambers. In a similar manner, assay input lines 318 are provided in a layer of the microfluidic device underlying the layer containing the assay chambers. Samples flow from the sample input lines 212 illustrated in Fig. 3 to sample input lines 316 and up through one or more vias (not shown) passing from the sample input lines 316 to the sample chambers 310. Although the sample input lines 316 are illustrated as branching into three input lines as the fluid passes to the sample chambers, this particular number is not required by the present invention and other numbers of sample input lines, for example, 1 input line, 2, 4, or more than 4 sample input lines may be utilized. Similar design criteria are applicable to the three fluid lines 314 connecting the sample chambers and corresponding assay chambers. The sample input lines 316 provide a continuous flow path in the row direction of the figure, enabling a single sample to be distributed evenly among multiple sample chambers, for example, the top row of sample chambers or the bottom row of sample chambers.

[0155] Utilizing interface valves 330 and containment valves 340, each of the sample chambers can be isolated from each of the other sample chambers as well as the assay chambers. The assay chambers can be isolated from the other assay

IDT Exhibit 1020.2673

chambers using the containment valves. Both the isolation and containment valves are actuated by application of pressure to a corresponding control line present on the carrier or by other means, for example, electrostatic actuation.

[0156] Fig. 4 illustrates assay input lines 318, which provide for assay flow from assay input lines 222 illustrated in Fig. 3 to the assay chambers 312. When the containment valves are in the open position, assays are able to flow from the assay input lines to the assay chambers 312. In a specific embodiment, the assays flow through vias connecting the input lines and the assay chambers in a manner similar to the filling of the sample chambers. The loading of assays along the columns of the microfluidic device provide a different assay for each row of samples, resulting in M x N pairwise combinations.

[0157] Opening of the interface valves 330 enables the samples and the assays to mix in pairwise combinations via free interface diffusion. After the samples and assays are mixed, thermocycling can be performed to form reaction products. Reaction products are recovered from the microfluidic device by opening harvest valves 350, which enable the reaction products to flow into portions 360 of the sample input lines adjacent the sample chambers. Using sample input lines 316 and on-chip pumps (not shown), reaction products flow through the sample input lines toward recovery ports on the carrier.

[0158] In the embodiment illustrated in Fig. 4, samples load from a first side of the microfluidic device. The assays load from an adjacent side of the microfluidic device. After processing, a harvesting reagent is input from the first side of the device using the sample input lines and reaction products exit the microfluidic device out fluid lines running toward the side of the microfluidic device opposing the first side. In this embodiment, the remaining side of the microfluidic device is not used for sample or assay loading or reaction product unloading. Other configurations are included within the scope of the present invention and the example configuration illustrated in Fig. 4 is merely provided by way of example. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0159] A benefit provided by the systems described herein is that the volume of samples and assays used in the reactions are fixed, regardless of the pipetting volume dispensed into the sample input and the assay input ports. If the volume in

IDT Exhibit 1020.2674

the sample and/or assay input ports is above a predetermined threshold sufficient to fill the sample/assay input lines and the sample/assay chambers, then application of pressure to the sample/assay input ports will result in complete filling of the sample/assay chambers. The completely filled chambers thus provide a fixed reaction volume not available with conventional microtiter plate techniques.

[0160] Although systems have been developed by the present assignee to perform many simultaneous binding assays, including, but not limited to immunological experiments such as ELISA assays, embodiments of the present invention provide for dilation pumping "on-chip" as well as separate sample and assay chambers. Thus, pairwise combinations of samples and assays are possible using embodiments described herein that are not possible with previously developed techniques. Additional description of binding assays is provided in U.S. Patent Application Publication No. 2007/0074972, filed on September 13, 2006, the disclosure of which is hereby incorporated by reference in its entirety for all purposes.

[0161] Embodiments of the present invention provide a system suitable for PCR sample preparation that features reduced cost, time, and labor in the preparation of amplicon libraries from an input DNA template. In a typical use case, the first amplification will be used to generate libraries for next-generation sequencing. Utilizing embodiments of the present invention, samples and encoded primers are combined with amplicon-specific (AS) primers to create a mixture that is suitable for desired reactions. Based on an MxN architecture of the microfluidic device, each of the M samples is combined with each of the N AS primers (i.e., assays) to form MxN pairwise combinations. That is, one reaction site is provided for each sample and assay pair. After the completion of the reaction (e.g., PCR), the reaction products are recovered from the system, typically using a harvest reagent that flows through the microfluidic device. In a specific embodiment, reaction products associated with each sample are recovered in a separate reaction pool, enabling further processing or study of the pool containing a given sample reacted with each of the various assays.

[0162] Thus, in embodiments described herein, a microfluidic device is provided in which independent sample inputs are combined with primer inputs in an MxN array configuration. Thus, each reaction is a unique combination of a particular sample and a particular primer. As described more fully throughout the present specification, samples are loaded into sample chambers in the microfluidic device

IDT Exhibit 1020.2675

through sample input lines arranged as columns in one implementation. AS primers or assays are loaded into assay chambers in the microfluidic device through assay input lines arranged as rows crossing the columns. The sample chambers and the assay chambers are in fluidic isolation during loading. After the loading process is completed, an interface valve operable to obstruct a fluid line passing between pairs of sample and assay chambers is opened to enable free interface diffusion of the pairwise combinations of samples and assays. Precise mixture of the samples and assays enables reactions to occur between the various pairwise combinations, producing a reaction product including a set of specific PCR reactions for which each sample has been effectively coded with a unique barcode. The reaction products are harvested and can then be used for subsequent sequencing processes. The terms "assay" and "sample" as used herein are descriptive of particular uses of the devices in some embodiments. However, the uses of the devices are not limited to the use of "sample(s)" and "assay(s)" in all embodiments. For example, in other embodiments, "sample(s)" may refer to "a first reagent" or a plurality of "first reagents" and "assay(s)" may refer to "a second reagent" or a plurality of "second reagents." The MxN character of the devices enable the combination of any set of first reagents to be combined with any set of second reagents.

[0163] According to one particular process implemented using an embodiment of the present invention, after 25 cycles of PCR, the reaction products from the MxN pairwise combinations will be recovered from the microfluidic device in discrete pools, one for each of the M samples. Typically, the discrete pools are contained in a sample input port provided on the carrier. In some processes, the reaction products may be harvested on a "per amplicon" basis for purposes of normalization. Utilizing embodiments of the present invention, it is possible to achieve results (for replicate experiments assembled from the same input solutions of samples and assays) for which the copy number of amplification products varies by no more than ± 25% within a sample and no more than ± 25% between samples. Thus, the amplification products recovered from the microfluidic device will be representative of the input samples as measured by the distribution of specific known genotypes. Preferably, output sample concentration will be greater than 2,000 copies/amplicon/microliter and recovery of reaction products will be performed in less than two hours.

IDT Exhibit 1020.2676

[0164] Applications in which embodiments of the present invention can be used include sequencer-ready amplicon preparation and long-range PCR amplicon library production. For the sequencer-ready amplicon preparation, multiple-forward primer and 3-primer combination protocols can be utilized.

[0165] Fig. 5A is simplified schematic diagram of a microfluidic device according to another embodiment of the present invention. The microfluidic device illustrated in Fig.5A shares common features as well as differences with the microfluidic device illustrated in Fig. 2. Samples are loaded into the microfluidic device through 48 vias and corresponding sample input lines provided at one edge of the microfluidic device (i.e., the bottom edge in Fig. 5A). Samples flow through the sample input lines into the array portion 430 of the microfluidic device. The assays are loaded from vias and assay input lines on two sides of the microfluidic device (i.e., the left and right sides in Fig. 5A). Additional discussion of the unit cells present in the array portion 430 is provided in relation to Fig. 6. Reaction products are removed through the sample input lines and are recovered in the sample input ports 120 provided on the carrier 100. Thus, in Fig. 5A, loading of samples and recovery of reaction products are illustrated as flowing through the bottom side of the microfluidic device.

[0166] Fig. 5B is a simplified schematic diagram of portions of the microfluidic device illustrated in Fig. 5A. The portions illustrated in Fig. 5B include sample input lines 410, assay input lines for even assays (assay input lines 420), assay input lines for odd assays (assay input lines 422), and a harvesting reagent input lines 430. In an embodiment, the sample input lines 410 are in fluid communication with vias 412 that are aligned with sample input lines 140, which are in fluid communication with sample input ports 120 as illustrated in Fig. 2. In other embodiments, additional sample input lines (not shown) are provided to enable fluid communication between the sample input ports 120 and the sample input lines 410. Thus, pressurization of the bank of sample input ports will result in flow of the various samples into the illustrated sample input lines 410.

[0167] As discussed in relation to Fig. 6 below, sample input lines 410 are in fluid communication with sample input lines 516 and sample chambers 510 present in the microfluidic device. For an array with 48 sample chambers and 48 assay chambers, the 48 sample input lines 410 illustrated in Fig. 5B will provide samples to

IDT Exhibit 1020.2677

48 separate columns of sample chambers, two of which are illustrated in Fig. 6. It should be noted that the various fluid lines illustrated in Fig. 5B can be integrated in carrier 100, integrated in the microfluidic device 110, or present in one or more other structures, depending on the particular implementation. Thus, the illustration of the sample input lines 410 in Fig. 5B is not intended to limit the scope of the present invention but merely to illustrate fluid lines suitable for providing controlled fluid flow to the various chambers of the microfluidic device.

[0168] In an embodiment, even assay input lines 420 and odd assay input lines 422 are in fluid communication with vias 424 that are aligned with assay input lines 140, which are in fluid communication with assay input ports 122 as illustrated in Fig. 2. In other embodiments, additional assay input lines (not shown) are provided to enable fluid communication between the assay input ports 122 and the assay input lines 420 and 422. Thus, pressurization of the bank of assay input ports 122 will result in flow of the various assays into the illustrated assay input lines. After flowing through the input lines illustrated in Fig. 5B, the various fluids will eventually enter into the unit cells illustrated in Fig. 6.

[0169] As discussed above, the various fluid lines can be integrated into the carrier, the microfluidic device, or other suitable structure. In a 48 sample x 48 assay array configuration, the 24 even assay input lines 420 will provide inputs to half of the rows of assay input lines 518 shown in Fig. 6. The 24 odd assay input lines 422 will provide inputs to half of the rows of assay inputs lines 518 shown in Fig. 6. Thus, although loading of assays is only illustrated from the right side of the array in Fig. 6, actual implementation will typically load assays from both sides in an even/odd configuration. In some embodiments, additional vias are provide for loading of hydration fluids or the like. Moreover, in some embodiments, in order to provide compatibility with existing carriers, some fluid lines are unused or modified to provide for such compatibility.

[0170] The harvesting reagent input line 430 provides for harvesting reagent used in recovering reaction products from the microfluidic device. The harvesting reagent input line 430 illustrated in Fig. 5B is in fluid communication with the harvesting reagent input port 136 illustrated in Fig. 2 and passes along the microfluidic device adjacent to the even assay input lines to the top portion of the device and then branches off into a plurality of harvesting reagent input lines. The

IDT Exhibit 1020.2678

harvesting reagent multiplexor has a substantially equal volume for every sample input line to provide uniform pumping during the reaction product recovery operation. It should be noted that the particular branching system illustrated in Fig. 5B is merely provided as an example and other branching systems are included within the scope of the present invention. The harvesting reagent input lines 430 are in fluid communication with sample input lines 516 discussed in relation to Fig. 6. As discussed in relation to Figs. 9A-D, embodiments utilize a separate harvesting reagent input line for each column of sample chambers, for example, 48 harvesting reagent input lines for an microfluidic device with a 48 x 48 array configuration. Additionally, although the harvesting reagent input line enters the microfluidic device at a location adjacent the even assay input lines 420, this is not required by embodiments of the present invention and other configurations are within the scope of the present invention. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0171] As described in relation to Fig. 8C, harvesting reagent flows from a harvesting reagent input port on the carrier, through the harvesting reagent input lines 430 and into one end of the sample input lines 516. As discussed more fully throughout the present specification, the sample input lines function both as input lines and reaction product recovery lines. For loading, the sample flow path is from the sample input ports 120, through input lines 140, through vias 412, through sample input lines 410, through sample input lines 516, to the reaction chambers 510 and to the loading bowls 830. For reaction product recovery, the product flow path is from the harvesting reagent input port 136,through the harvesting reagent input lines 430, through the sample input lines 516, through the sample input lines 410, through the vias 412, and to the sample input ports 120, which serve during harvesting, as a fluid recovery well. Thus, the use of the term "input" lines should be considered in the context of the particular process being performed, since the "input" lines can serve to both load samples and recover or remove reaction products from the microfluidic device and the carrier.

[0172] By applying pressure to the bank of sample input ports 120 and the bank of assay input ports 122, samples and reagents can be loaded through the illustrated sample and assay input lines into sample and assay chambers present in the microfluidic device. By applying pressure to the harvesting reagent input port 136,

IDT Exhibit 1020.2679

the reaction products can be recovered from the sample chambers and delivered to the sample input ports. Valves present in the microfluidic device are utilized to control the flow of samples, assays, and reaction products, as described more fully throughout the present specification. Fig. 5B only illustrates a portion of the sample input lines, assay input lines, and harvesting reagent input lines and additional portions of these input lines are illustrated in Fig. 5A and Fig. 6.

[0173] Fig. 6 is a simplified schematic diagram of several unit cells of the microfluidic device illustrated in Fig. 5A. In Fig. 6, four unit cells from the array portion 430 shown in Fig. 5A are illustrated for purposes of clarity. The unit cell section of the microfluidic device includes sample chambers 510 and assay chambers 512. Fluid lines 514 provide for fluid communication between the sample chambers and the assay chambers when the interface valves 530 are in the open position. In a specific embodiment, sample input lines 516 are provided in a layer of the microfluidic device underlying the layer containing the sample chambers. In a similar manner, assay input lines 518 are provided in a layer of the microfluidic device underlying the layer containing the assay chambers. Samples flow from the sample input lines 410 illustrated in Fig. 5B to sample input lines 516 and up through one or more vias passing from the sample input lines to the sample chambers. Although two sample lines are illustrated for each sample chambers, this particular number is not required by the present invention and other numbers of sample input lines, for example, 1 input line, 3, 4, or more than 4 sample input lines may be utilized. The sample input lines 516 provide a continuous flow path in the column direction of the figure, enabling a single sample to be distributed evenly among multiple sample chambers.

[0174] Fig. 6 illustrates assay input lines 518, which provide for assay flow from assay input lines 420 and 422 illustrated in Fig. 5B to the assay chambers 512. Although Fig. 6 only illustrates assay input lines entering the unit cells from the right side of the figure, it will be evident to one of skill in the art that in the implementation illustrated in Fig. 5B, even and odd assays are loaded from opposing sides of the microfluidic device. The illustration provided in Fig. 6 is merely simplified for purposes of clarity. In a specific embodiment, the assays load through vias connecting the input lines and the assay chambers in a manner similar to the filling of the sample chambers. The loading of assays along the rows of the microfluidic device

IDT Exhibit 1020.2680

provide a different assay for each of the samples, resulting in a number of pairwise combinations appropriate for an M x N array.

[0175] As described more fully throughout the present specification, reaction products are recovered from the microfluidic device using the sample input lines 516 and pumps (not shown). Containment valves 540 provide for containment between the various sample and assay chambers in each row. Utilizing the interface valves 530 and the containment valves 540, each of the sample chambers can be isolated from each of the other sample chambers as well as the assay chambers. The assay chambers can be isolated from the other assay chambers using the containment valves. Both the isolation and containment valves are actuated by application of pressure to a corresponding control line in fluid communication with sources 130-134 or by other means, for example, electrostatic actuation.

[0176] In Fig. 6, four sample chambers 510 are illustrated in an array configuration. The four illustrated chambers are merely shown by way of example and implementations of the present invention are not limited to the four illustrated chambers, but typically provide 2,304 chambers in a 48 x 48 array configuration, 4,096 chambers in a 64 x 64 array configuration, 9,216 chambers in a 96 x 96 array configuration, or the like.

[0177] Embodiments of the present invention provide unit cells with dimensions on the order of several hundred microns, for example unit cells with dimension of 500 x 500 μm, 525 x 525 μm, 550 x 550 μm, 575 x 575 μm, 600 x 600 μm, 625 x 625 μm, 650 x 650 μm, 675 x 675, μm, 700 x 700 μm, or the like. The dimensions of the sample chambers and the assay chambers are selected to provide amounts of materials sufficient for desired processes while reducing sample and assay usage. As examples, sample chambers can have dimensions on the order of 100-400 μm in width x 200-600 μm in length x 100-500 μm in height. For example, the width can be 100 μm, 125 μm, 150 μm, 175 μm, 200 μm, 225 μm, 250 μm, 275 μm, 300 μm, 325 μm, 350 μm, 375 μm, 400 μm, or the like. For example, the length can be 200 μm, 225 μm, 250 μm, 275 μm, 300 μm, 325 μm, 350 μm, 375 μm, 400 μm, 425 μm, 450 μm, 475 μm, 500 μm, 525 μm, 550 μm, 575 μm, 600 μm, or the like. For example, the height can be 100 μm, 125 μm, 150 μm, 175 μm, 200 μm, 225 μm, 250 μm, 275 μm, 300 μm, 325 μm, 350 μm, 375 μm, 400 μm, 425 μm, 450 μm, 475 μm, 500 μm, 525 μm, 550 μm, 575 μm, 600 μm, or the like. Assay chambers can have

IDT Exhibit 1020.2681

similar dimensional ranges, typically providing similar steps sizes over smaller ranges than the smaller chamber volumes. In some embodiments, the ratio of the sample chamber volume to the assay chamber volume is about 5:1, 10:1, 15:1, 20:1, 25:1, or 30:1. Smaller chamber volumes than the listed ranges are included within the scope of the invention and are readily fabricated using microfluidic device fabrication techniques.

[0178] Higher density microfluidic devices will typically utilize smaller chamber volumes in order to reduce the footprint of the unit cells. In applications for which very small sample sizes are available, reduced chamber volumes will facilitate testing of such small samples.

[0179] The dimensions of the interface valves 530 are selected to provide for complete obstruction of the fluid lines 514 connecting the sample and assay chambers. In some embodiments, the valve dimensions range from about 10-200 μm x 10-200 μm, for example, 50 x 50 μm, 50 x 65 μm, 50 x 80 μm, 50 x 100 μm, 65 x 50 μm, 65 x 65 μm, 65 x 80 μm, 65 x 100 μm, 80 x 50 μm, 80 x 65 μm, 80 x 80 μm, 80 x 100 μm, 100 x 50 μm, 100 x 65 μm, 100 x 80 μm, 100 x 100 μm, or the like. The sample input lines may have various widths depending on the number of sample input lines and the sample chamber volumes, and desired flow rates for loading and product recovery. As examples, the sample input lines may have a cross-section of 1-20 μm in height and 50-100 μm in width. For example, the sample input lines may have heights of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 μm and widths of 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 μm.

[0180] Other device parameters, including layer to layer alignment, ranging from 20-100 μm, and via size, ranging from 50-200 microns, are selected to provide desired system performance characteristics. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0181] In some embodiments, an extra assay inlet is provided at the side of the microfluidic device adjacent the harvesting reagent input lines. Additionally, no assay inlet is provided at the side of the microfluidic device adjacent the sample input ports on the carrier. In this configuration, the extra assay inlet can be used for dehydration chamber loading. Typically, loading of the dehydration chambers will

IDT Exhibit 1020.2682

use more than 5 μl of assay solution. Alternatively, a separate dehydration solution could be used to keep assay volumes uniform across the microfluidic device.

[0182] Fig. 7 is a simplified flowchart of a method of operating a microfluidic device according to an embodiment of the present invention. In the illustrated embodiment, the microfluidic device includes at least one assay chamber, at least one sample chamber, and at least one harvesting port. In a particular embodiment, the microfluidic device includes a plurality of assay chambers and a plurality of sample chambers. The method 600 includes closing a fluid line between the assay chamber and the sample chamber (610). Referring to Fig. 6, interface valves 530 are operable to close fluid lines 514 passing between the sample chambers 510 to the assay chambers 512. In some embodiments, the interface valves 530 are "push-up" valves as described more fully below. The interface valves are formed by the intersection of control line 532 or control channel that is at least partially contained in a first layer of the microfluidic device. The fluid lines are at least partially contained in a second layer of the microfluidic device. The control lines 532 are in fluid communication with one or more pressure actuators or accumulators as illustrated in Fig. 2.

[0183] The intersection of the control line 532 with the fluid line 514 forms a valve at the intersection, referred to as an interface valve 530 because the valve prevents mixing at the interface between the sample and the assay. The interface valve 530 is actuated in response to fluid pressure in the control line and is operative to prevent fluid flow through the fluid lines. Generally, the multilayer microfluidic device discussed herein includes a number of elastomeric layers and the valves 530 include a deflectable membrane between the first layer and the second layer. In a "push-up" configuration, the deflectable membrane of the valve is deflectable into the fluid line 514 positioned above the intersection with the control line 532. In this configuration, the deflectable membrane deflects up into the fluid line to close the fluid line at the position of the valve, thus the reference to "push-up" valves. Releasing the pressure in the control line will result in the deflectable membrane returning to the undeflected position and thereby opening of the closed valve. Additional description of microfluidic devices including valves is provided in U.S. Patent Application No. 2005/0226742, the entire disclosure of which is hereby incorporated by reference in its entirety for all purposes.

IDT Exhibit 1020.2683

[0184] As illustrated in Fig. 6, actuation of control lines 532 will obstruct fluid lines 514. Typically, the control lines 532 are actuated concurrently by application of pressure to a pressure accumulator. Referring once again to Fig. 7, after closing of the interface valves 530, samples flow into sample chambers 510 via sample input lines 516 (612). As illustrated in Fig. 6, each sample chamber 510 is in fluid communication with multiple (e.g., two) sample input lines 516. In other embodiments, other numbers of sample input lines can be utilized. One of ordinary skill in the art would recognize many variations, modifications, and alternatives. Typically, the sample input lines, which are at least partially contained in a second layer of the microfluidic device, pass under the sample chambers, which are at least partially contained in a third layer of the microfluidic device. A via (not illustrated) passing from the sample input line up to the sample chamber, provide for fluid flow from the sample input line to the sample chamber. As shown in Fig. 6, samples flow in, for example, up the columns, past containment valves 540, which are open, through the vias extending out of the plane of the figure, and into the various sample chambers. Fluids such as air present in the sample chambers are expelled during loading of the samples as a result of the permeability of the elastomeric material used to fabricate the microfluidic device.

[0185] Referring once again to Fig. 7, assays flow into assay chambers 512 via assay input lines 518 (614). The assays flow through assay input lines 518, past containment valves 540, which are open, and through vias (not shown) passing from the assay input lines to the assay chambers. The closure of the interface valves 530 prevent the samples in the sample chambers and the assays in the assay chambers from mixing. Once the sample and assay are loaded, the fluid line between the assay chamber and the sample chamber is opened (616). In embodiments of the present invention, multiple fluid lines 514 are opened concurrently by opening of interface valves 530. At least a portion of the sample and at least a portion of the assay are combined to form a mixture (618). The mixture of the sample and assay is formed throughout the sample and assay chambers as well as the fluid lines connecting these chambers. Free interface diffusion is a process in which mixing is slow and the rate of species equilibration depends on the species' diffusion constants. Small molecules such as salts have large diffusion constants, and hence equilibrate quickly. Large molecules (e.g., proteins) have small diffusion constants, and equilibrate more slowly.

IDT Exhibit 1020.2684

[0186] The mixture is reacted to form a reaction product (620). A typical reaction included within the scope of the present invention is PCR, which involves thermocycling of the microfluidic device through a number of cycles as will be evident to one of skill in the art. The fluid line between the assay chamber and the sample chamber is closed (622) by actuation of interface valves 530. Closure of the interface valves separates the reaction product present in the sample chambers from the reaction products present in the assay chambers. Additionally, the containment valves 540 can be closed during thermocycling in order to prevent precipitation during the thermocycling process. A harvesting reagent flows from the harvesting port to the sample chamber (624) in order to begin the process of harvesting the reaction products present in the sample chambers. The harvesting port 136 is an example of a fluid input port useful in the harvesting process. As illustrated in Fig. 4, the reaction products flow down through the sample input lines 516 toward the sample input ports from which the samples were originally provides. Thus, in the illustrated method, removing the reaction products from the microfluidic device includes flowing the reaction products through at least a portion of the sample input line that was used to load the samples to the sample input port. Thus, the reaction product are removed from the microfluidic device (626) and output to the sample input ports, for example, sample input ports 120 illustrated in Fig. 2.

[0187] Dilation pumping is used in the illustrated embodiment to remove the reaction products from the microfluidic device as discussed in additional detail in relation to Figs. 9A-D. Referring once again to Fig. 2, control ports 130 and 132 or pressure accumulators 150 and 152 can be used to actuate the valves used in dilation pumping. Thus, embodiments provide valves for dilation pumping on the microfluidic device, which provides for removal of the reaction products from the microfluidic device. This contrasts with conventional designs in which such valves were not provided as part of the microfluidic device.

[0188] It should be appreciated that the specific steps illustrated in Fig. 7 provide a particular method of operating a microfluidic device according to an embodiment of the present invention. Other sequences of steps may also be performed according to alternative embodiments. For example, alternative embodiments of the present invention may perform the steps outlined above in a different order. Moreover, the individual steps illustrated in Fig. 7 may include

IDT Exhibit 1020.2685

multiple sub-steps that may be performed in various sequences as appropriate to the individual step. Furthermore, additional steps may be added or removed depending on the particular applications. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0189] Figs. 8A - 8D are simplified schematic diagrams illustrating fluid flow through unit cells of a microfluidic device during operation according to an embodiment of the present invention. Referring to Fig. 8A, the microfluidic device is illustrated during the sample and assay loading process. Containment valves 540 are open, allowing fluid flow along the sample input lines 516 to the various sample chambers 510. The open state of the containment valves also enables the assays to flow in through the assay input lines 518 to the various assay chambers 512. Using a single sample for each set of sample input lines (m of M samples) enables loading of a single sample in each column of sample chambers. Additionally, using a single assay for each assay input line (n of N assays) enables loading of a single assay in each row of assay chambers. The closed state of the interface valves 530 prevent mixing of the samples and the assays.

[0190] Fig. 8B illustrates the microfluidic device during a sample and assay mixing process as well as a subsequent reaction process (e.g., amplification). The containment valves 540 are closed, preventing fluid flow along the sample input lines 516. The closure of the containment valves thus isolates the sample chambers along a column from one another (with each column potentially containing a different sample). The closing of the containment valves 540 additionally isolates the assay chambers from other assay chambers in each row. After this chamber isolation is provided, the interface valves 530 are opened, enabling the samples and assays to mix via free interface diffusion (FID) and form M x N pairwise combinations. Although the materials in the pairs of sample/assay chambers are illustrated with the same shading in Fig. 8B, it will be appreciated that four different pairwise combinations are illustrated, one for each pair of sample/assay chambers. The multiple steps involved in mixing the samples and assays, then performing PCR amplification, are represented by a single drawing in Fig. 8B although it will be apparent to one of skill in the art that numerous steps, for example, multiple thermocycling steps, are involved in these processes.

IDT Exhibit 1020.2686

[0191] Fig. 8C illustrates the microfluidic device during sample chamber isolation and initial loading of the harvesting reagent. The interface valves 530 are closed to maintain isolation between the sample chambers and the assay chambers. Thus, the reaction products present in the sample chambers are recovered while the reaction products in the assay chambers are not recovered. The containment valves 540 are opened to allow the harvesting reagent to flow into the sample input lines 516 from the harvesting reagent input lines 430 illustrated in Fig. 5B. The harvesting reagent flows through the sample input lines, and into the sample chambers. In the illustrated embodiment, the reaction products are removed as the harvesting reagent flows through the sample input lines and the sample chambers in response to a dilation pumping process described in additional detail in relation to Figs. 9A-9D. As illustrated in Fig. 8C, the harvesting reagent has only reached the middle region of the upper reaction chambers. As the dilation pumping process continues, the harvesting reagent will be progressively introduced into subsequent sample chambers, thereby displacing the reaction products. Eventually, the reaction products associated with each sample will be recovered as a pooled fluid including the reaction products and the harvesting reagent in the sample input port from which the sample was originally dispensed.

[0192] It should be noted that the straight line representing the interface between the harvesting reagent and the reaction products is shown merely for purposes of simplicity and it will be apparent to one of skill in the art that in practice, a more complicated interface will be present.

[0193] Fig. 8D illustrates the microfluidic device during final loading of the harvesting reagent and recovery of the reaction products. As the dilation pumping process continues, harvesting reagent is introduced into additional sample chambers progressively farther from the harvesting reagent input lines. The state of the recovery process illustrated in Fig. 8D shows that the reaction products have been flushed from the sample chambers, which are now filled with harvesting reagent. Although only four sample chambers are illustrated in Fig. 8D, it will be appreciated that recovery of the reaction products from all the sample chambers in the array is provided by the embodiments described herein.

[0194] Figs. 9A-9D are simplified schematic diagrams illustrating fluid flow through a microfluidic device during operation according to an embodiment of the

IDT Exhibit 1020.2687

present invention. Fig. 9A illustrates a portion of a microfluidic device according to an embodiment of the present invention during loading of samples and assays. The illustrated portion includes a harvesting reagent input line 810, vent and loading bowl portions 830, and isolation valve 840. As described more fully throughout the present specification, valves 820 and 822 are used to perform dilation pumping of reaction products present in sample chambers 510. As illustrated in Fig. 9A, valve 820 is closed and valve 822 is open. Valves 820 and 822 are typically "push-up" valves described elsewhere in the present specification.

[0195] Samples are loaded into sample chambers 510 and assays are loaded into assay chambers 512 as described in relation to Fig. 6. Interface valves 530 are closed, preventing mixing of the samples and assays. Vents and loading bowls are provided in some embodiments to allow for reductions in effects related to depletion fronts. The inventors have observed that in loading samples into microfluidic devices (e.g., through the vertical sample input lines illustrated in Fig. 9A), binding of a portion of the sample present at the leading edge of the flow path to the material of the microfluidic device will produce a depletion front in which one or more components of the sample are depleted as a result of this binding process. The provision of the vents and loading bowls 830 enables the user to push the depletion front through the various sample chambers of the microfluidic device and store the depleted sample material in the loading bowls 830. Eventually, as the depleted sample material is flushed through the device into the loading bowls, the sample contained in the microfluidic device will be substantially undepleted.

[0196] Isolation valve 840 is open during the sample and assay loading process to enable the depletion front to flow into the loading bowls 830. Valve 822 is open, allowing the samples to flow through the sample input lines to the various sample chambers. Since valve 820 is closed, samples are not allowed to pass into the harvesting reagent input line 810. It should be noted that containment valves 540 are illustrated in the closed state in Fig. 9A. Containment valves are open during sample and assay loading and then are closed as illustrated after sample and assay loading is complete. The containment valves isolate the various pairs of reaction and assay chambers from other pairs containing the various pairwise combinations.

[0197] In Fig. 9A, only a single column of the microfluidic device is illustrated for purposes of clarity. It is understood that additional columns are

IDT Exhibit 1020.2688

provided by the microfluidic devices as illustrated, for example, in Fig. 6. Moreover, much of the column is not illustrated for the purposes of clarity. The two sets of sample/assay chambers illustrated are those at the top and bottom of Fig. 4A, respectively. The set adjacent valve 820 is the topmost set and the set adjacent valve 822 is the bottommost set. Thus, these diagrams are merely representative and not intended to limit the scope of the present invention.

[0198] Fig. 9B illustrates mixing of the samples and assays and a subsequent reaction (e.g., amplification) process. In order to mix the samples and reagents, interface valves 530 are placed in the open position as shown. Closure of containment valves 540 seals the reaction products in the sample and assay chambers along with the connecting fluid lines. As illustrated in Fig. 9B, isolation valve 840 is closed, preventing fluid flow between the sample input lines and the loading bowls 830. Actuation of valve 840 to place it in the open or closed position is performed using a pressure accumulator (not shown) in some embodiments and using other actuation techniques in other embodiments, for example, mechanical, electrostatic, electromechanical, thermodynamic, piezoelectric, or the like. Such additional techniques may also be applicable to other valves described herein. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0199] Although Fig. 9B illustrates mixing and reaction using a single image, one of skill in the art will appreciate that multiple thermal cycles may be used to amplify DNA using the PCR process. Thus, this simple figure is intended to show mixing and subsequent reactions that can occur in the microfluidic device.

[0200] Fig. 9C illustrates a portion of the microfluidic device in a first stage of a reaction product harvesting process. A harvesting reagent flows from a harvesting port as illustrated in Fig. 5A through harvesting reagent input line 810 toward the sample chambers. As shown in Fig. 9A, valve 820 is open, allowing the harvesting reagent to flow into the topmost sample chamber. Closure of the interface valves 530 prevents the harvesting reagent from flowing into the assay chambers, which also contain reaction products. The extent to which the harvesting reagent initially fills the sample input lines and sample input chambers is limited by the closure of valve 822. As illustrated, the harvesting reagent has only partially filled a portion of the first sample chamber. The illustration of a side of the sample chamber being filled with harvesting reagent is merely provided by way of example, since the flow of the

IDT Exhibit 1020.2689

harvesting reagent into the sample chamber is actually through vias extending from the plane of the figure. Closure of isolation valve 840 prevents harvesting reagent from flowing into the loading bowls, although other embodiments may enable such a flow if desired.

[0201] Fluid pressure resulting from the flow of the harvesting reagent into the array portion of the microfluidic device results in expansion of the sample input lines and sample chambers above the valve 822. The pump cycle is initiated by this pressurization of the sample chambers. As described below, closing of valve 820 and opening of valve 822 will enable the pressurized harvesting reagent and reaction products to be recovered from the microfluidic device as it flows through the microfluidic device.

[0202] Fig. 9D illustrates a portion of the microfluidic device in a second stage of the reaction product harvesting process. Although a second stage is illustrated, this is not intended to imply that the second stage immediately follows the first stage. As described below, the second stage is typically separated from the first stage by one or more intermediate stages of dilation pumping.

[0203] Dilation pumping (also know as volumetric capacitive pumping) is a method of operating a properly configured integrated fluidic circuit (microfluidic device) to obtain precise, low rate, low volume pumping through all configured elements of the microfluidic device. Dilation pumping is unique to microfluidic circuits that utilize channels that have one or more channel walls formed from an elastomeric material. As an example, the flow of the harvesting reagent through the sample input lines and sample chambers is considered volumetric capacitive pumping. Pumping proceeds by the closure of valves 822 and the opening of valves 820. As discussed above, harvesting reagent ports (not illustrated) are pressurized to introduce the harvesting reagent into the topmost sample input lines and sample chambers, which can be considered as a channel. The pressurization of microfluidic channels with at least one channel wall formed from an elastomeric material results in expansion of the elastomeric wall(s) outward from the channel with a resulting increase in channel volume that is proportional to the fluidic pressure (or gaseous pressure in alternate embodiments) within the channel, the elastic properties of the elastomeric channel wall material such as Young's modulus, and the length and cross sectional area of the channel. The sample input lines and sample chambers are

IDT Exhibit 1020.2690

allowed to pressurize and then valves 820 is closed as illustrated in Fig. 9D. Following closure of valves 820, valves 822 are opened. The pumped volume through the sample input lines and the sample chambers is equal to the expanded volume of the channel when under pressure minus the native volume of the channel when pressure is released and the expanded elastomeric channel wall(s) is allow to relax. Dilation pumping is continued through repetitive cycles of closing 822, opening 820, pressurizing the sample input lines and sample chambers, closing 820, and opening 822. In this manner, continuous or discontinuous low volume pumping may be accomplished at precisely controlled flow rates.

[0204] Thus embodiments provide a method of dilation pumping that includes closing a first valve disposed between the sample chamber and the sample input port (i.e., valve 822), opening a second valve disposed between the harvesting port and the sample chamber (i.e., valve 820), closing the second valve, opening the first valve, and repeating these steps a predetermined number of times. Between the steps of opening the second valve and closing the second valve, the harvesting reagent flows into the sample input lines and sample chambers, pressurizing the channel as described above. After the dilation pumping process is complete, harvesting reagent substantially fills the sample input lines and sample chambers (e.g., recovery rates > 95%), thereby pooling the reaction products associated with a given sample in the sample input port from which the given sample was initially dispensed.

[0205] Dilation pumping provides benefits not typically available using conventional techniques. For example, dilation pumping enables for a slow removal of the reaction products from the microfluidic device. In an exemplary embodiment, the reaction products are recovered at a fluid flow rate of less than 100 μl per hour. In this example, for 48 reaction products distributed among the reaction chambers in each column, with a volume of each reaction product of about 1.5 μl, removal of the reaction products in a period of about 30 minutes, will result in a fluid flow rate of 72 μl/hour. (i.e., 48 * 1.5 / 0.5 hour). In other embodiments, the removal rate of the reaction products is performed at a rate of less than 90 μl/hr, 80 μl/hr, 70 μl/hr, 60 μl/hr, 50 μl/hr, 40 μl/hr, 30 μl/hr, 20 μl/hr, 10 μl/hr, 9 μl/hr, less than 8 μl/hr, less than 7 μl/hr, less than 6 μl/hr, less than 5 μl/hr, less than 4 μl/hr, less than 3 μl/hr, less than 2 μl/hr, less than 1 μl/hr, or less than 0.5 μl/hr.

IDT Exhibit 1020.2691

[0206] Dilation pumping results in clearing of substantially a high percentage and potentially all the reaction products present in the microfluidic device. Some embodiments remove more than 75% of the reaction products present in the reaction chambers (e.g., sample chambers) of the microfluidic device. As an example, some embodiments remove more than 80%, 85%, 90%, 92 %, 95%, 96%, 97%, 98%, or 99% of the reaction products present in the reaction chambers.

[0207] In some embodiments, a harvesting valve is provided on the microfluidic device to obstruct the flow of harvesting reagent through the device. Application of a pressure source to a harvesting input port results in flow of harvesting fluid (e.g., a harvesting liquid) through harvest reagent input lines up to the harvesting valve. The permeability of the materials utilized to fabricate the microfluidic device enables such a harvesting fluid to fill the harvest reagent input lines, typically expelling air initially present in such lines. The presence of the harvesting valve will obstruct the flow of the harvest reagent at the location of the harvesting valve. Actuation (i.e., opening) of the harvesting valve will result in the harvesting fluid flowing through the harvest reagent input lines downstream of the harvesting valve. In other embodiments, a harvesting valve is replaced with one or more other suitable valves as appropriate to the particular application. For example, in the embodiment illustrated in Figs. 9A-9D, valve 820 serves to prevent flow of harvesting reagent until the dilation pumping process is initiated.

[0208] Fabrication methods using elastomeric materials and methods for design of devices and their components have been described in detail in the scientific and patent literature. See, *e.g.*, Unger *et al.* (2000) Science 288:113-116; U.S. Pat. Nos. US 6,960,437 (Nucleic acid amplification utilizing microfluidic devices); 6,899,137 (Microfabricated elastomeric valve and pump systems); 6,767,706 (Integrated active flux microfluidic devices and methods); 6,752,922 (Microfluidic chromatography); 6,408,878 (Microfabricated elastomeric valve and pump systems); 6,645,432 (Microfluidic systems including three-dimensionally arrayed channel networks); U.S. Patent Application Publication Nos. 2004/0115838; 2005/0072946; 2005/0000900; 2002/0127736; 2002/0109114; 2004/0115838; 2003/0138829; 2002/0164816; 2002/0127736; and 2002/0109114; PCT Publication Nos. WO 2005/084191; WO 05/030822A2; and WO 01/01025; Quake & Scherer, 2000, "From micro to nanofabrication with soft materials" Science 290: 1536-40; Unger *et al.*,

IDT Exhibit 1020.2692

2000, "Monolithic microfabricated valves and pumps by multilayer soft lithography" Science 288:113-116; Thorsen *et al.*, 2002, "Microfluidic large-scale integration" Science 298:580-584; Chou *et al.*, 2000, "Microfabricated Rotary Pump" Biomedical Microdevices 3:323-330; Liu *et al.*, 2003, "Solving the "world-to-chip" interface problem with a microfluidic matrix" Analytical Chemistry 75, 4718-23, Hong et al, 2004, "A nanoliter-scale nucleic acid processor with parallel architecture" Nature Biotechnology 22:435-39.

[0209] According to certain embodiments describer herein, the detection and/or quantification of one or more target nucleic acids from one or more samples may generally be carried out on a microfluidic device by obtaining a sample, optionally pre-amplifying the sample, and distributing the optionally pre-amplified sample, or aliquots thereof, into reaction chambers of a microfluidic device containing the appropriate buffers, primers, optional probe(s), and enzyme(s), subjecting these mixtures to amplification, and querying the aliquots for the presence of amplified target nucleic acids. The sample aliquots may have a volume of less than 1 picoliter or, in various embodiments, in the range of about 1 picoliter to about 500 nanoliters, in a range of about 2 picoliters to about 50 picoliters, in a range of about 5 picoliters to about 25 picoliters, in the range of about 100 picoliters to about 20 nanoliters, in the range of about 1 nanoliter to about 20 nanoliters, and in the range of about 5 nanoliters to about 15 nanoliters. In many embodiments, sample aliquots account for the majority of the volume of the amplification mixtures. Thus, amplification mixtures can have a volume of less than 1 picoliter or, in various embodiments about 2, about 5 about 7, about 10, about 15, about 20, about 25, about 50, about 100, about 250, about 500, and about 750 picoliters; or about 1, about 2, about 5, about 7, about 15, about 20, about 25, about 50, about 250, and about 500 nanoliters. The amplification mixtures can also have a volume within any range bounded by any of these values (e.g., about 2 picoliters to about 50 picoliters).

[0210] In certain embodiments, multiplex detection is carried out in individual amplification mixture, *e.g.*, in individual reaction chambers of a microfluidic device, which can be used to further increase the number of samples and/or targets that can be analyzed in a single assay or to carry out comparative methods, such as comparative genomic hybridization (CGH). In various embodiments, up to 2, 3, 4, 5, 6, 7, 8, 9, 10,

IDT Exhibit 1020.2693

50, 100, 500, 1000, 5000, 10000 or more amplification reactions are carried out in each individual reaction chamber.

[0211] In specific embodiments, the assay usually has a dynamic range of at least 3 orders of magnitude, more often at least 4, at least 5, at least 6, at least 7, or at least 8 orders of magnitude.

Quantitative Real-Time PCR and Other Detection and Quantification Methods

[0212] Any method of detection and/or quantification of nucleic acids can be used in the invention to detect amplification products. In one embodiment, PCR (polymerase chain reaction) is used to amplify and/or quantify target nucleic acids. In other embodiments, other amplification systems or detection systems are used, including, e.g., systems described in U.S. Pat. No. 7,118,910 (which is incorporated herein by reference in its entirety for its description of amplification/detection systems) and Invader assays; PE BioSystems). In particular embodiments, real-time quantification methods are used. For example, "quantitative real-time PCR" methods can be used to determine the quantity of a target nucleic acid present in a sample by measuring the amount of amplification product formed during the amplification process itself.

[0213] Fluorogenic nuclease assays are one specific example of a real-time quantification method that can be used successfully in the methods described herein. This method of monitoring the formation of amplification product involves the continuous measurement of PCR product accumulation using a dual-labeled fluorogenic oligonucleotide probe--an approach frequently referred to in the literature as the "TaqMan® method." See U.S. Pat. No. 5,723,591; Heid et al., 1996, Real-time quantitative PCR Genome Res. 6:986-94, each incorporated herein by reference in their entireties for their descriptions of fluorogenic nuclease assays. It will be appreciated that while "TaqMan® probes" are the most widely used for qPCR, the invention is not limited to use of these probes; any suitable probe can be used.

[0214] Other detection/quantification methods that can be employed in the present invention include FRET and template extension reactions, molecular beacon detection, Scorpion detection, Invader detection, and padlock probe detection.

IDT Exhibit 1020.2694

[0215] FRET and template extension reactions utilize a primer labeled with one member of a donor/acceptor pair and a nucleotide labeled with the other member of the donor/acceptor pair. Prior to incorporation of the labeled nucleotide into the primer during a template-dependent extension reaction, the donor and acceptor are spaced far enough apart that energy transfer cannot occur. However, if the labeled nucleotide is incorporated into the primer and the spacing is sufficiently close, then energy transfer occurs and can be detected. These methods are particularly useful in conducting single base pair extension reactions in the detection of single nucleotide polymorphisms and are described in U.S. Patent No. 5,945,283 and PCT Publication WO 97/22719.

[0216] With molecular beacons, a change in conformation of the probe as it hybridizes to a complementary region of the amplified product results in the formation of a detectable signal. The probe itself includes two sections: one section at the 5' end and the other section at the 3' end. These sections flank the section of the probe that anneals to the probe binding site and are complementary to one another. One end section is typically attached to a reporter dye and the other end section is usually attached to a quencher dye. In solution, the two end sections can hybridize with each other to form a hairpin loop. In this conformation, the reporter and quencher dye are in sufficiently close proximity that fluorescence from the reporter dye is effectively quenched by the quencher dye. Hybridized probe, in contrast, results in a linearized conformation in which the extent of quenching is decreased. Thus, by monitoring emission changes for the two dyes, it is possible to indirectly monitor the formation of amplification product. Probes of this type and methods of their use are described further, for example, by Piatek et al., 1998, Nat. Biotechnol. 16:359-63; Tyagi, and Kramer, 1996, Nat. Biotechnology 14:303-308; and Tyagi, et al., 1998, Nat. Biotechnol. 16:49-53 (1998).

[0217] The Scorpion detection method is described, for example, by Thelwell et al. 2000, Nucleic Acids Research, 28:3752-3761 and Solinas et al., 2001, "Duplex Scorpion primers in SNP analysis and FRET applications" Nucleic Acids Research 29:20. Scorpion primers are fluorogenic PCR primers with a probe element attached at the 5'-end via a PCR stopper. They are used in real-time amplicon-specific detection of PCR products in homogeneous solution. Two different formats are possible, the "stem–loop" format and the "duplex" format. In both cases the probing

IDT Exhibit 1020.2695

mechanism is intramolecular. The basic elements of Scorpions in all formats are: (i) a PCR primer; (ii) a PCR stopper to prevent PCR read-through of the probe element; (iii) a specific probe sequence; and (iv) a fluorescence detection system containing at least one fluorophore and quencher. After PCR extension of the Scorpion primer, the resultant amplicon contains a sequence that is complementary to the probe, which is rendered single-stranded during the denaturation stage of each PCR cycle. On cooling, the probe is free to bind to this complementary sequence, producing an increase in fluorescence, as the quencher is no longer in the vicinity of the fluorophore. The PCR stopper prevents undesirable read-through of the probe by Taq DNA polymerase.

[0218] Invader assays (Third Wave Technologies, Madison, WI) are used particularly for SNP genotyping and utilize an oligonucleotide, designated the signal probe, that is complementary to the target nucleic acid (DNA or RNA) or polymorphism site. A second oligonucleotide, designated the Invader Oligo, contains the same 5' nucleotide sequence, but the 3' nucleotide sequence contains a nucleotide polymorphism. The Invader Oligo interferes with the binding of the signal probe to the target nucleic acid such that the 5' end of the signal probe forms a "flap" at the nucleotide containing the polymorphism. This complex is recognized by a structure specific endonuclease, called the Cleavase enzyme. Cleavase cleaves the 5' flap of the nucleotides. The released flap binds with a third probe bearing FRET labels, thereby forming another duplex structure recognized by the Cleavase enzyme. This time, the Cleavase enzyme cleaves a fluorophore away from a quencher and produces a fluorescent signal. For SNP genotyping, the signal probe will be designed to hybridize with either the reference (wild type) allele or the variant (mutant) allele. Unlike PCR, there is a linear amplification of signal with no amplification of the nucleic acid. Further details sufficient to guide one of ordinary skill in the art are provided by, for example, Neri, B.P., et al., *Advances in Nucleic Acid and Protein Analysis* 3826:117-125, 2000) and U.S. Patent No. 6,706,471.

[0219] Padlock probes (PLPs) are long (e.g., about 100 bases) linear oligonucleotides. The sequences at the 3' and 5' ends of the probe are complementary to adjacent sequences in the target nucleic acid. In the central, noncomplementary region of the PLP there is a "tag" sequence that can be used to identify the specific PLP. The tag sequence is flanked by universal priming sites,

IDT Exhibit 1020.2696

which allow PCR amplification of the tag. Upon hybridization to the target, the two ends of the PLP oligonucleotide are brought into close proximity and can be joined by enzymatic ligation. The resulting product is a circular probe molecule catenated to the target DNA strand. Any unligated probes (i.e., probes that did not hybridize to a target) are removed by the action of an exonuclease. Hybridization and ligation of a PLP requires that both end segments recognize the target sequence. In this manner, PLPs provide extremely specific target recognition.

[0220] The tag regions of circularized PLPs can then be amplified and resulting amplicons detected. For example, TaqMan® real-time PCR can be carried out to detect and quantify the amplicon. The presence and amount of amplicon can be correlated with the presence and quantity of target sequence in the sample. For descriptions of PLPs see, e.g., Landegren et al., 2003, Padlock and proximity probes for in situ and array-based analyses: tools for the post-genomic era, *Comparative and Functional Genomics* 4:525-30; Nilsson et al., 2006, Analyzing genes using closing and replicating circles *Trends Biotechnol*. 24:83-8; Nilsson et al., 1994, Padlock probes: circularizing oligonucleotides for localized DNA detection, *Science* 265:2085-8.

[0221] In particular embodiments, fluorophores that can be used as detectable labels for probes include, but are not limited to, rhodamine, cyanine 3 (Cy 3), cyanine 5 (Cy 5), fluorescein, Vic™, Liz™., Tamra™, 5-Fam™, 6-Fam™, and Texas Red (Molecular Probes). (Vic™, Liz™, Tamra™, 5-Fam™, 6-Fam™ are all available from Applied Biosystems, Foster City, Calif.).

[0222] Devices have been developed that can perform a thermal cycling reaction with compositions containing a fluorescent indicator, emit a light beam of a specified wavelength, read the intensity of the fluorescent dye, and display the intensity of fluorescence after each cycle. Devices comprising a thermal cycler, light beam emitter, and a fluorescent signal detector, have been described, e.g., in U.S. Pat. Nos. 5,928,907; 6,015,674; and 6,174,670.

[0223] In some embodiments, each of these functions can be performed by separate devices. For example, if one employs a Q-beta replicase reaction for amplification, the reaction may not take place in a thermal cycler, but could include a

IDT Exhibit 1020.2697

light beam emitted at a specific wavelength, detection of the fluorescent signal, and calculation and display of the amount of amplification product.

[0224] In particular embodiments, combined thermal cycling and fluorescence detecting devices can be used for precise quantification of target nucleic acids. In some embodiments, fluorescent signals can be detected and displayed during and/or after one or more thermal cycles, thus permitting monitoring of amplification products as the reactions occur in "real-time." In certain embodiments, one can use the amount of amplification product and number of amplification cycles to calculate how much of the target nucleic acid sequence was in the sample prior to amplification.

[0225] According to some embodiments, one can simply monitor the amount of amplification product after a predetermined number of cycles sufficient to indicate the presence of the target nucleic acid sequence in the sample. One skilled in the art can easily determine, for any given sample type, primer sequence, and reaction condition, how many cycles are sufficient to determine the presence of a given target nucleic acid.

[0226] According to certain embodiments, one can employ an internal standard to quantify the amplification product indicated by the fluorescent signal. See, e.g., U.S. Pat. No. 5,736,333.

[0227] In various embodiments, employing preamplification, the number of preamplification cycles is sufficient to add one or more nucleotide tags to the target nucleotide sequences, so that the relative copy numbers of the tagged target nucleotide sequences is substantially representative of the relative copy numbers of the target nucleic acids in the sample. For example, preamplification can be carried out for 2-20 cycles to introduce the sample-specific or set-specific nucleotide tags. In other embodiments, detection is carried out at the end of exponential amplification, i.e., during the "plateau" phase, or endpoint PCR is carried out. In this instance, preamplification will normalize amplicon copy number across targets and across samples. In various embodiments, preamplification and/or amplification can be carried out for about: 2, 4, 10, 15, 20, 25, 30, 35, or 40 cycles or for a number of cycles falling within any range bounded by any of these values.

IDT Exhibit 1020.2698

Labeling Strategies

[0228] Any suitable labeling strategy can be employed in the methods of the invention. Where the assay mixture is aliquoted, and each aliquot is analyzed for presence of a single amplification product, a universal detection probe can be employed in the amplification mixture. In particular embodiments, real-time PCR detection can be carried out using a universal qPCR probe. Suitable universal qPCR probes include double-stranded DNA dyes, such as SYBR Green, Pico Green (Molecular Probes, Inc., Eugene, OR), Eva Green (Biotinum), ethidium bromide, and the like (see Zhu et al., 1994, *Anal. Chem.* 66:1941-48). Suitable universal qPCR probes also include sequence-specific probes that bind to a nucleotide sequence present in all amplification products. Binding sites for such probes can be conveniently introduced into the tagged target nucleic acids during amplification.

[0229] Alternatively, one or more target-specific qPCR probes (i.e., specific for a target nucleotide sequence to be detected) is employed in the amplification mixtures to detect amplification products. Target-specific probes could be useful, e.g., when only a few target nucleic acids are to be detected in a large number of samples. For example, if only three targets were to be detected, a target-specific probe with a different fluorescent label for each target could be employed. By judicious choice of labels, analyses can be conducted in which the different labels are excited and/or detected at different wavelengths in a single reaction. See, e.g., Fluorescence Spectroscopy (Pesce et al., Eds.) Marcel Dekker, New York, (1971); White et al., Fluorescence Analysis: A Practical Approach, Marcel Dekker, New York, (1970); Berlman, Handbook of Fluorescence Spectra of Aromatic Molecules, 2nd ed., Academic Press, New York, (1971); Griffiths, Colour and Constitution of Organic Molecules, Academic Press, New York, (1976); Indicators (Bishop, Ed.). Pergamon Press, Oxford, 19723; and Haugland, Handbook of Fluorescent Probes and Research Chemicals, Molecular Probes, Eugene (1992).

Removal of Undesired Reaction Components

[0230] It will be appreciated that reactions involving complex mixtures of nucleic acids in which a number of reactive steps are employed can result in a variety of unincorporated reaction components, and that removal of such unincorporated reaction components, or reduction of their concentration, by any of a variety of clean-

IDT Exhibit 1020.2699

up procedures can improve the efficiency and specificity of subsequently occurring reactions. For example, it may be desirable, in some embodiments, to remove, or reduce the concentration of preamplification primers prior to carrying out the amplification steps described herein.

[0231] In certain embodiments, the concentration of undesired components can be reduced by simple dilution. For example, preamplified samples can be diluted about 2-, 5-, 10-, 50-, 100-, 500-, 1000-fold prior to amplification to improve the specificity of the subsequent amplification step.

[0232] In some embodiments, undesired components can be removed by a variety of enzymatic means. Alternatively, or in addition to the above-described methods, undesired components can be removed by purification. For example, a purification tag can be incorporated into any of the above-described primers (e.g., into the barcode nucleotide sequence) to facilitate purification of the tagged target nucleotides.

[0233] In particular embodiments, clean-up includes selective immobilization of the desired nucleic acids. For example, desired nucleic acids can be preferentially immobilized on a solid support. In an illustrative embodiment, an affinity moiety, such as biotin (e.g., photo-biotin), is attached to desired nucleic acid, and the resulting biotin-labeled nucleic acids immobilized on a solid support comprising an affinity moiety-binder such as streptavidin. Immobilized nucleic acids can be queried with probes, and non-hybridized and/or non-ligated probes removed by washing (See, e.g., Published P.C.T. Application WO 03/006677 and USSN 09/931,285.) Alternatively, immobilized nucleic acids can be washed to remove other components and then released from the solid support for further analysis. This approach can be used, for example, in recovering target amplicons from amplification mixtures after the addition of primer binding sites for DNA sequencing. In particular embodiments, an affinity moiety, such as biotin, can be attached to an amplification primer such that amplification produces an affinity moiety-labeled (e.g., biotin-labeled) amplicon. Thus, for example, where three primers are employed to add barcode and nucleotide tag elements to a target nucleotide sequence, as described above, at least one of the barcode or reverse primers can include an affinity moiety. Where four primers (two inner primers and two outer primers) are employed to add desired element to a target nucleotide sequence, at least one of the outer primers can include an affinity moiety.

IDT Exhibit 1020.2700

Data Output and Analysis

[0234] In certain embodiments, when the methods of the invention are carried out on a matrix-type microfluidic device, the data can be output as a heat matrix (also termed "heat map"). In the heat matrix, each square, representing a reaction chamber on the DA matrix, has been assigned a color value which can be shown in gray scale, but is more typically shown in color. In gray scale, black squares indicate that no amplification product was detected, whereas white squares indicate the highest level of amplification produce, with shades of gray indicating levels of amplification product in between. In a further aspect, a software program may be used to compile the data generated in the heat matrix into a more reader-friendly format.

Applications

[0235] The methods of the invention are applicable to any technique aimed at detecting the presence or amount of one or more target nucleic acids in a nucleic acid sample. Thus, for example, these methods are applicable to identifying the presence of particular polymorphisms (such as SNPs), alleles, or haplotypes, or chromosomal abnormalities, such as amplifications, deletions, or aneuploidy. The methods may be employed in genotyping, which can be carried out in a number of contexts, including diagnosis of genetic diseases or disorders, pharmacogenomics (personalized medicine), quality control in agriculture (e.g., for seeds or livestock), the study and management of populations of plants or animals (e.g., in aquaculture or fisheries management or in the determination of population diversity), or paternity or forensic identifications. The methods of the invention can be applied in the identification of sequences indicative of particular conditions or organisms in biological or environmental samples. For example, the methods can be used in assays to identify pathogens, such as viruses, bacteria, and fungi). The methods can also be used in studies aimed at characterizing environments or microenvironments, e.g., characterizing the microbial species in the human gut.

[0236] These methods can also be employed in determinations DNA or RNA copy number. Determinations of aberrant DNA copy number in genomic DNA is useful, for example, in the diagnosis and/or prognosis of genetic defects and diseases, such as cancer. Determination of RNA "copy number," i.e., expression level is useful for expression monitoring of genes of interest, e.g., in different individuals, tissues, or

IDT Exhibit 1020.2701

cells under different conditions (e.g., different external stimuli or disease states) and/or at different developmental stages.

[0237] In addition, the methods can be employed to prepare nucleic acid samples for further analysis, such as, e.g., DNA sequencing.

[0238] Finally, nucleic acid samples can be tagged as a first step, prior subsequent analysis, to reduce the risk that mislabeling or cross-contamination of samples will compromise the results. For example, any physician's office, laboratory, or hospital could tag samples immediately after collection, and the tags could be confirmed at the time of analysis. Similarly, samples containing nucleic acids collected at a crime scene could be tagged as soon as practicable, to ensure that the samples could not be mislabeled or tampered with. Detection of the tag upon each transfer of the sample from one party to another could be used to establish chain of custody of the sample.

Kits

[0239] Kits according to the invention include one or more reagents useful for practicing one or more assay methods of the invention. A kit generally includes a package with one or more containers holding the reagent(s) (e.g., primers and/or probe(s)), as one or more separate compositions or, optionally, as admixture where the compatibility of the reagents will allow. The kit can also include other material(s) that may be desirable from a user standpoint, such as a buffer(s), a diluent(s), a standard(s), and/or any other material useful in sample processing, washing, or conducting any other step of the assay.

[0240] Kits according to the invention generally include instructions for carrying out one or more of the methods of the invention. Instructions included in kits of the invention can be affixed to packaging material or can be included as a package insert. While the instructions are typically written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to, electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), RF tags, and the like. As used herein, the term "instructions" can include the address of an internet site that provides the instructions.

IDT Exhibit 1020.2702

[0241] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

[0242] In addition, all other publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

EXAMPLES

Example 1

Multi-Primer Amplification Method For Barcoding Of Target Nucleic Acids in Preparation for DNA Sequencing

[0243] Genomic DNA samples (BioChain, USA) at 100 and 0 ng/ml (negative control ["NTC"]) were amplified for 25 cycles 7900HT Fast Real-Time PCR System (Applied Biosystems, USA) with the following primer pairs at 200 nM per primer: 1) 454 tails; 2) A5 specific primers; and 3) the three primers shown in Fig. 10. PCR was performed in 15 μl reaction volumes containing 7.5 μl of FastStart TaqMan® Probe Master (Roche Diagnostics, USA), 0.75 μl of DA sample loading reagent (Fluidigm Corp, USA) and 6.75 μl of sample. Thermal cycling condition included an initial hot start at 50 °C for 2 minutes and at 94 °C for 10 minutes, followed by 25 cycles at 94 °C for 15 s, 70 °C for 5s, 60°C for 30 s and 72°C for 90 s. The resulting amplification products were run on an electrophoresis gel (Invitrogen, USA) using 8 μl of the reaction mixture per lane following the manufacturer instruction. See Fig. 11, which shows that the 3-primer method produced an amplicon of the correct size. The amplicons generated from PCR amplification were purified using Ampure Beads® and then re-amplified on a PCR plate with 454 tail primers, followed by Sanger sequencing with either 454 tail primers, which showed that the 3-primer method generated an amplicon having the correct sequence.

IDT Exhibit 1020.2703

Example 2

Multi-Primer Amplification Method For Quantifying Target Nucleic Acids in Preparation for DNA Sequencing

[0244] Primers for preparing genomic DNA for sequencing using various DNA conventional DNA sequencing methods are shown below.

ShotGun Forward: 5'-CCATCTCATCCCTGCGTGTC-3' (SEQ ID NO:1)

ShotGun Reverse: 5'-CCTATCCCCTGTGTGCCTTG-3' (SEQ ID NO:2)

ShotGun UPR Forward: 5'-GGCGGCGACCATCTCATCCCTGCGTGTC-3' (SEQ ID NO:3)

MID Forward: 5'-GCCTCCCTCGCGCCATCAG-3' (SEQ ID NO:4)

MID Reverse: 5'-GCCTTGCCAGCCCGCTCAG-3' (SEQ ID NO:5)

MID UPR Forward: 5'-GGCGGCGAGCCTCCCTCGCGCCATCAG-3' (SEQ ID NO:6)

Solexa Forward: 5'-ACACTCTTTCCCTACACGA-3' (SEQ ID NO:7)

Solexa Reverse: 5'-CAAGCAGAAGACGGCATA-3' (SEQ ID NO:8)

Solexa UPR Forward: 5'-GGCGGCGAACACTCTTTCCCTACACGA-3' (SEQ ID NO:9)

Solid Forward: 5'-CCACTACGCCTCCGCTTTCCTCTCTATG-3' (SEQ ID NO:10)

Solid Reverse: 5'-CTGCCCCGGGTTCCTCATTCT-3' (SEQ ID NO:11)

Solid UPR Forward: 5'-GGCGGCGACCACTACGCCTCCGCTTTCCTCTCTATG-3' (SEQ ID NO:12)

IDT Exhibit 1020.2704

454 Titanium Forward: 5'-CCATCTCATCCCTGCGTG-3' (SEQ ID NO:13)

454 Titanium Reverse: 5'-CCTATCCCCTGTGTGCCTTG-3' (SEQ ID NO:14)

454 Titanium UPR Forward: 5'-GGCGGCGACCATCTCATCCCTGCGTG-3' (SEQ ID NO:15)

Solexa smRNA Forward: 5'-TAATGATACGGCGACCACC-3' (SEQ ID NO:16)

Solexa smRNA Reverse: 5'-ACAAGCAGAAGACGGCATAC-3' (SEQ ID NO:17)

Solexa smRNA UPL Forward: 5'-GGCGGCGATAATGATACGGCGACCAC-3' (SEQ ID NO:18)

[0245] The properties of these primers is shown in Table 1 below.

Table 1

454-standard (ShotGun)	Primer	Length (nt)	CG%	Tm (°C)	Primer-Dimer
	ShotGun Forward:	20	60	68.4	No self/cross-dimer, 1.5°C diff in Tm
	ShotGun Reverse:	20	60	66.9	
	ShotGun UPR Forward:	28	67.8	84.8	
454-MID					
	MID Forward:	19	73.6	74.9	4-bases of self-dimer(F.UPL) & cross-dimer(F./UPL , R/UPL)
	MID Reverse:	19	73.6	74.9	High GC
	MID UPR Forward:	27	77.7	88.5	
Solexa					
	Solexa Forward:	19	47.3	57.8	No dimer, 2.1°C diff in Tm
	Solexa Reverse:	18	50	60.6	Low GC
	Solexa UPR Forward:	27	59.2	78.4	
Solid					
	Solid Forward:	28	57.1	74.7	Strong self-dimer & cross-dimer
	Solid Reverse:	21	61.9	72.5	variety of GC & Tm
	Solid UPR Forward:	36	63.8	85.6	

IDT Exhibit 1020.2705

[0246] The reaction mixture used for amplification of genomic DNA to incorporate primer sequences is given below in Table 2.

Table 2

	Add V µl of TE into dry probe tube	
100uM stock solution	V= "Total nmol" value of the dry probe * 10	
10X Fluidigm Assay		
	100 µMol	
Forward:	4	2000 nM
UPR Forward:	4	2000 nM
Reverse:	8	4000 nM
TE:	184	
Total:	200	

Example 3

Additional Illustrative Primers For Barcoding Of Target Nucleic Acids in Preparation for 454 DNA Sequencing

[0247] Tables 3 and 4 below show additional illustrative primers for barcoding of target nucleic acids in preparation for 454 DNA sequencing. "454F" refers to a 454 forward primer binding site; "454R" refers to 454 reverse primer binding site. "BC" refers to a nucleotide barcode. "TAG" refers to a nucleotide tag. "P53" refers to a target-specific primer sequence.

Table 3

Sequence Name	Sequence	SEQ ID
454F-BC1-TAG8	GCCTCCCTCGCGCCATCAGGCATGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:19)
454F-BC2-TAG8	GCCTCCCTCGCGCCATCAGCGTACGACACTGACGACA TGGTTCTACA	(SEQ ID NO:20)
454F-BC3-TAG8	GCCTCCCTCGCGCCATCAGGTCAGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:21)
454F-BC4-TAG8	GCCTCCCTCGCGCCATCAGAGCTGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:22)
454F-BC5-TAG8	GCCTCCCTCGCGCCATCAGTGCATCACACTGACGACA TGGTTCTACA	(SEQ ID NO:23)
454F-BC6-TAG8	GCCTCCCTCGCGCCATCAGCTGATGACACTGACGACA TGGTTCTACA	(SEQ ID NO:24)
454F-BC7-TAG8	GCCTCCCTCGCGCCATCAGGTAGTCACACTGACGACA TGGTTCTACA	(SEQ ID NO:25)

IDT Exhibit 1020.2706

454F-BC8-TAG8	GCCTCCCTCGCGCCATCAGGTCGATACACTGACGACA TGGTTCTACA	(SEQ ID NO:26)
454F-BC9-TAG8	GCCTCCCTCGCGCCATCAGGATACGACACTGACGACA TGGTTCTACA	(SEQ ID NO:27)
454F-BC10-TAG8	GCCTCCCTCGCGCCATCAGTGATGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:28)
454F-BC11-TAG8	GCCTCCCTCGCGCCATCAGAGCTGAACACTGACGACA TGGTTCTACA	(SEQ ID NO:29)
454F-BC12-TAG8	GCCTCCCTCGCGCCATCAGACTGTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:30)
454F-BC13-TAG8	GCCTCCCTCGCGCCATCAGTGCATGACACTGACGACA TGGTTCTACA	(SEQ ID NO:31)
454F-BC14-TAG8	GCCTCCCTCGCGCCATCAGAGTCTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:32)
454F-BC15-TAG8	GCCTCCCTCGCGCCATCAGTGTCTGACACTGACGACA TGGTTCTACA	(SEQ ID NO:33)
454F-BC16-TAG8	GCCTCCCTCGCGCCATCAGGCTAGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:34)
454F-BC17-TAG8	GCCTCCCTCGCGCCATCAGGATAGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:35)
454F-BC18-TAG8	GCCTCCCTCGCGCCATCAGGCTACTACACTGACGACA TGGTTCTACA	(SEQ ID NO:36)
454F-BC19-TAG8	GCCTCCCTCGCGCCATCAGCTATGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:37)
454F-BC20-TAG8	GCCTCCCTCGCGCCATCAGGCTATGACACTGACGACA TGGTTCTACA	(SEQ ID NO:38)
454F-BC21-TAG8	GCCTCCCTCGCGCCATCAGCGTGCAACACTGACGACA TGGTTCTACA	(SEQ ID NO:39)
454F-BC22-TAG8	GCCTCCCTCGCGCCATCAGATAGCTACACTGACGACA TGGTTCTACA	(SEQ ID NO:40)
454F-BC23-TAG8	GCCTCCCTCGCGCCATCAGTGTAGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:41)
454F-BC24-TAG8	GCCTCCCTCGCGCCATCAGGTGCTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:42)
454F-BC25-TAG8	GCCTCCCTCGCGCCATCAGGTCATGACACTGACGACA TGGTTCTACA	(SEQ ID NO:43)
454F-BC26-TAG8	GCCTCCCTCGCGCCATCAGATCGTGACACTGACGACA TGGTTCTACA	(SEQ ID NO:44)
454F-BC27-TAG8	GCCTCCCTCGCGCCATCAGTGTACGACACTGACGACA TGGTTCTACA	(SEQ ID NO:45)
454F-BC28-TAG8	GCCTCCCTCGCGCCATCAGAGTGTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:46)
454F-BC29-TAG8	GCCTCCCTCGCGCCATCAGTGACAGACACTGACGACA TGGTTCTACA	(SEQ ID NO:47)
454F-BC30-TAG8	GCCTCCCTCGCGCCATCAGGATCACACACTGACGACA TGGTTCTACA	(SEQ ID NO:48)
454F-BC31-TAG8	GCCTCCCTCGCGCCATCAGCTAGAGACACTGACGACA TGGTTCTACA	(SEQ ID NO:49)
454F-BC32-TAG8	GCCTCCCTCGCGCCATCAGCTAGTCACACTGACGACA TGGTTCTACA	(SEQ ID NO:50)
454F-BC33-TAG8	GCCTCCCTCGCGCCATCAGAGCTAGACACTGACGACA TGGTTCTACA	(SEQ ID NO:51)
454F-BC34-TAG8	GCCTCCCTCGCGCCATCAGTGACTGACACTGACGACA TGGTTCTACA	(SEQ ID NO:52)
454F-BC35-TAG8	GCCTCCCTCGCGCCATCAGTGATAGACACTGACGACA TGGTTCTACA	(SEQ ID NO:53)
454F-BC36-TAG8	GCCTCCCTCGCGCCATCAGCGTATCACACTGACGACA TGGTTCTACA	(SEQ ID NO:54)
454F-BC37-TAG8	GCCTCCCTCGCGCCATCAGGTCTGAACACTGACGACA	(SEQ ID NO:55)

IDT Exhibit 1020.2707

	TGGTTCTACA	
454F-BC38-TAG8	GCCTCCCTCGCGCCATCAGCATGACACACTGACGACA TGGTTCTACA	(SEQ ID NO:56)
454F-BC39-TAG8	GCCTCCCTCGCGCCATCAGCGATGAACACTGACGACA TGGTTCTACA	(SEQ ID NO:57)
454F-BC40-TAG8	GCCTCCCTCGCGCCATCAGGCTGATACACTGACGACA TGGTTCTACA	(SEQ ID NO:58)
454F-BC41-TAG8	GCCTCCCTCGCGCCATCAGCAGTACACACTGACGACA TGGTTCTACA	(SEQ ID NO:59)
454F-BC42-TAG8	GCCTCCCTCGCGCCATCAGGCGACTACACTGACGACA TGGTTCTACA	(SEQ ID NO:60)
454F-BC43-TAG8	GCCTCCCTCGCGCCATCAGGTACGAACACTGACGACA TGGTTCTACA	(SEQ ID NO:61)
454F-BC44-TAG8	GCCTCCCTCGCGCCATCAGACGCTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:62)
454F-BC45-TAG8	GCCTCCCTCGCGCCATCAGAGCATCACACTGACGACA TGGTTCTACA	(SEQ ID NO:63)
454F-BC46-TAG8	GCCTCCCTCGCGCCATCAGGATGCTACACTGACGACA TGGTTCTACA	(SEQ ID NO:64)
454F-BC47-TAG8	GCCTCCCTCGCGCCATCAGGTCTGCACACTGACGACA TGGTTCTACA	(SEQ ID NO:65)
454F-BC48-TAG8	GCCTCCCTCGCGCCATCAGATGCGAACACTGACGACA TGGTTCTACA	(SEQ ID NO:66)

Table 4

Sequence Name	Sequence	SEQ ID
TAG8-P53-1+	ACACTGACGACATGGTTCTACAACTGTCCAGCTTTGT GCC	(SEQ ID NO:67)
TAG8-P53-2+	ACACTGACGACATGGTTCTACAGATCATCATAGGAGT TGCATTGTTG	(SEQ ID NO:68)
TAG8-P53-3+	ACACTGACGACATGGTTCTACACGGACCTTTGTCCTT CCT	(SEQ ID NO:69)
TAG8-P53-4+	ACACTGACGACATGGTTCTACAATGCAAACCTCAATC CCTCC	(SEQ ID NO:70)
TAG8-P53-5+	ACACTGACGACATGGTTCTACAAGTTTCTTCCCATGC ACCTG	(SEQ ID NO:71)
TAG8-P53-6+	ACACTGACGACATGGTTCTACAGTGAATCCCCGTCTC TACTAAAA	(SEQ ID NO:72)
TAG8-P53-7+	ACACTGACGACATGGTTCTACATGTTTCCCATTTGCG GTTATGA	(SEQ ID NO:73)
TAG8-P53-8+	ACACTGACGACATGGTTCTACAAGTTGTGGGACTGCT TTATACATT	(SEQ ID NO:74)
454R-P53-1-	GCCTTGCCAGCCCGCTCAGTCCTCTGCCTAGGCGTT	(SEQ ID NO:75)
454R-P53-2-	GCCTTGCCAGCCCGCTCAGGAAATGTAAATGTGGAGC CAAACA	(SEQ ID NO:76)
454R-P53-3-	GCCTTGCCAGCCCGCTCAGACTCATTCTTGAAAATAC CTCCGG	(SEQ ID NO:77)
454R-P53-4-	GCCTTGCCAGCCCGCTCAGAAATGCCACCTCGATTTA GGAAA	(SEQ ID NO:78)
454R-P53-5-	GCCTTGCCAGCCCGCTCAGTCACCCTCCCGAATAGCT	(SEQ ID NO:79)
454R-P53-6-	GCCTTGCCAGCCCGCTCAGAGTGTAAAATGGTACAAC CGCT	(SEQ ID NO:80)
454R-P53-7-	GCCTTGCCAGCCCGCTCAGCCTCTTAAGATACTGTAA ACTCTGTAAAGC	(SEQ ID NO:81)
454R-P53-8-	GCCTTGCCAGCCCGCTCAGATTGTGCCATTGTACTCT	(SEQ ID NO:82)

IDT Exhibit 1020.2708

	AGCC	
TAG8-P53-9+	ACACTGACGACATGGTTCTACACTTCCTTTCTCTACTG AATGCTTTTAATTT	(SEQ ID NO:83)
TAG8-P53-10+	ACACTGACGACATGGTTCTACATCTTACACAAACTCT TCAGAAAACAGA	(SEQ ID NO:84)
TAG8-P53-11+	ACACTGACGACATGGTTCTACAGTACCAAAACCAAA CAAGGACAT	(SEQ ID NO:85)
TAG8-P53-12+	ACACTGACGACATGGTTCTACAGGTGAAACGCCATCT CTACTAA	(SEQ ID NO:86)
TAG8-P53-13+	ACACTGACGACATGGTTCTACATCATGATTGTAGCTG ATTCAACATTCA	(SEQ ID NO:87)
TAG8-P53-14+	ACACTGACGACATGGTTCTACAACTAGCATGCTGAAA CCCC	(SEQ ID NO:88)
TAG8-P53-15+	ACACTGACGACATGGTTCTACATCAGGAGATCGAGA CCATCC	(SEQ ID NO:89)
TAG8-P53-16+	ACACTGACGACATGGTTCTACATCATGCCTGTAATCC CAGC	(SEQ ID NO:90)
454R-P53-9-	GCCTTGCCAGCCCGCTCAGACCTCAAATGATCCCCTG C	(SEQ ID NO:91)
454R-P53-10-	GCCTTGCCAGCCCGCTCAGATTACAGGCGTGAGCCAC	(SEQ ID NO:92)
454R-P53-11-	GCCTTGCCAGCCCGCTCAGTTTTGAGATGAAGTCTTG CTCTGT	(SEQ ID NO:93)
454R-P53-12-	GCCTTGCCAGCCCGCTCAGTAAAGACCAGTCTGACTA TGTTGC	(SEQ ID NO:94)
454R-P53-13-	GCCTTGCCAGCCCGCTCAGACCATGCCCGGCTAATTT T	(SEQ ID NO:95)
454R-P53-14-	GCCTTGCCAGCCCGCTCAGAGTTCACGCCATTCTCCT G	(SEQ ID NO:96)
454R-P53-15-	GCCTTGCCAGCCCGCTCAGCACTACGCCCGGCTAATT TT	(SEQ ID NO:97)
454R-P53-16-	GCCTTGCCAGCCCGCTCAGTGGCCCCATTAGGACATG TAT	(SEQ ID NO:98)
TAG8-P53-17+	ACACTGACGACATGGTTCTACATTGTCCCATTGCACT CCAG	(SEQ ID NO:99)
TAG8-P53-18+	ACACTGACGACATGGTTCTACATGGGCAACAAGAGT GAAACT	(SEQ ID NO:100)
TAG8-P53-19+	ACACTGACGACATGGTTCTACAAAATAAATATAGCA GGGTTGCAGGT	(SEQ ID NO:101)
TAG8-P53-20+	ACACTGACGACATGGTTCTACATGCATTTCTCTTGGC TCCC	(SEQ ID NO:102)
TAG8-P53-21+	ACACTGACGACATGGTTCTACAACTTTCCTCAACTCT ACATTTCCC	(SEQ ID NO:103)
TAG8-P53-22+	ACACTGACGACATGGTTCTACATCAGTGCAAACAACA GAAAAGTG	(SEQ ID NO:104)
TAG8-P53-23+	ACACTGACGACATGGTTCTACACATGTTTCTTAGCAA ATCTGATGACA	(SEQ ID NO:105)
TAG8-P53-24+	ACACTGACGACATGGTTCTACATCTGTGGTCCCAGCT ACT	(SEQ ID NO:106)
454R-P53-17-	GCCTTGCCAGCCCGCTCAGTTTCACCATGTTAGGTTG GTCTC	(SEQ ID NO:107)
454R-P53-18-	GCCTTGCCAGCCCGCTCAGTGTAGGTTAAATCCAAAT ACTATACCGTC	(SEQ ID NO:108)
454R-P53-19-	GCCTTGCCAGCCCGCTCAGTCTCAAATCTTCAGTAGC AACTAAAATCT	(SEQ ID NO:109)
454R-P53-20-	GCCTTGCCAGCCCGCTCAGTCCCGACCTCAGGTGATC	(SEQ ID NO:110)
454R-P53-21-	GCCTTGCCAGCCCGCTCAGTGGTCTTGAACTCCCAAC TTC	(SEQ ID NO:111)
454R-P53-22-	GCCTTGCCAGCCCGCTCAGCCTCCGACTCCCAAAGTG	(SEQ ID NO:112)

IDT Exhibit 1020.2709

454R-P53-23-	GCCTTGCCAGCCCGCTCAGACTACAGCCTCGGACTCC	(SEQ ID NO:113)
454R-P53-24-	GCCTTGCCAGCCCGCTCAGATCTTGCACGAAGTTATG CAACTA	(SEQ ID NO:114)
TAG8-P53-25+	ACACTGACGACATGGTTCTACAACCACTGCACTCCAG C	(SEQ ID NO:115)
TAG8-P53-26+	ACACTGACGACATGGTTCTACAACAAGGAAAAGTAT CAGACAATGTAAGT	(SEQ ID NO:116)
TAG8-P53-27+	ACACTGACGACATGGTTCTACAACGGTAGCTCACACC TGTAAT	(SEQ ID NO:117)
TAG8-P53-28+	ACACTGACGACATGGTTCTACATGGAAGTCCCTCTCT GATTGT	(SEQ ID NO:118)
TAG8-P53-29+	ACACTGACGACATGGTTCTACAACTGACTTTCTGCTC TTGTCTTTC	(SEQ ID NO:119)
TAG8-P53-30+	ACACTGACGACATGGTTCTACAATTCTGGGACAGCCA AGTC	(SEQ ID NO:120)
TAG8-P53-31+	ACACTGACGACATGGTTCTACAAGGAGTTCAAGACC AGCCT	(SEQ ID NO:121)
TAG8-P53-32+	ACACTGACGACATGGTTCTACATCTGTCTCCTTCCTCT TCCTAC	(SEQ ID NO:122)
454R-P53-25-	GCCTTGCCAGCCCGCTCAGCCTCTTCCCCAAAAGCTC T	(SEQ ID NO:123)
454R-P53-26-	GCCTTGCCAGCCCGCTCAGTCTCGAACTCCTTACTTC AGGT	(SEQ ID NO:124)
454R-P53-27-	GCCTTGCCAGCCCGCTCAGCCCAACACCATGCCAGTG	(SEQ ID NO:125)
454R-P53-28-	GCCTTGCCAGCCCGCTCAGTCCCCAGCCCTCCAG	(SEQ ID NO:126)
454R-P53-29-	GCCTTGCCAGCCCGCTCAGATTGAAGTCTCATGGAAG CCAG	(SEQ ID NO:127)
454R-P53-30-	GCCTTGCCAGCCCGCTCAGTCAAGTGATCTTCCCACC TCA	(SEQ ID NO:128)
454R-P53-31-	GCCTTGCCAGCCCGCTCAGACAACCTCCGTCATGTGC	(SEQ ID NO:129)
454R-P53-32-	GCCTTGCCAGCCCGCTCAGACCCATTTACTTTGCACA TCTCA	(SEQ ID NO:130)
TAG8-P53-33+	ACACTGACGACATGGTTCTACATTAAGGGTGGTTGTC AGTGG	(SEQ ID NO:131)
TAG8-P53-34+	ACACTGACGACATGGTTCTACATTGCAGTGAGCTGAG ATCAC	(SEQ ID NO:132)
TAG8-P53-35+	ACACTGACGACATGGTTCTACAATCTCCTTACTGCTC CCACT	(SEQ ID NO:133)
TAG8-P53-36+	ACACTGACGACATGGTTCTACATTTTATCACCTTTCCT TGCCTCTT	(SEQ ID NO:134)
TAG8-P53-37+	ACACTGACGACATGGTTCTACAACTCGTCGTAAGTTG AAAATATTGTAAGT	(SEQ ID NO:135)
TAG8-P53-38+	ACACTGACGACATGGTTCTACATCCCAAAGTGCTGGG ATTAC	(SEQ ID NO:136)
TAG8-P53-39+	ACACTGACGACATGGTTCTACATCCATCCTCCCAGCT CAG	(SEQ ID NO:137)
TAG8-P53-40+	ACACTGACGACATGGTTCTACAATCTCAGCTCACTGC AGC	(SEQ ID NO:138)
454R-P53-33-	GCCTTGCCAGCCCGCTCAGAGCCAACCTAGGAGATA ACACA	(SEQ ID NO:139)
454R-P53-34-	GCCTTGCCAGCCCGCTCAGAGGCTCCATCTACTCCCA A	(SEQ ID NO:140)
454R-P53-35-	GCCTTGCCAGCCCGCTCAGTTGATAAGAGGTCCCAAG ACTTAGTA	(SEQ ID NO:141)
454R-P53-36-	GCCTTGCCAGCCCGCTCAGTGGGTGACAGAGTGAGA CT	(SEQ ID NO:142)
454R-P53-37-	GCCTTGCCAGCCCGCTCAGACATCACTGTAATCCAGC CTG	(SEQ ID NO:143)

IDT Exhibit 1020.2710

454R-P53-38-	GCCTTGCCAGCCCGCTCAGAGATCATGCCACTGCACT C	(SEQ ID NO:144)
454R-P53-39-	GCCTTGCCAGCCCGCTCAGGGCATGTGCCTGTAGTCC	(SEQ ID NO:145)
454R-P53-40-	GCCTTGCCAGCCCGCTCAGTGGTCTTGAACTCCTGAC CT	(SEQ ID NO:146)
TAG8-P53-41+	ACACTGACGACATGGTTCTACAAAACAGCATGGTTGC ATGAAAG	(SEQ ID NO:147)
TAG8-P53-42+	ACACTGACGACATGGTTCTACAAGTCGCATGCACATG TAGTC	(SEQ ID NO:148)
TAG8-P53-43+	ACACTGACGACATGGTTCTACAAAAAGTCAGCTGTAT AGGTACTTGAAG	(SEQ ID NO:149)
TAG8-P53-44+	ACACTGACGACATGGTTCTACACCTCAGTGTATCCAC AGAACA	(SEQ ID NO:150)
TAG8-P53-45+	ACACTGACGACATGGTTCTACAATGCATGCCTGTAAT CCCAG	(SEQ ID NO:151)
TAG8-P53-46+	ACACTGACGACATGGTTCTACAAACTCATGTTCAAGA CAGAAGGG	(SEQ ID NO:152)
TAG8-P53-47+	ACACTGACGACATGGTTCTACAATTTTCTCTAACTTC AAGGCCCATAT	(SEQ ID NO:153)
TAG8-P53-48+	ACACTGACGACATGGTTCTACATGGATCCACCAAGAC TTGTTTTAT	(SEQ ID NO:154)
454R-P53-41-	GCCTTGCCAGCCCGCTCAGGATTACAGGTGTGAGCCA CT	(SEQ ID NO:155)
454R-P53-42-	GCCTTGCCAGCCCGCTCAGACAGTACCTGAGTTAAAA GATGGTTC	(SEQ ID NO:156)
454R-P53-43-	GCCTTGCCAGCCCGCTCAGTGAGACCCTCCAGCTCTG	(SEQ ID NO:157)
454R-P53-44-	GCCTTGCCAGCCCGCTCAGATCTTCCCTTACCCCATTT TACTTTATT	(SEQ ID NO:158)
454R-P53-45-	GCCTTGCCAGCCCGCTCAGTTCAAAGACCCAAAACCC AAAATG	(SEQ ID NO:159)
454R-P53-46-	GCCTTGCCAGCCCGCTCAGGTCAAGTTCTAGACCCCA TGTAATA	(SEQ ID NO:160)
454R-P53-47-	GCCTTGCCAGCCCGCTCAGTGTGGTCCCAGCTACTCC	(SEQ ID NO:161)
454R-P53-48-	GCCTTGCCAGCCCGCTCAGAGCAAAGTTTTATTGTAA AATAAGAGATCGAT	(SEQ ID NO:162)

Example 4

4-Primer Barcoding Of Target Nucleic Acids in Preparation for 454 DNA Sequencing Using a Microfluidic Device that Permits Recovery of Amplication Products

[0248] Target-specific primers were designed for 48 genomic regions associated with prostate cancer. In addition to the target-specific regions, the primers were designed to contain additional tag sequences at the 5' end. Forward primers contained the sequence ACACTGACGACATGGTTCTACA (SEQ ID NO:163). Reverse primers contained the sequence TACGGTAGCAGAGACTTGGTCT (SEQ ID NO:164). The sequences of the primers containing both tag sequences and the target-specific regions are listed in Table 5.

IDT Exhibit 1020.2711

Table 5

Assay #	Assay Name	Tagged Forward primer sequence	Tagged Reverse primer sequence	Amplicon position	Amplicon size (no tags)	SEQ ID NO.
1	MSMB-1	ACACTGACGACATGGTTCTACAGTGGTTGCCCTCTCCAGTA	TACGGTAGCAGAGACTTGGTCTGCACACGCATATTAAAATAGGAA	chr10:51219512+51219668	157	(SEQ ID NO:165)
2	MSMB-2	ACACTGACGACATGGTTCTACATCATTCTCCACCCTGACCTT	TACGGTAGCAGAGACTTGGTCTTTCATCTGCAGACAGGTCCA	chr10:51225703+51225910	208	(SEQ ID NO:166)
3	MSMB-3	ACACTGACGACATGGTTCTACAAGGCCTTGTTCTCATTGCAT	TACGGTAGCAGAGACTTGGTCTCCAGCACTGGCTTGAGACTT	chr10:51226702+51226884	183	(SEQ ID NO:167)
4	MSMB-4	ACACTGACGACATGGTTCTACAGGGTCCTTTCTCTTCTAACAGG	TACGGTAGCAGAGACTTGGTCTAGGCCAGAGGAGAATGAGG	chr10:51232232+51232460	229	(SEQ ID NO:168)
5	HNF1B-1	ACACTGACGACATGGTTCTACACAGAGGGTGATGGTGTGGA	TACGGTAGCAGAGACTTGGTCTATGACCCTGCCAAATGACAC	chr17:33121423+33121560	138	(SEQ ID NO:169)
6	HNF1B-5	ACACTGACGACATGGTTCTACATGCTTCCCATTCTTCTTCTCC	TACGGTAGCAGAGACTTGGTCTTGGAAACTGCTCTTTGTGGTC	chr17:33138980+33139231	252	(SEQ ID NO:170)
7	HNF1B-6	ACACTGACGACATGGTTCTACATGCCTCTTATCTTATCAGCTCCA	TACGGTAGCAGAGACTTGGTCTTGGTGGCACTAATGTTCCCTA	chr17:33144574+33144827	254	(SEQ ID NO:171)
8	HNF1B-7	ACACTGACGACATGGTTCTACATAAGATCCGTGGCAAGAACC	TACGGTAGCAGAGACTTGGTCTGAGGTCCGTGTCTACAACTGG	chr17:33165634+33165867	234	(SEQ ID NO:172)
9	HNF1B-8	ACACTGACGACATGGTTCTACAGTCCATGGCCAGCTTTTG	TACGGTAGCAGAGACTTGGTCTCCCCTCACTCACCATCTCC	chr17:33165796+33165970	175	(SEQ ID NO:173)
10	HNF1B-9	ACACTGACGACATGGTTCTACAGGGTTCCTGGGTCTGTGTA	TACGGTAGCAGAGACTTGGTCTAGTCCGATGATGCCTGCT	chr17:33167605+33167819	215	(SEQ ID NO:174)
11	HNF1B-10	ACACTGACGACATGGTTCTACACTTCTTGTTGGTGGGCTCAG	TACGGTAGCAGAGACTTGGTCTTGAGTGAAGGCTACAGACCCTA	chr17:33167782+33167975	194	(SEQ ID NO:175)

IDT Exhibit 1020.2712

12	HNF1B-11	ACACTGACGACATGGTTCTACATGAGAGGGCAAAGGTCACTT	TACGGTAGCAGAGACTTGGTCTAGAGGGAGGTGGTCGATGT	chr17:33173490+33173681	192	(SEQ ID NO:176)
13	HNF1B-12	ACACTGACGACATGGTTCTACAGTTGAGATGCTGGGAGAGGT	TACGGTAGCAGAGACTTGGTCTTCTCCCACTAGTACCCTAACCATC	chr17:33173623+33173782	160	(SEQ ID NO:177)
14	MYC-1	ACACTGACGACATGGTTCTACAGACCCGCTTCTCTGAAAGG	TACGGTAGCAGAGACTTGGTCTGCATTCGACTCATCTCAGCA	chr8:128817980+128818121	142	(SEQ ID NO:178)
15	MYC-2	ACACTGACGACATGGTTCTACACAGGTTTCCGCACCAAGA	TACGGTAGCAGAGACTTGGTCTCAGCAGCTCGAATTTCTTCC	chr8:128819612+128819858	247	(SEQ ID NO:179)
16	MYC-6	ACACTGACGACATGGTTCTACAAACCTTGCTAAAGGAGTGATTTCT	TACGGTAGCAGAGACTTGGTCTCCTCTTGGCAGCAGGATAGT	chr8:128821784+128822038	255	(SEQ ID NO:180)
17	MYC-7	ACACTGACGACATGGTTCTACAACGTCTCCACACATCAGCAC	TACGGTAGCAGAGACTTGGTCTAACTCCGGGATCTGGTCAC	chr8:128821968+128822217	250	(SEQ ID NO:181)
18	MYC-8	ACACTGACGACATGGTTCTACACCAGAGGAGGAACGAGCTAA	TACGGTAGCAGAGACTTGGTCTTTCTGTTAGAAGGAATCGTTTTCC	chr8:128822158+128822420	263	(SEQ ID NO:182)
19	JAZF1-2	ACACTGACGACATGGTTCTACATTCCATGTGGTTATGCCAAG	TACGGTAGCAGAGACTTGGTCTCTCCTGACAGTCCTTGCACTT	chr7:27846803+27847046	244	(SEQ ID NO:183)
20	JAZF1-4	ACACTGACGACATGGTTCTACACAATAAGCAGCAGATATAAGGTTGTT	TACGGTAGCAGAGACTTGGTCTCTTTGTGTTAGGTAGCCTCATATATTC	chr7:27998002+27998196	195	(SEQ ID NO:184)
21	NCOA4-1	ACACTGACGACATGGTTCTACATTCAAAGGTGGTTTTTGGTTG	TACGGTAGCAGAGACTTGGTCTGCCCTGTGTCAAGAGTCCAG	chr10:51249073+51249337	265	(SEQ ID NO:185)
22	NCOA4-2	ACACTGACGACATGGTTCTACATTGGGAAACATCATTCTTTGG	TACGGTAGCAGAGACTTGGTCTACCAGAAGCCATGCTCAAAC	chr10:51250503+51250748	246	(SEQ ID NO:186)
23	NCOA4-3	ACACTGACGACATGGTTCTACATGGTGTCATTGTGGCTAGTTG	TACGGTAGCAGAGACTTGGTCTTGATCTTATCCTAGCAACACAGAAG	chr10:51250847+51251096	250	(SEQ ID NO:187)

IDT Exhibit 1020.2713

24	NCOA4-4	ACACTGACGACATGGTTCTACATGAAGTTGATGAAACAGATATTCCTT	TACGGTAGCAGAGACTTGGTCTAGAAGTGCCCAGTGAAGCAT	chr10:51251218+51251418	201	(SEQ ID NO:188)
25	NCOA4-5	ACACTGACGACATGGTTCTACATTGGCAGCATAGCATAAATAACA	TACGGTAGCAGAGACTTGGTCTCCCAAAGGAAGTATAAGCCAAG	chr10:51252141+51252337	197	(SEQ ID NO:189)
26	NCOA4-6	ACACTGACGACATGGTTCTACACTGCATTTGACATTCCTTGTTT	TACGGTAGCAGAGACTTGGTCTTCCACCTACTGCTGTGTCTACTG	chr10:51252768+51252994	227	(SEQ ID NO:190)
27	NCOA4-7	ACACTGACGACATGGTTCTACAGCAGACAGAATCTCCAAAGCA	TACGGTAGCAGAGACTTGGTCTTCTGATAGGTCCATCTCATCTTGA	chr10:51254556+51254815	260	(SEQ ID NO:191)
28	NCOA4-8	ACACTGACGACATGGTTCTACAGGTTGGAGATCAAGAGCTTCCT	TACGGTAGCAGAGACTTGGTCTTGGTCATTCAGGCACTTCAG	chr10:51254768+51255022	255	(SEQ ID NO:192)
29	NCOA4-9	ACACTGACGACATGGTTCTACAGAAACCAGCCCAAAGGTGT	TACGGTAGCAGAGACTTGGTCTCCTTCTTTCTTCAGAAGCCACT	chr10:51254962+51255214	253	(SEQ ID NO:193)
30	NCOA4-10	ACACTGACGACATGGTTCTACAGAATTGTGAGAAGGAGGCTCTG	TACGGTAGCAGAGACTTGGTCTTGGGACTTCCTTCTTTGTATGG	chr10:51255167+51255432	266	(SEQ ID NO:194)
31	NCOA4-11	ACACTGACGACATGGTTCTACACCTTGTCGGAGTGGCTTATC	TACGGTAGCAGAGACTTGGTCTCCAGTGCTATTTTGATGTTTATGC	chr10:51255385+51255633	249	(SEQ ID NO:195)
32	NCOA4-13	ACACTGACGACATGGTTCTACAGGAGCTTTAAGGCAGGGAAA	TACGGTAGCAGAGACTTGGTCTTTGGCAAGCTGCAGTCAC	chr10:51259156+51259310	155	(SEQ ID NO:196)
33	NUDT11-1	ACACTGACGACATGGTTCTACAAGCGAGGCAGACAAATAGAAG	TACGGTAGCAGAGACTTGGTCTGTACTGACTGTCACGGAGCTG	chrX:51255496+51255748	253	(SEQ ID NO:197)
34	SLC22A3-4	ACACTGACGACATGGTTCTACATCTGCATTCTGGCATGTCTC	TACGGTAGCAGAGACTTGGTCTTCCCCGTATTAATGCATGGTAT	chr6:160738955+160739163	209	(SEQ ID NO:198)
35	SLC22A3-5	ACACTGACGACATGGTTCTACAAGGTGAGCTCTTTTCCTGTCTT	TACGGTAGCAGAGACTTGGTCTTTGTTGGCTATCTGGCCCTA	chr6:160748030+160748274	245	(SEQ ID NO:199)

IDT Exhibit 1020.2714

36	SLC22A3-6	ACACTGACGACATGGTTCTACATGCTTCTGTGACCTCTTGTGT	TACGGTAGCAGAGACTTGGTCTGTCTGTTTGGAGTCTAATTTCTGC	chr6:160749740+160750007	268	(SEQ ID NO:200)
37	SLC22A3-7	ACACTGACGACATGGTTCTACACATAACTCACAACAGCCTCCTTC	TACGGTAGCAGAGACTTGGTCTAATCAATTCACCAGCTTTAGCAA	chr6:160751720+160751920	201	(SEQ ID NO:201)
38	SLC22A3-10	ACACTGACGACATGGTTCTACAGTGGTGGAACTGCCAGGA	TACGGTAGCAGAGACTTGGTCTGGCTCCCTATACTTGATTGTGG	chr6:160778107+160778308	202	(SEQ ID NO:202)
39	SLC22A3-11	ACACTGACGACATGGTTCTACACCTCCCTTTCAAACTTTCTGTG	TACGGTAGCAGAGACTTGGTCTCGCTGGTCTACAGAGTTACTTAGGA	chr6:160783754+160783942	189	(SEQ ID NO:203)
40	SLC22A3-12	ACACTGACGACATGGTTCTACATGATTATCTTGAAGTCACTTGTTGAA	TACGGTAGCAGAGACTTGGTCTTGAAGGCTCTTAAGAATAGCAAATG	chr6:160784591+160784798	208	(SEQ ID NO:204)
41	SLC22A3-13	ACACTGACGACATGGTTCTACAGTGTCTTCCTGGAGCGGTAA	TACGGTAGCAGAGACTTGGTCTTTCCCTGTGGATATTCAATTTTTCT	chr6:160788700+160788934	235	(SEQ ID NO:205)
42	SLC22A3-14	ACACTGACGACATGGTTCTACATCTTTCCTAAAGACTTTCTCCTTTG	TACGGTAGCAGAGACTTGGTCTATCTCTGCAAGGCACAGCTT	chr6:160791984+160792152	169	(SEQ ID NO:206)
43	KLK3-1	ACACTGACGACATGGTTCTACAAGTCCTGGGGAATGAAGGTT	TACGGTAGCAGAGACTTGGTCTGGAAAGAGCCTCAGCTTGAC	chr19:56049936+56050140	205	(SEQ ID NO:207)
44	KLK3-2	ACACTGACGACATGGTTCTACAGTTCCTCCTGTCAACCCTGA	TACGGTAGCAGAGACTTGGTCTCCTCTGGGACACAGACACCT	chr19:56051260+56051515	256	(SEQ ID NO:208)
45	KLK3-3	ACACTGACGACATGGTTCTACATCCTTATCATCCTCGCTCCT	TACGGTAGCAGAGACTTGGTCTTTCACAGCATCCGTGAGC	chr19:56053051+56053300	250	(SEQ ID NO:209)
46	KLK3-4	ACACTGACGACATGGTTCTACAACTCCAGCCACGACCTCAT	TACGGTAGCAGAGACTTGGTCTCCCTCAGACCCAGGCATC	chr19:56053237+56053436	200	(SEQ ID NO:210)
47	KLK3-5	ACACTGACGACATGGTTCTACAGGTCCAGCCCACAACAGT	TACGGTAGCAGAGACTTGGTCTCCCAGCCCAGAATTAAGGT	chr19:56053490+56053729	240	(SEQ ID NO:211)
48	KLK3-8	ACACTGACGACATGGTTCTACA	TACGGTAGCAGAGACTTGGT	chr19:56054924+56055115	192	(SEQ ID NO:212)

IDT Exhibit 1020.2715

		TCTTCCAAAGC TGGGAACTG	CTGGGCACAT GGTTCACTGC			

Preparation of Reaction Mixtures

[0249] Primers were synthesized by IDT at 10nmol scale, and provided resuspended in water at a concentration of 100uM. The forward and reverse primer for each region in Table 5 were combined in separate wells in a 96-well PCR plate (USA scientific) to a final concentration of 1μM of each primer in PCR-quality water (Teknova) containing 0.05% Tween-20.

[0250] 48 human genomic DNA samples from the HapMap sample collection were resuspended at 50ng/μl in low-EDTA TE buffer (Teknova), and prepared for PCR as follows.

[0251] A pre-sample mixture was prepared as follows:

Table 6

Pre-sample mixture	**Volume per sample (μl)**	**Volume for 64 samples (μl)**
Faststart High Fidelity reaction Buffer with $MgCl_2$	0.5	32
DMSO	0.1	6.4
PCR-Grade Nucleotide Mixture	0.1	6.4
Faststart High-Fidelity Enzyme Blend (Roche 04 738 292 001)	0.05	3.2
20x Access Array Loading Reagent (PN: 100-0883)	0.25	16
20x Evagreen (Biotium-31000)	0.25	16
20x ROX dye (Invitrogen 12223-012)	0.25	16
PCR-Grade water	0.5	32
Total	2	128

IDT Exhibit 1020.2716

[0252] For each sample, a sample mixture containing forward and reverse barcode primers, genomic DNA, and pre-sample mix was prepared in an individual well in a 96-well PCR plate.

Table 7

Sample Mixture	Volume (μl)
Pre-sample Mixture	2
2 μM forward barcode primer	0.5
2 μm reverse barcode primer	0.5
Genomic DNA (50ng/μl)	1
PCR-grade water	1

[0253] Each sample was mixed with one pair of barcode primers selected from Table 8.

Table 8

	Reverse barcode primer (454B-BC#-CS1)	Reverse barcode primer SEQ ID NO.	Forward barcode primer (454A-BC#-CS2)	Forward barcode primer SEQ ID NO.
1	GCCTTGCCAGCCCGCTCAGGCATGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:213)	GCCTCCCTCGCGCCATCAGGCATGCACACTGACGACATGGTTCTACA	(SEQ ID NO:214)
2	GCCTTGCCAGCCCGCTCAGCGTACGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:215)	GCCTCCCTCGCGCCATCAGCGTACGACACTGACGACATGGTTCTACA	(SEQ ID NO:216)
3	GCCTTGCCAGCCCGCTCAGGTCAGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:217)	GCCTCCCTCGCGCCATCAGGTCAGCACACTGACGACATGGTTCTACA	(SEQ ID NO:218)
4	GCCTTGCCAGCCCGCTCAGAGCTGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:219)	GCCTCCCTCGCGCCATCAGAGCTGCACACTGACGACATGGTTCTACA	(SEQ ID NO:220)
5	GCCTTGCCAGCCCGCTCAGTGCATCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:221)	GCCTCCCTCGCGCCATCAGTGCATCACACTGACGACATGGTTCTACA	(SEQ ID NO:222)
6	GCCTTGCCAGCCCGCTCAGCTGATGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:223)	GCCTCCCTCGCGCCATCAGCTGATGACACTGACGACATGGTTCTACA	(SEQ ID NO:224)
7	GCCTTGCCAGCCCGCTCAGGTAGTCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:225)	GCCTCCCTCGCGCCATCAGGTAGTCACACTGACGACATGGTTCTACA	(SEQ ID NO:226)

IDT Exhibit 1020.2717

8	GCCTTGCCAGCCCGCTCAGGTCGATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:227)	GCCTCCCTCGCGCCATCAGGTCGATACACTGACGACATGGTTCTACA	(SEQ ID NO:228)
9	GCCTTGCCAGCCCGCTCAGGATACGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:229)	GCCTCCCTCGCGCCATCAGGATACGACACTGACGACATGGTTCTACA	(SEQ ID NO:230)
10	GCCTTGCCAGCCCGCTCAGTGATGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:231)	GCCTCCCTCGCGCCATCAGTGATGCACACTGACGACATGGTTCTACA	(SEQ ID NO:232)
11	GCCTTGCCAGCCCGCTCAGAGCTGATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:233)	GCCTCCCTCGCGCCATCAGAGCTGAACACTGACGACATGGTTCTACA	(SEQ ID NO:234)
12	GCCTTGCCAGCCCGCTCAGACTGTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:235)	GCCTCCCTCGCGCCATCAGACTGTAACACTGACGACATGGTTCTACA	(SEQ ID NO:236)
13	GCCTTGCCAGCCCGCTCAGTGCATGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:237)	GCCTCCCTCGCGCCATCAGTGCATGACACTGACGACATGGTTCTACA	(SEQ ID NO:238)
14	GCCTTGCCAGCCCGCTCAGAGTCTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:239)	GCCTCCCTCGCGCCATCAGAGTCTAACACTGACGACATGGTTCTACA	(SEQ ID NO:240)
15	GCCTTGCCAGCCCGCTCAGTGTCTGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:241)	GCCTCCCTCGCGCCATCAGTGTCTGACACTGACGACATGGTTCTACA	(SEQ ID NO:242)
16	GCCTTGCCAGCCCGCTCAGGCTAGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:243)	GCCTCCCTCGCGCCATCAGGCTAGCACACTGACGACATGGTTCTACA	(SEQ ID NO:244)
17	GCCTTGCCAGCCCGCTCAGGATAGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:245)	GCCTCCCTCGCGCCATCAGGATAGCACACTGACGACATGGTTCTACA	(SEQ ID NO:246)
18	GCCTTGCCAGCCCGCTCAGGCTACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:247)	GCCTCCCTCGCGCCATCAGGCTACTACACTGACGACATGGTTCTACA	(SEQ ID NO:248)
19	GCCTTGCCAGCCCGCTCAGCTATGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:249)	GCCTCCCTCGCGCCATCAGCTATGCACACTGACGACATGGTTCTACA	(SEQ ID NO:250)
20	GCCTTGCCAGCCCGCTCAGGCTATGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:251)	GCCTCCCTCGCGCCATCAGGCTATGACACTGACGACATGGTTCTACA	(SEQ ID NO:252)
21	GCCTTGCCAGCCCGCTCAGCGTGCATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:253)	GCCTCCCTCGCGCCATCAGCGTGCAACACTGACGACATGGTTCTACA	(SEQ ID NO:254)
22	GCCTTGCCAGCCCGCTCAGATAGCTTACGGTAGCAGAGACTTGGT	(SEQ ID NO:255)	GCCTCCCTCGCGCCATCAGATAGCTACACTGACGACATGGTTCTACA	(SEQ ID NO:256)

IDT Exhibit 1020.2718

	CT			
23	GCCTTGCCAGCCCGC TCAGTGTAGCTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:257)	GCCTCCCTCGCGCCATCAGTGTA GCACACTGACGACATGGTTCTAC A	(SEQ ID NO:258)
24	GCCTTGCCAGCCCGC TCAGGTGCTATACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:259)	GCCTCCCTCGCGCCATCAGGTGC TAACACTGACGACATGGTTCTAC A	(SEQ ID NO:260)
25	GCCTTGCCAGCCCGC TCAGGTCATGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:261)	GCCTCCCTCGCGCCATCAGGTCA TGACACTGACGACATGGTTCTAC A	(SEQ ID NO:262)
26	GCCTTGCCAGCCCGC TCAGATCGTGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:263)	GCCTCCCTCGCGCCATCAGATCG TGACACTGACGACATGGTTCTAC A	(SEQ ID NO:264)
27	GCCTTGCCAGCCCGC TCAGTGTACGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:265)	GCCTCCCTCGCGCCATCAGTGTA CGACACTGACGACATGGTTCTAC A	(SEQ ID NO:266)
28	GCCTTGCCAGCCCGC TCAGAGTGTATACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:267)	GCCTCCCTCGCGCCATCAGAGTG TAACACTGACGACATGGTTCTAC A	(SEQ ID NO:268)
29	GCCTTGCCAGCCCGC TCAGTGACAGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:269)	GCCTCCCTCGCGCCATCAGTGAC AGACACTGACGACATGGTTCTAC A	(SEQ ID NO:270)
30	GCCTTGCCAGCCCGC TCAGGATCACTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:271)	GCCTCCCTCGCGCCATCAGGATC ACACACTGACGACATGGTTCTAC A	(SEQ ID NO:272)
31	GCCTTGCCAGCCCGC TCAGCTAGAGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:273)	GCCTCCCTCGCGCCATCAGCTAG AGACACTGACGACATGGTTCTAC A	(SEQ ID NO:274)
32	GCCTTGCCAGCCCGC TCAGCTAGTCTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:275)	GCCTCCCTCGCGCCATCAGCTAG TCACACTGACGACATGGTTCTAC A	(SEQ ID NO:276)
33	GCCTTGCCAGCCCGC TCAGAGCTAGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:277)	GCCTCCCTCGCGCCATCAGAGCT AGACACTGACGACATGGTTCTAC A	(SEQ ID NO:278)
34	GCCTTGCCAGCCCGC TCAGTGACTGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:279)	GCCTCCCTCGCGCCATCAGTGAC TGACACTGACGACATGGTTCTAC A	(SEQ ID NO:280)
35	GCCTTGCCAGCCCGC TCAGTGATAGTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:281)	GCCTCCCTCGCGCCATCAGTGAT AGACACTGACGACATGGTTCTAC A	(SEQ ID NO:282)
36	GCCTTGCCAGCCCGC TCAGCGTATCTACGG TAGCAGAGACTTGGT CT	(SEQ ID NO:283)	GCCTCCCTCGCGCCATCAGCGTA TCACACTGACGACATGGTTCTAC A	(SEQ ID NO:284)
37	GCCTTGCCAGCCCGC TCAGGTCTGATACGG	(SEQ ID NO:285)	GCCTCCCTCGCGCCATCAGGTCT GAACACTGACGACATGGTTCTAC	(SEQ ID NO:286)

IDT Exhibit 1020.2719

	TAGCAGAGACTTGGTCT		A	
38	GCCTTGCCAGCCCGCTCAGCATGACTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:287)	GCCTCCCTCGCGCCATCAGCATGACACACTGACGACATGGTTCTACA	(SEQ ID NO:288)
39	GCCTTGCCAGCCCGCTCAGCGATGATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:289)	GCCTCCCTCGCGCCATCAGCGATGAACACTGACGACATGGTTCTACA	(SEQ ID NO:290)
40	GCCTTGCCAGCCCGCTCAGGCTGATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:291)	GCCTCCCTCGCGCCATCAGGCTGATACACTGACGACATGGTTCTACA	(SEQ ID NO:292)
41	GCCTTGCCAGCCCGCTCAGCAGTACTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:293)	GCCTCCCTCGCGCCATCAGCAGTACACACTGACGACATGGTTCTACA	(SEQ ID NO:294)
42	GCCTTGCCAGCCCGCTCAGGCGACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:295)	GCCTCCCTCGCGCCATCAGGCGACTACACTGACGACATGGTTCTACA	(SEQ ID NO:296)
43	GCCTTGCCAGCCCGCTCAGGTACGATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:297)	GCCTCCCTCGCGCCATCAGGTACGAACACTGACGACATGGTTCTACA	(SEQ ID NO:298)
44	GCCTTGCCAGCCCGCTCAGACGCTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:299)	GCCTCCCTCGCGCCATCAGACGCTAACACTGACGACATGGTTCTACA	(SEQ ID NO:300)
45	GCCTTGCCAGCCCGCTCAGAGCATCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:301)	GCCTCCCTCGCGCCATCAGAGCATCACACTGACGACATGGTTCTACA	(SEQ ID NO:302)
46	GCCTTGCCAGCCCGCTCAGGATGCTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:303)	GCCTCCCTCGCGCCATCAGGATGCTACACTGACGACATGGTTCTACA	(SEQ ID NO:304)
47	GCCTTGCCAGCCCGCTCAGGTCTGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:305)	GCCTCCCTCGCGCCATCAGGTCTGCACACTGACGACATGGTTCTACA	(SEQ ID NO:306)
48	GCCTTGCCAGCCCGCTCAGATGCGATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:307)	GCCTCCCTCGCGCCATCAGATGCGAACACTGACGACATGGTTCTACA	(SEQ ID NO:308)

Running the Access Array IFC

[0254] The containment and interface accumulator reservoirs were filled with 300 µl of Control Line Fluid (Fluidigm PN 89000020) and the H1-H4 reagent wells were loaded with 500 µl of 0.05% Tween-20 in PCR-grade water prior to Access Array IFC loading. 5µl of each sample mixture was loaded into the sample ports, and

5 μl of each primer mixture was loaded into the primer inlets on the Access Array IFC.

[0255] The Access Array IFC was thermal cycled and imaged using a BioMark ™ Real-Time PCR system manufactured by Fluidigm Corporation. The Access Array IFC thermal cycling protocol contains a thermal mix step [50 °C for 2 min, 70 °C for 20 min], a hotstart step [95 °C for 10 min], a 35 cycle touch down PCR strategy [2 cycles of 95 °C for 15 sec and 63 °C for 1 min, 2 cycles of 95 °C for 15 sec and 62 °C for 1 min, 2 cycles of 95 °C for 15 sec and 61 °C for 1 min, 2 cycles of 95 °C for 15 sec and 60 °C for 1 min, 2 cycles of 95 °C for 15 sec and 58 °C for 1 min, 25 cycles of 95 °C for 15 sec and 72 °C for 1 min], and an elongation step [72 °C for 3 min]. The real-time data was analyzed with Fluidigm Real-Time PCR Analysis software to obtain C_T values for each reaction chamber.

[0256] After amplification, the PCR products were harvested from the Access Array IFC using the Post-PCR IFC Loader AX. Before harvesting, each sample port was filled with 2 μl of 0.05% Tween-20. Residual solution was removed from the H1-H4 reagent wells, and they were refilled with 600 μl of 1X Access Array Harvesting Reagent (0.05% tween-20). After harvesting ,each sample port became a PCR product outlet that contained 10 μl (± 10%) of 48 pooled PCR products. The pooled PCR products were removed from the Access Array IFC and stored in a microtiter plate at 4 °C.

[0257] 1 μl of each PCR product pool for each sample was taken and loaded onto an Agilent 1K Bioanalyzer chip. Figure 12 shows the electropherograms from each of the 48 individual product pools. Figure 13 shows the distribution of product size within a single product pool. All products fall within the predicted size range, and there is no evidence of any small-sized PCR by-products.

[0258] PCR Products for each sample were pooled based on concentrations calculated from the Agilent Bioanalyzer traces. The product pool was purified using AMPure beads (Agencourt) according to the manufacturer's instructions.

[0259] The purified product pool was subjected to emulsion PCR followed by pyrosequencing on a 454 FLX sequencer (Roche Analytical Sciences) according to manufacturer's instructions. The sequence file output by the sequencer was then analyzed for the presence of barcoded PCR products.

IDT Exhibit 1020.2721

[0260] The number of sequences obtained for each barcode were counted, and plotted (Figure 14(A)). On average ~3400 sequences were counted for each barcode. All samples were represented at >50% of average and < 2-fold of average.

[0261] The number of sequences counted for each individual PCR product in each sample were then analyzed (figure 14(B)). Only one of 2304 PCR products was not observed on the sequencer. The vast majority of sequences were present at >50% of average and <2 fold of average. 2303/2304 products were counted >5 times. Figure 14(C) shows the distribution of PCR products from all 2304 PCR reactions in the Access Array IFC. >95% of sequences were measured between 50% and 2 fold of the average coverage. >99% of sequences were measured between 50% and 2 fold of the average coverage.

Example 5

Multi-Primer Amplification Using Four Outer Primers With Different Combinations Of Primer Binding Site And Nucleotide Tags

[0262] Sets of primer pairs were designed to amplify specific regions from the EGFR and MET genes. These were then combined in an Access Array IFC with human genomic DNA and four outer primers (Figure 15).

Preparation of Reaction Mixtures

[0263] Primers were synthesized by Eurofins MWG Operon at 10nmol scale and provided resuspended in water at a concentration of 100μM. The forward and reverse primer for each region in Table 9 were combined in separate wells in a 96-well PCR plate (USA scientific) to a final concentration of 1μM of each primer in PCR-quality water (Teknova) containing 0.05% Tween-20.

Table 9

Assay	Forward Primer	Forward Primer SEQ ID NO.	Reverse Primer	Reverse Primer SEQ ID NO.
EGFR_Exon3	ACACTGACGACATGGTTCTACATTCTTAGACCATCCAGGAGGTG	(SEQ ID NO:309)	TACGGTAGCAGAGACTTGGTCTCCAGCCTCTCACCCTGTAAA	(SEQ ID NO:310)
EGFR_Exon4	ACACTGACGACATGGTTCTACAAGCTGGAAAGAGTGCTC	(SEQ ID NO:311)	TACGGTAGCAGAGACTTGGTCTTAGGAGCTGGAGGCAGAGAT	(SEQ ID NO:312)

IDT Exhibit 1020.2722

	ACC			
EGFR_Exon5	ACACTGACGACATGGTTCTACAGCGTCATCAGTTTCTCATCATT	(SEQ ID NO:313)	TACGGTAGCAGAGACTTGGTCTACATGGGTCTGAGGCTGTTC	(SEQ ID NO:314)
EGFR_Exon6	ACACTGACGACATGGTTCTACACCCTGGGAAATGATCCTACC	(SEQ ID NO:315)	TACGGTAGCAGAGACTTGGTCTTCTTACCAGGCAGTCGCTCT	(SEQ ID NO:316)
EGFR_Exon7	ACACTGACGACATGGTTCTACACCAGCGTGTCCTCTCTCCT	(SEQ ID NO:317)	TACGGTAGCAGAGACTTGGTCTGACAAGGATGCCTGACCAGT	(SEQ ID NO:318)
EGFR_Exon8	ACACTGACGACATGGTTCTACACAAAGGAGGATGGAGCCTTTC	(SEQ ID NO:319)	TACGGTAGCAGAGACTTGGTCTGATGTGTTCCTTTGGAGGTGG	(SEQ ID NO:320)
EGFR_Exon9	ACACTGACGACATGGTTCTACATCCAACAAATGTGAACGGAAT	(SEQ ID NO:321)	TACGGTAGCAGAGACTTGGTCTCAAGCAACTGAACCTGTGACTC	(SEQ ID NO:322)
EGFR_Exon10	ACACTGACGACATGGTTCTACAGATCAATAATCACCCTGTTGTTTG	(SEQ ID NO:323)	TACGGTAGCAGAGACTTGGTCTTTCCAAGGGAACAGGAAATATG	(SEQ ID NO:324)
EGFR_Exon11	ACACTGACGACATGGTTCTACATCCTACGTGGTGTGTGTCTGA	(SEQ ID NO:325)	TACGGTAGCAGAGACTTGGTCTGCTTTGGCTGTGGTCAACTT	(SEQ ID NO:326)
EGFR_Exon12	ACACTGACGACATGGTTCTACACCACATGATTTTTCTTCTCTCCA	(SEQ ID NO:327)	TACGGTAGCAGAGACTTGGTCTCGGTGACTTACTGCAGCTGTT	(SEQ ID NO:328)
EGFR_Exon13	ACACTGACGACATGGTTCTACAGCTCTGTCACTGACTGCTGTG	(SEQ ID NO:329)	TACGGTAGCAGAGACTTGGTCTGCTATAACAACAACCTGGAGCCT	(SEQ ID NO:330)
EGFR_Exon14	ACACTGACGACATGGTTCTACAGCTGACGGGTTTCCTCTTC	(SEQ ID NO:331)	TACGGTAGCAGAGACTTGGTCTGACGTGGATAGCAGCAAGG	(SEQ ID NO:332)
EGFR_Exon15	ACACTGACGACATGGTTCTACAGCATGAACATTTTTCTCCACCT	(SEQ ID NO:333)	TACGGTAGCAGAGACTTGGTCTTTCTGTTCTCCTTCACTTTCCAC	(SEQ ID NO:334)
EGFR_Exon16	ACACTGACGACATGGTTCTACATTTCTCTTTCACTTCCTACAGATGC	(SEQ ID NO:335)	TACGGTAGCAGAGACTTGGTCTCCACAGCAGTGTGGTCATTC	(SEQ ID NO:336)
EGFR_Exon17	ACACTGACGACATGGTTCTACATGGAATCTGTCAGCAACCTC	(SEQ ID NO:337)	TACGGTAGCAGAGACTTGGTCTCCCAGGACTGGCACTCA	(SEQ ID NO:338)
EGFR_Exon18	ACACTGACGACATGGTTCTACAGCTGAGGTGACCCTTGTCTC	(SEQ ID NO:339)	TACGGTAGCAGAGACTTGGTCTCCCACCAGACCATGAGAGG	(SEQ ID NO:340)
EGFR_Exon19	ACACTGACGACATGGTTCTACATCAC	(SEQ ID NO:341)	TACGGTAGCAGAGACTTGGTCTCCACACAGCAAAGCAG	(SEQ ID NO:342)

IDT Exhibit 1020.2723

	AATTGCCAGTTAACGTCT		AAAC	
EGFR_Exon20	ACACTGACGACATGGTTCTACACCACACTGACGTGCCTCTC	(SEQ ID NO:343)	TACGGTAGCAGAGACTTGGTCTCCGTATCTCCCTTCCCTGAT	(SEQ ID NO:344)
EGFR_Exon21	ACACTGACGACATGGTTCTACACCTCACAGCAGGGTCTTCTC	(SEQ ID NO:345)	TACGGTAGCAGAGACTTGGTCTCTGACCTAAAGCCACCTCCTT	(SEQ ID NO:346)
EGFR_Exon22	ACACTGACGACATGGTTCTACACACTGCCTCATCTCTCACCA	(SEQ ID NO:347)	TACGGTAGCAGAGACTTGGTCTCCAGCTTGGCCTCAGTACA	(SEQ ID NO:348)
EGFR_Exon23	ACACTGACGACATGGTTCTACACATGATCCCACTGCCTTCTT	(SEQ ID NO:349)	TACGGTAGCAGAGACTTGGTCTAGTGTGGACAGACCCACCA	(SEQ ID NO:350)
EGFR_Exon24	ACACTGACGACATGGTTCTACATTCCAGTGTTCTAATTGCACTGTT	(SEQ ID NO:351)	TACGGTAGCAGAGACTTGGTCTGAGGGACTCTTCCCAATGGA	(SEQ ID NO:352)
EGFR_Exon25	ACACTGACGACATGGTTCTACACTAATAGCCTCAAAATCTCTGCAC	(SEQ ID NO:353)	TACGGTAGCAGAGACTTGGTCTTTTGTTCAAATGAGTAGACACAGC	(SEQ ID NO:354)
EGFR_Exon26	ACACTGACGACATGGTTCTACACATTCCATGGGCAACTTCTC	(SEQ ID NO:355)	TACGGTAGCAGAGACTTGGTCTTTCTGGCTTATAAGGTGTTCATACA	(SEQ ID NO:356)
EGFR_Exon27	ACACTGACGACATGGTTCTACACCTTCCCTCATTTCCTCCTG	(SEQ ID NO:357)	TACGGTAGCAGAGACTTGGTCTTCCAGACAAGCCACTCACC	(SEQ ID NO:358)
EGFR_Exon28-1	ACACTGACGACATGGTTCTACAcctctgatttctttccactttca	(SEQ ID NO:359)	TACGGTAGCAGAGACTTGGTCTCTAATTTGGTGGCTGCCTTT	(SEQ ID NO:360)
EGFR_Exon28-2	ACACTGACGACATGGTTCTACATGTCAACAGCACATTCGACAG	(SEQ ID NO:361)	TACGGTAGCAGAGACTTGGTCTGGTCCTGGGTATCGAAAGAGT	(SEQ ID NO:362)
EGFR_Exon2	ACACTGACGACATGGTTCTACATTTCTTCCAGTTTGCCAAGG	(SEQ ID NO:363)	TACGGTAGCAGAGACTTGGTCTAGGAAAATCAAAGTCACCAACC	(SEQ ID NO:364)
MET_Exon1-1	ACACTGACGACATGGTTCTACACTCTCGCCTTGAACCTGTTT	(SEQ ID NO:365)	TACGGTAGCAGAGACTTGGTCTCAGCACAGGCCCAGTCTT	(SEQ ID NO:366)
MET_Exon1-2	ACACTGACGACATGGTTCTACATTCCTTGGTGCCACTAACTACA	(SEQ ID NO:367)	TACGGTAGCAGAGACTTGGTCTGGGAGAATATGCAGTGAACCTC	(SEQ ID NO:368)
MET_Exon2	ACACTGACGACATGGTTCTACATGGATTCACATTAACTCTATGACCA	(SEQ ID NO:369)	TACGGTAGCAGAGACTTGGTCTTTGCACAATACCAGATAGAACAGAC	(SEQ ID NO:370)
MET_Exon3	ACACTGACGACATGGTTCTACATGAG	(SEQ ID NO:371)	TACGGTAGCAGAGACTTGGTCTCGTCTATGGAAATTCC	(SEQ ID NO:372)

IDT Exhibit 1020.2724

	CTTGTTGGAATAAGGATG		CTGTG	
MET_Exon4	ACACTGACGACATGGTTCTACAGAAGCTCTTTCCACCCCTTC	(SEQ ID NO:373)	TACGGTAGCAGAGACTTGGTCTTGCCAGCTGTTAGAGATTCCT	(SEQ ID NO:374)
MET_Exon5	ACACTGACGACATGGTTCTACATGTCCTTGTAGGTTTTCCCAAA	(SEQ ID NO:375)	TACGGTAGCAGAGACTTGGTCTCCCCAGCAAAGCATTTTAAG	(SEQ ID NO:376)
MET_Exon6	ACACTGACGACATGGTTCTACAGAAAATTCCTTGGATTTGTCATG	(SEQ ID NO:377)	TACGGTAGCAGAGACTTGGTCTCATGATAGGATAGAATCTTCCTTACCA	(SEQ ID NO:378)
MET_Exon7	ACACTGACGACATGGTTCTACAGTTTTGTTTTTATCTCCCCTCCA	(SEQ ID NO:379)	TACGGTAGCAGAGACTTGGTCTTTCAAATTGACAGATGCAACAA	(SEQ ID NO:380)
MET_Exon8	ACACTGACGACATGGTTCTACAGGAACCATTGAGTTATATCCTTTTG	(SEQ ID NO:381)	TACGGTAGCAGAGACTTGGTCTTTGTTTTCTTATACCCATCAGAAGC	(SEQ ID NO:382)
MET_Exon9	ACACTGACGACATGGTTCTACATTGGTGGAAAGAACCTCTCAA	(SEQ ID NO:383)	TACGGTAGCAGAGACTTGGTCTCAGGTACCATGAAAGCCACA	(SEQ ID NO:384)
MET_Exon10	ACACTGACGACATGGTTCTACATGTTGCCAAGCTGTATTCTGTT	(SEQ ID NO:385)	TACGGTAGCAGAGACTTGGTCTTTTGAGCTGATGATTTAAGACAGTG	(SEQ ID NO:386)
MET_Exon12	ACACTGACGACATGGTTCTACAGGACCCAAAGTGCTACAACC	(SEQ ID NO:387)	TACGGTAGCAGAGACTTGGTCTCAAGAATCGACGACAATCTTAAAC	(SEQ ID NO:388)
MET_Exon13	ACACTGACGACATGGTTCTACAGCCCATGATAGCCGTCTTTA	(SEQ ID NO:389)	TACGGTAGCAGAGACTTGGTCTCAACAATGTCACAACCCACTG	(SEQ ID NO:390)
MET_Exon14	ACACTGACGACATGGTTCTACACCTTCATCTTACAGATCAGTTTCCT	(SEQ ID NO:391)	TACGGTAGCAGAGACTTGGTCTGCTTACTGGAAAATCGTATTTAACAAA	(SEQ ID NO:392)
MET_Exon15	ACACTGACGACATGGTTCTACAACGCAGTGCTAACCAAGTTCT	(SEQ ID NO:393)	TACGGTAGCAGAGACTTGGTCTTCCACAAGGGGAAAGTGTAAA	(SEQ ID NO:394)
MET_Exon16	ACACTGACGACATGGTTCTACATGTCTCCACCACTGGATTTCT	(SEQ ID NO:395)	TACGGTAGCAGAGACTTGGTCTGGCTTACAGCTAGTTTGCCAGT	(SEQ ID NO:396)
MET_Exon17	ACACTGACGACATGGTTCTACATGCTTTTCTAACTCTCTTTGACTGC	(SEQ ID NO:397)	TACGGTAGCAGAGACTTGGTCTTCCTCCTTGTCACTTAATTTGGA	(SEQ ID NO:398)
MET_Exon18	ACACTGACGACATGGTTCTACATTCTATTTCAGCCACGGGTAA	(SEQ ID NO:399)	TACGGTAGCAGAGACTTGGTCTAGAGGAGAAACTCAGAGATAACCAA	(SEQ ID NO:400)
MET_Exon19	ACACTGACGACAT	(SEQ ID	TACGGTAGCAGAGACTTGG	(SEQ ID

IDT Exhibit 1020.2725

	GGTTCTACACTCA CCTCATCTGTCCTG TTTCT	NO:401)	TCTGGCATTTCTGTAAAAG TAAAGAACG	NO:402)
MET_Exon20	ACACTGACGACAT GGTTCTACACCTG CCTTCAAAGGGTC TCT	(SEQ ID NO:403)	TACGGTAGCAGAGACTTGG TCTGTGTGGACTGTTGCTT TGACA	(SEQ ID NO:404)

[0264] A single human Genomic DNA sample (Coriell NA10830) was resuspended at 50ng/µl in low-EDTA TE buffer (Teknova) and prepared for PCR as follows.

[0265] A pre-sample mixture was prepared as follows:

Table 10

Pre-sample mixture	**Volume per sample (µl)**	**Volume for 64 samples (µl)**
Faststart High Fidelity reaction Buffer with $MgCl_2$	0.5	32
DMSO	0.1	6.4
PCR-Grade Nucleotide Mixture	0.1	6.4
Faststart High-Fidelity Enzyme Blend (Roche 04 738 292 001)	0.05	3.2
20x Access Array Loading Reagent (PN: 100-0883)	0.25	16
PCR-Grade water	0.5	32
Total	2.5	160

[0266] For each sample replicate, a sample mixture containing forward and reverse barcode primers, genomic DNA and pre-sample mix was prepared in an individual well in a 96-well PCR plate.

Table 11

Sample Mixture	**Volume (µl)**
Pre-sample Mixture	2
4 µM forward barcode primer	0.5

IDT Exhibit 1020.2726

4 μm reverse barcode primer	0.5
Genomic DNA (50ng/μl)	1
PCR-grade water	1

[0267] Four replicate samples were prepared by mixing each sample with one pair of barcode primers selected from Table 12.

Table 12

	Reverse barcode primer (454B-BC#-CS1)	Reverse barcode primer SEQ ID NO.	Forward barcode primer (454A-BC#-CS2)	Forward barcode primer SEQ ID NO.
1	GCCTTGCCAGCCC GCTCAGGCATGC TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:405)	GCCTCCCTCGCGCC ATCAGGCATGCAC ACTGACGACATGGT TCTACA	(SEQ ID NO:406)
2	GCCTTGCCAGCCC GCTCAGCGTACG TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:407)	GCCTCCCTCGCGCC ATCAGCGTACGAC ACTGACGACATGGT TCTACA	(SEQ ID NO:408)
3	GCCTTGCCAGCCC GCTCAGGTCAGC TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:409)	GCCTCCCTCGCGCC ATCAGGTCAGCAC ACTGACGACATGGT TCTACA	(SEQ ID NO:410)
4	GCCTTGCCAGCCC GCTCAGAGCTGC TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:411)	GCCTCCCTCGCGCC ATCAGAGCTGCAC ACTGACGACATGGT TCTACA	(SEQ ID NO:412)
5	GCCTTGCCAGCCC GCTCAGTGCATCT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:413)	GCCTCCCTCGCGCC ATCAGTGCATCACA CTGACGACATGGTT CTACA	(SEQ ID NO:414)
6	GCCTTGCCAGCCC GCTCAGCTGATGT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:415)	GCCTCCCTCGCGCC ATCAGCTGATGACA CTGACGACATGGTT CTACA	(SEQ ID NO:416)
7	GCCTTGCCAGCCC GCTCAGGTAGTCT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:417)	GCCTCCCTCGCGCC ATCAGGTAGTCACA CTGACGACATGGTT CTACA	(SEQ ID NO:418)
8	GCCTTGCCAGCCC GCTCAGGTCGATT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:419)	GCCTCCCTCGCGCC ATCAGGTCGATACA CTGACGACATGGTT CTACA	(SEQ ID NO:420)
9	GCCTTGCCAGCCC GCTCAGGATACG TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:421)	GCCTCCCTCGCGCC ATCAGGATACGAC ACTGACGACATGGT TCTACA	(SEQ ID NO:422)
10	GCCTTGCCAGCCC GCTCAGTGATGCT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:423)	GCCTCCCTCGCGCC ATCAGTGATGCACA CTGACGACATGGTT CTACA	(SEQ ID NO:424)

IDT Exhibit 1020.2727

11	GCCTTGCCAGCCC GCTCAGAGCTGA TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:425)	GCCTCCCTCGCGCC ATCAGAGCTGAAC ACTGACGACATGGT TCTACA	(SEQ ID NO:426)
12	GCCTTGCCAGCCC GCTCAGACTGTAT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:427)	GCCTCCCTCGCGCC ATCAGACTGTAACA CTGACGACATGGTT CTACA	(SEQ ID NO:428)
13	GCCTTGCCAGCCC GCTCAGTGCATGT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:429)	GCCTCCCTCGCGCC ATCAGTGCATGACA CTGACGACATGGTT CTACA	(SEQ ID NO:430)
14	GCCTTGCCAGCCC GCTCAGAGTCTAT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:431)	GCCTCCCTCGCGCC ATCAGAGTCTAACA CTGACGACATGGTT CTACA	(SEQ ID NO:432)
15	GCCTTGCCAGCCC GCTCAGTGTCTGT ACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:433)	GCCTCCCTCGCGCC ATCAGTGTCTGACA CTGACGACATGGTT CTACA	(SEQ ID NO:434)
16	GCCTTGCCAGCCC GCTCAGGCTAGC TACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:435)	GCCTCCCTCGCGCC ATCAGGCTAGCAC ACTGACGACATGGT TCTACA	(SEQ ID NO:436)

Running the Access Array IFC

[0268] The containment and interface accumulator reservoirs were filled with 300 µl of Control Line Fluid (Fluidigm PN 89000020), and the H1-H4 reagent wells were loaded with 500 µl of 0.05% Tween-20 in PCR-grade water prior to Access Array IFC loading. 5µl of each sample mixture was loaded into the sample ports, and 5 µl of each primer mixture was loaded into the primer inlets on the Access Array IFC.

[0269] The Access Array IFC was thermal cycled and imaged using an IFC Stand-Alone Thermal Cycler (Fluidigm Corporation). The thermal cycling protocol contains a thermal mix step [50 °C for 2 min, 70 °C for 20 min], a hotstart step [95 °C for 10 min], a 35 cycle PCR strategy [2 cycles of 95 °C for 15 sec and 60 °C for 4 min, 33 cycles of 95 °C for 15 sec, 60 °C for 15 sec, 72°C for 1 min, and an elongation step [72 °C for 3 min].

[0270] After amplification, the PCR products were harvested from the Access Array IFC using the Post-PCR IFC Loader AX. Before harvesting, each sample port was filled with 2 µl of 0.05% Tween-20. Residual solution was removed from the H1-H4 reagent wells, and they were refilled with 600 µl of 1X Access Array

IDT Exhibit 1020.2728

Harvesting Reagent (0.05% tween-20). After harvesting each sample port became a PCR product outlet that contained 10 µl (± 10%) of 48 pooled PCR products. The pooled PCR products were removed from the Access Array IFC and stored in a microtiter plate at 4 °C.

[0271] PCR products for each sample were pooled based on concentrations calculated from the Agilent Bioanalyzer traces. The purified product pool was subjected to emulsion PCR followed by pyrosequencing on a 454 FLX sequencer (Roche Analytical Sciences) according to manufacturer's instructions. Emulsion PCR reactions were run with beads containing both A and B primer sequences attached, enabling sequence reads for both strands of the amplicon.

[0272] The number of sequences counted for each individual PCR product in each sample were analyzed to demonstrate representation of the PCR products shown in Figure 15. Sequences could be counted for each of the amplicons shown in Figure 15B, by summing tag 5 sequences from emulsion A with tag 8 sequences from emulsion B (Figure 16A) or tag 5 sequences from emulsion B with tag 8 sequences from emulsion A. Figure 16 shows that representation of all amplicons mostly lies between 2x and 0.5x of average coverage. Furthermore, representation of both amplicons for each primer pair is very similar, although the amplicon represented in Figure 16B shows less variation within samples.

Example 6

4-Primer Barcoding Of Target Nucleic Acids for Illumina DNA Sequencing Using a Microfluidic Device that Permits Recovery of Amplication Products

[0273] Sequences designed for a 4-primer tagging scheme to be used on the Illumina Genome Analyzer II are shown in Tables 13 and 14. The tag sequence is the inner primer sequence.

Table 13

Inner primers		
Tag Sequence	Target-Specific Sequence (Forward)	Oligonucleotide Sequence
ACACTCTTTCCCTACACGACGCTCTTCCGATCT (SEQ ID NO:437)	ACTGTCCAGCTTTGTGCC (SEQ ID NO:438)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGTCCAGCTTTGTGCC (SEQ ID NO:439)
ACACTCTTTCCCTACACGACGCTCTTCCGATCT	GATCATCATAGGAGTTGCATTGTTG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCATCATAGGAGTTGCATTGTTG

IDT Exhibit 1020.2729

CT (SEQ ID NO:440)	(SEQ ID NO:441)	(SEQ ID NO:442)
Tag Sequence	**Target-Specific Sequence (Reverse)**	**Oligonucleotide Sequence**
CTCGGCATTCCTGCTG AACCGCTCTTCCGAT CT (SEQ ID NO:443)	TCCTCTGCCTAGG CGTT (SEQ ID NO:444)	CTCGGCATTCCTGCTGAACCGCTCTTCCG ATCTTCCTCTGCCTAGGCGTT (SEQ ID NO:445)
CTCGGCATTCCTGCTG AACCGCTCTTCCGAT CT (SEQ ID NO:446)	GAAATGTAAATGT GGAGCCAAACA (SEQ ID NO:447)	CTCGGCATTCCTGCTGAACCGCTCTTCCG ATCTGAAATGTAAATGTGGAGCCAAACA (SEQ ID NO:448)

Table 14

Barcode Primers	**Direction**	
ILMN_PE1sh_F	Forward	AATGATACGGCGACCAC**CGAGATC**TACACTCTTT CCCTACACGA (SEQ ID NO:449)
ILMN_PE2sh_R	Reverse	CAAGCAGAAGACGGCATA**CGAGATC**GGTCTCGG CATTCCTGCTGAAC (SEQ ID NO:450)

[0274] The successful amplication of a PCR product using the 4-primer strategy designed for use on the Illumina GA II sequencer is shown in Figure 17.

Example 7

Barcoding Of Target Nucleic Acids for Titanium Chemistry on the 454 FLX Sequencer (Roche Analytical Sciences)

[0275] Table 15 shows forward barcode sequences for use with Titanium chemistry on the 454 FLX Sequencer (Roche Analytical Sciences). Table 16 shows reverse barcode sequences for use with Titanium chemistry on the 454 FLX Sequencer (Roche Analytical Sciences).

Table 15

Well	**Barcode**	**Forward Oligo Name**	**Forward Oligo sequence**	**SEQ ID NO.**
A1	TI-MID1	TI-F-MID1-TAG8	CGTATCGCCTCCCTCGCGCCATCA GACGAGTGCGTACACTGACGACA TGGTTCTACA	(SEQ ID NO:451)
B1	TI-MID2	TI-F-MID2-TAG8	CGTATCGCCTCCCTCGCGCCATCA GACGCTCGACAACACTGACGACA TGGTTCTACA	(SEQ ID NO:452)

IDT Exhibit 1020.2730

C1	TI-MID3	TI-F-MID3-TAG8	CGTATCGCCTCCCTCGCGCCATCAGAGACGCACTCACACTGACGACATGGTTCTACA	(SEQ ID NO:453)
D1	TI-MID67	TI-F-MID67-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCGATAGTGAACACTGACGACATGGTTCTACA	(SEQ ID NO:454)
E1	TI-MID5	TI-F-MID5-TAG8	CGTATCGCCTCCCTCGCGCCATCAGATCAGACACGACACTGACGACATGGTTCTACA	(SEQ ID NO:455)
F1	TI-MID6	TI-F-MID6-TAG8	CGTATCGCCTCCCTCGCGCCATCAGATATCGCGAGACACTGACGACATGGTTCTACA	(SEQ ID NO:456)
G1	TI-MID7	TI-F-MID7-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCGTGTCTCTAACACTGACGACATGGTTCTACA	(SEQ ID NO:457)
H1	TI-MID8	TI-F-MID8-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCTCGCGTGTCACACTGACGACATGGTTCTACA	(SEQ ID NO:458)
A2	TI-MID10	TI-F-MID10-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCTCTATGCGACACTGACGACATGGTTCTACA	(SEQ ID NO:459)
B2	TI-MID11	TI-F-MID11-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTGATACGTCTACACTGACGACATGGTTCTACA	(SEQ ID NO:460)
C2	TI-MID13	TI-F-MID13-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCATAGTAGTGACACTGACGACATGGTTCTACA	(SEQ ID NO:461)
D2	TI-MID14	TI-F-MID14-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCGAGAGATACACACTGACGACATGGTTCTACA	(SEQ ID NO:462)
E2	TI-MID15	TI-F-MID15-TAG8	CGTATCGCCTCCCTCGCGCCATCAGATACGACGTAACACTGACGACATGGTTCTACA	(SEQ ID NO:463)
F2	TI-MID16	TI-F-MID16-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCACGTACTAACACTGACGACATGGTTCTACA	(SEQ ID NO:464)
G2	TI-MID17	TI-F-MID17-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCGTCTAGTACACACTGACGACATGGTTCTACA	(SEQ ID NO:465)
H2	TI-MID18	TI-F-MID18-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCTACGTAGCACACTGACGACATGGTTCTACA	(SEQ ID NO:466)
A3	TI-MID19	TI-F-MID19-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTGTACTACTCACACTGACGACATGGTTCTACA	(SEQ ID NO:467)
B3	TI-MID20	TI-F-MID20-TAG8	CGTATCGCCTCCCTCGCGCCATCAGACGACTACAGACACTGACGACATGGTTCTACA	(SEQ ID NO:468)
C3	TI-MID21	TI-F-MID21-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCGTAGACTAGACACTGACGACATGGTTCTACA	(SEQ ID NO:469)
D3	TI-MID22	TI-F-MID22-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACGAGTATGACACTGACGACATGGTTCTACA	(SEQ ID NO:470)
E3	TI-MID23	TI-F-MID23-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACTCTCGTGACACTGACGACATGGTTCTACA	(SEQ ID NO:471)
F3	TI-MID24	TI-F-MID24-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTAGAGACGAGACACTGACGACA	(SEQ ID NO:472)

IDT Exhibit 1020.2731

			TGGTTCTACA	
G3	TI-MID25	TI-F-MID25-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCGTCGCTCGACACTGACGACATGGTTCTACA	(SEQ ID NO:473)
H3	TI-MID26	TI-F-MID26-TAG8	CGTATCGCCTCCCTCGCGCCATCAGACATACGCGTACACTGACGACATGGTTCTACA	(SEQ ID NO:474)
A4	TI-MID27	TI-F-MID27-TAG8	CGTATCGCCTCCCTCGCGCCATCAGACGCGAGTATACACTGACGACATGGTTCTACA	(SEQ ID NO:475)
B4	TI-MID28	TI-F-MID28-TAG8	CGTATCGCCTCCCTCGCGCCATCAGACTACTATGTACACTGACGACATGGTTCTACA	(SEQ ID NO:476)
C4	TI-MID68	TI-F-MID68-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCGCTGCGTAACACTGACGACATGGTTCTACA	(SEQ ID NO:477)
D4	TI-MID30	TI-F-MID30-TAG8	CGTATCGCCTCCCTCGCGCCATCAGAGACTATACTACACTGACGACATGGTTCTACA	(SEQ ID NO:478)
E4	TI-MID31	TI-F-MID31-TAG8	CGTATCGCCTCCCTCGCGCCATCAGAGCGTCGTCTACACTGACGACATGGTTCTACA	(SEQ ID NO:479)
F4	TI-MID32	TI-F-MID32-TAG8	CGTATCGCCTCCCTCGCGCCATCAGAGTACGCTATACACTGACGACATGGTTCTACA	(SEQ ID NO:480)
G4	TI-MID33	TI-F-MID33-TAG8	CGTATCGCCTCCCTCGCGCCATCAGATAGAGTACTACACTGACGACATGGTTCTACA	(SEQ ID NO:481)
H4	TI-MID34	TI-F-MID34-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCACGCTACGTACACTGACGACATGGTTCTACA	(SEQ ID NO:482)
A5	TI-MID35	TI-F-MID35-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCAGTAGACGTACACTGACGACATGGTTCTACA	(SEQ ID NO:483)
B5	TI-MID36	TI-F-MID36-TAG8	CGTATCGCCTCCCTCGCGCCATCAGCGACGTGACTACACTGACGACATGGTTCTACA	(SEQ ID NO:484)
C5	TI-MID37	TI-F-MID37-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACACACACTACACTGACGACATGGTTCTACA	(SEQ ID NO:485)
D5	TI-MID38	TI-F-MID38-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACACGTGATACACTGACGACATGGTTCTACA	(SEQ ID NO:486)
E5	TI-MID39	TI-F-MID39-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACAGATCGTACACTGACGACATGGTTCTACA	(SEQ ID NO:487)
F5	TI-MID40	TI-F-MID40-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTACGCTGTCTACACTGACGACATGGTTCTACA	(SEQ ID NO:488)
G5	TI-MID69	TI-F-MID69-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCTGACGTCAACACTGACGACATGGTTCTACA	(SEQ ID NO:489)
H5	TI-MID42	TI-F-MID42-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCGATCACGTACACTGACGACATGGTTCTACA	(SEQ ID NO:490)
A6	TI-MID43	TI-F-MID43-TAG8	CGTATCGCCTCCCTCGCGCCATCAGTCGCACTAGTACACTGACGACATGGTTCTACA	(SEQ ID NO:491)
B6	TI-MID44	TI-F-MID44-	CGTATCGCCTCCCTCGCGCCATCA	(SEQ ID NO:492)

IDT Exhibit 1020.2732

		TAG8	GTCTAGCGACTACACTGACGACAT GGTTCTACA	
C6	TI-MID45	TI-F-MID45-TAG8	CGTATCGCCTCCCTCGCGCCATCA GTCTATACTATACACTGACGACAT GGTTCTACA	(SEQ ID NO:493)
D6	TI-MID46	TI-F-MID46-TAG8	CGTATCGCCTCCCTCGCGCCATCA GTGACGTATGTACACTGACGACAT GGTTCTACA	(SEQ ID NO:494)
E6	TI-MID47	TI-F-MID47-TAG8	CGTATCGCCTCCCTCGCGCCATCA GTGTGAGTAGTACACTGACGACA TGGTTCTACA	(SEQ ID NO:495)
F6	TI-MID48	TI-F-MID48-TAG8	CGTATCGCCTCCCTCGCGCCATCA GACAGTATATAACACTGACGACA TGGTTCTACA	(SEQ ID NO:496)
G6	TI-MID49	TI-F-MID49-TAG8	CGTATCGCCTCCCTCGCGCCATCA GACGCGATCGAACACTGACGACA TGGTTCTACA	(SEQ ID NO:497)
H6	TI-MID50	TI-F-MID50-TAG8	CGTATCGCCTCCCTCGCGCCATCA GACTAGCAGTAACACTGACGACA TGGTTCTACA	(SEQ ID NO:498)

Table 16

Well	Barcode	Reverse Oligo Name	Reverse Oligo Sequence	SEQ ID NO.
A1	TI-MID1	TI-R-MID1-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GACGAGTGCGTTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:499)
B1	TI-MID2	TI-R-MID2-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GACGCTCGACATACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:500)
C1	TI-MID3	TI-R-MID3-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GAGACGCACTCTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:501)
D1	TI-MID67	TI-R-MID67-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GTCGATAGTGATACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:502)
E1	TI-MID5	TI-R-MID5-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GATCAGACACGTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:503)
F1	TI-MID6	TI-R-MID6-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GATATCGCGAGTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:504)
G1	TI-MID7	TI-R-MID7-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GCGTGTCTCTATACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:505)
H1	TI-MID8	TI-R-MID8-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GCTCGCGTGTCTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:506)
A2	TI-MID10	TI-R-MID10-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GTCTCTATGCGTACGGTAGCAGAG ACTTGGTCT	(SEQ ID NO:507)
B2	TI-MID11	TI-R-MID11-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GTGATACGTCTTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:508)
C2	TI-MID13	TI-R-MID13-TAG5	CTATGCGCCTTGCCAGCCCGCTCA GCATAGTAGTGTACGGTAGCAGA GACTTGGTCT	(SEQ ID NO:509)

IDT Exhibit 1020.2733

D2	TI-MID14	TI-R-MID14-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCGAGAGATACTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:510)
E2	TI-MID15	TI-R-MID15-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGATACGACGTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:511)
F2	TI-MID16	TI-R-MID16-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCACGTACTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:512)
G2	TI-MID17	TI-R-MID17-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCGTCTAGTACTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:513)
H2	TI-MID18	TI-R-MID18-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCTACGTAGCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:514)
A3	TI-MID19	TI-R-MID19-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTGTACTACTCTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:515)
B3	TI-MID20	TI-R-MID20-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACGACTACAGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:516)
C3	TI-MID21	TI-R-MID21-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCGTAGACTAGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:517)
D3	TI-MID22	TI-R-MID22-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACGAGTATGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:518)
E3	TI-MID23	TI-R-MID23-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACTCTCGTGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:519)
F3	TI-MID24	TI-R-MID24-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTAGAGACGAGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:520)
G3	TI-MID25	TI-R-MID25-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCGTCGCTCGTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:521)
H3	TI-MID26	TI-R-MID26-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACATACGCGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:522)
A4	TI-MID27	TI-R-MID27-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACGCGAGTATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:523)
B4	TI-MID28	TI-R-MID28-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACTACTATGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:524)
C4	TI-MID68	TI-R-MID68-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCGCTGCGTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:525)
D4	TI-MID30	TI-R-MID30-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGAGACTATACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:526)
E4	TI-MID31	TI-R-MID31-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGAGCGTCGTCTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:527)
F4	TI-MID32	TI-R-MID32-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGAGTACGCTATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:528)
G4	TI-MID33	TI-R-MID33-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGATAGAGTACTTACGGTAGCAGA	(SEQ ID NO:529)

IDT Exhibit 1020.2734

			GACTTGGTCT	
H4	TI-MID34	TI-R-MID34-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCACGCTACGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:530)
A5	TI-MID35	TI-R-MID35-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCAGTAGACGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:531)
B5	TI-MID36	TI-R-MID36-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGCGACGTGACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:532)
C5	TI-MID37	TI-R-MID37-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACACACACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:533)
D5	TI-MID38	TI-R-MID38-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACACGTGATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:534)
E5	TI-MID39	TI-R-MID39-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACAGATCGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:535)
F5	TI-MID40	TI-R-MID40-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTACGCTGTCTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:536)
G5	TI-MID69	TI-R-MID69-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCTGACGTCATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:537)
H5	TI-MID42	TI-R-MID42-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCGATCACGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:538)
A6	TI-MID43	TI-R-MID43-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCGCACTAGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:539)
B6	TI-MID44	TI-R-MID44-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCTAGCGACTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:540)
C6	TI-MID45	TI-R-MID45-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTCTATACTATTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:541)
D6	TI-MID46	TI-R-MID46-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTGACGTATGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:542)
E6	TI-MID47	TI-R-MID47-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGTGTGAGTAGTTACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:543)
F6	TI-MID48	TI-R-MID48-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACAGTATATATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:544)
G6	TI-MID49	TI-R-MID49-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACGCGATCGATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:545)
H6	TI-MID50	TI-R-MID50-TAG5	CTATGCGCCTTGCCAGCCCGCTCAGACTAGCAGTATACGGTAGCAGAGACTTGGTCT	(SEQ ID NO:546)

IDT Exhibit 1020.2735

Example 8

Muliplex Barcoding Of Target Nucleic Acids

[0276] Three pools of 10 primers were assembled from the primers listed in Table 9. PCR conditions were identical to those listed in Example 4, with the exception that primer concentrations were varied. Figure 18 shows the results of PCR reactions of three pools of 10 sets of PCR primers (A, B, C) when the PCR reactions were run for template-specific primers only and in 4-primer mode. The presence of higher molecular weight products in the 4-primer strategy dmonestrates successful 4-primer assembly.

IDT Exhibit 1020.2736

CLAIMS

What is claimed is:

1. A method for amplifying a plurality of target nucleic acids in a plurality of samples, the method comprising:

preparing an amplification mixture for each target nucleic acid, said amplification mixture comprising:

a forward primer comprising a target-specific portion;

a reverse primer comprising a target-specific portion,

wherein the forward primer additionally comprises a first nucleotide tag and/or the reverse primer additionally comprises a second nucleotide tag; and

at least one barcode primer comprising a barcode nucleotide sequence and a first and/or second nucleotide tag-specific portion, wherein the barcode primer is in excess of the forward and/or reverse primer(s);

subjecting each amplification mixture to amplification to produce a plurality of target amplicons comprising tagged target nucleotide sequences, each comprising first and/or second nucleotide tags flanking the target nucleotide sequence, and at least one barcode nucleotide sequence at the 5' or 3' end of the target amplicon.

2. The method of claim 1, wherein the forward primer additionally comprises a first nucleotide tag.

3. The method of claims 1 or 2, wherein the reverse primer additionally comprises a second nucleotide tag.

4. The method of claim 1, wherein the concentration of the barcode primer in the amplification mixtures is at least 4-fold the concentration of the forward and/or reverse primer(s).

5. The method of claim 4, wherein the concentration of the barcode primer in the amplification mixtures is at least 50-fold the concentration of the forward and/or reverse primer(s).

IDT Exhibit 1020.2737

6. The method of claim 1, wherein the first and/or second nucleotide tags and/or the barcode nucleotide sequence are selected so as to avoid substantial annealing to the target nucleic acids.

7. The method of claim 1, wherein the barcode nucleotide sequence identifies a particular sample.

8. The method of claims 1 or 7, wherein the barcode primer comprises a barcode nucleotide sequence and a first nucleotide tag-specific portion.

9. The method of claim 8, wherein a plurality of forward primers comprise the same first nucleotide tag.

10. The method of claim 9, wherein the all forward primers used to amplify target sequences in each sample comprise the same first nucleotide tag.

11. The method of claim 7, wherein the forward and reverse primers for each target are initially combined separately from the sample, and each barcode primer is initially combined with its corresponding sample.

12. The method of claim 11, wherein T targets are amplified in S samples, T and S being integers greater than one, the method additionally comprising preparing S×T amplification mixtures wherein the initially combined forward and reverse primers are added to the initially combined samples and barcode primers.

13. The method of claim 1, wherein the amplification is carried out for at least 3 cycles to introduce the first and second nucleotide tags and the barcode nucleotide sequence.

14. The method of claim 13, wherein the amplification is carried out for between 5 and 50 cycles.

15. The method of claims 1, wherein the amplification is carried out for a sufficient number of cycles to normalize target amplicon copy number across targets and across samples.

16. The method of claim 1, wherein at least 50 percent of the target amplicons are present at greater than 50 percent of the average number of copies of

IDT Exhibit 1020.2738

target amplicons and less than 2-fold the average number of copies of target amplicons.

17. A method for amplifying a plurality of target nucleic acids in a plurality of samples, the method comprising:

preparing an amplification mixture for each target nucleic acid, said amplification mixture comprising:

a forward primer comprising a target-specific sequence; and

a reverse primer comprising a target-specific sequence;

subjecting each amplification mixture to amplification to produce a plurality of target nucleotide sequences;

tagging the target nucleotide sequences to produce a plurality of target amplicons, each comprising first and/or second nucleotide tags flanking the target nucleotide sequence;

wherein at least 50 percent of the target amplicons are present at greater than 50 percent of the average number of copies of target amplicons and less than 2-fold the average number of copies of target amplicons.

18. The method of claim 16 or 17, wherein at least 70 percent of the target amplicons are present at greater than 50 percent of the average number of copies of target amplicons and less than 2-fold the average number of copies of target amplicons.

19. The method of claims 16, 17, or 18, wherein the average length of the target amplicons is at least 25 bases.

20. The method of claims 16, 17, or 18, wherein the average length of the target amplicons is at least 50 bases.

21. The method of claims 16, 17, or 18, wherein the average length of the target amplicons is at least 100 bases.

22. The method of claims 16, 17, or 18, wherein the average length of the target amplicons is at least 200 bases.

IDT Exhibit 1020.2739

23. The method of claims 16, 17, or 18, wherein the average length of the target amplicons is at least 1 kilobase.

24. The method of claim 16 or 17, wherein at least 90 percent of the target amplicons are present at greater than 50 percent of the average number of copies of target amplicons and less than 2-fold the average number of copies of target amplicons.

25. The method of any of the preceding claims, wherein the volume of the amplification mixtures is in the range of about 1 picoliter to about 50 nanoliters.

26. The method of claim 17, wherein the volume of the amplification mixtures is in the range of about 5 picoliters to about 25 nanoliters.

27. The method of claims 1 or 17, wherein the amplification mixtures are formed in, or distributed into, separate compartments of a microfluidic device prior to amplification.

28. The method of claim 27, wherein the microfluidic device is fabricated, at least in part, from an elastomeric material.

29. The method of claims 1 or 17, wherein the amplification is carried out by polymerase chain reaction (PCR).

30. The method of claim 1 or 17, additionally comprising recovering the target amplicons from the amplification mixtures.

31. The method of claim 30, wherein the target amplicons are recovered in a volume and/or copy number that varies less than about 50% among the recovered target amplicons.

32. The method of claim 30, additionally comprising subjecting at least one target amplicon to amplification using primers specific for the first and second nucleotide tags to produce a target amplicon lacking the barcode nucleotide sequence.

IDT Exhibit 1020.2740

33. The method of claims 1 or 17, wherein the target nucleic acids comprise genomic DNA.

34. The method of claim 1, wherein one or more of the forward primer, reverse primer, and barcode primer comprises at least one additional primer binding site.

35. The method of claim 34, wherein the barcode primer comprises at least a first additional primer binding site upstream of the barcode nucleotide sequence, which is upstream of the first nucleotide tag.

36. The method of claim 35, wherein the reverse primer comprises at least a second additional primer binding site downstream of the second nucleotide tag.

37. The method of claim 36, wherein the first and second additional primer binding sites are capable of being bound by DNA sequencing primers.

38. The method of claim 17, wherein the first and second nucleotide tags are capable of being bound by DNA sequencing primers.

39. The method of claim 37 or 38, additionally comprising subjecting at least one target amplicon to DNA sequencing.

40. The method of claims 1 or 17, additionally comprising quantifying the amount of target amplicons in the amplification mixtures.

41. The method of claim 40, wherein said quantifying comprises recovering the target amplicons and subjecting them to digital amplification.

42. The method of claim 41, wherein said digital amplification comprises:

distributing the preamplified target amplicons into discrete reaction mixtures, wherein each reaction mixture, on average, includes no more than one amplicon per reaction mixture; and

subjecting the reaction mixtures to amplification.

IDT Exhibit 1020.2741

43. The method of claim 42, wherein said digital amplification comprises real-time PCR.

44. The method of claim 42, wherein said digital amplification comprises endpoint PCR.

45. The method of claim 42, wherein the amplification mixtures are formed in, or distributed into, separate compartments of a microfluidic device prior to amplification.

46. The method of claim 45, wherein the microfluidic device is fabricated, at least in part, from an elastomeric material.

47. The method of claims 1 or 17, wherein the presence of a target amplicon is determined by quantitative real-time polymerase chain reaction (qPCR).

48. The method of claims 1 or 17, wherein a universal qPCR probe is employed in the amplification mixtures to detect target amplicons.

49. The method of claims 1 or 17, wherein one or more target-specific qPCR probes is employed in the amplification mixtures to detect target amplicons.

50. The method of claims 1 or 17, wherein the presence of a target amplicon is detected using a fluorogenic nuclease assay.

51. The method of claims 1 or 17, wherein the presence of a target amplicon is detected using a dual-labeled fluorogenic hydrolysis oligonucleotide probe.

52. The method of claims 1 or 17, additionally comprising determining the amount of each target nucleic acid present in each sample.

53. The method of claims 1 or 17, wherein the method is performed in determining the copy numbers of the target nucleic acids in each sample.

54. The method of claims 1 or 17, wherein the method is performed in determining the genotypes at loci corresponding to the target nucleic acids.

IDT Exhibit 1020.2742

55. The method of claims 1 or 17, wherein the method is performed in determining the expression levels of the target nucleic acids.

56. The method of claims 1 or 17, wherein the method is performed to prepare target nucleic acids for sequencing.

57. A microfluidic device comprising:

a plurality of first input lines;

a plurality of second input lines;

a plurality of sets of first chambers, wherein each set of first chambers is in fluid communication with one of the plurality of first input lines;

a plurality of sets of second chambers, wherein each set of second chambers is in fluid communication with one of the plurality of second input lines;

a plurality of first pump elements in fluid communication with a first portion of the plurality of second input lines; and

a plurality of second pump elements in fluid communication with a second portion of the plurality of second input lines.

58. The microfluidic device of claim 57 wherein the first chambers comprise assay chambers and the second chambers comprise sample chambers.

59. The microfluidic device of claim 57 further comprising a set of valves preventing fluid communication between sample chambers making up a set of sample chambers.

60. The microfluidic device of claim 59 further comprising a set of valves preventing fluid communication between assay chambers making up a set of assay chambers and a sample chamber from each of the sets of sample chambers.

61. The microfluidic device of claim 57 wherein the plurality of first pump elements comprise a plurality of first valves.

62. The microfluidic device of claim 57 wherein the plurality of second pump elements comprise a plurality of second valves.

IDT Exhibit 1020.2743

63. The microfluidic device of claim 57 wherein the microfluidic device comprises elastomeric material.

64. The microfluidic device of claim 57 wherein for a set of the plurality of sets of first chambers:

one of the first chambers is in fluid communication with a second chamber of a first set of second chambers via a fluid line;

another of the first chambers is in fluid communication with a second chamber of a second set of second chambers via another fluid line;

a first valve is operable to obstruct fluid flow through the fluid line; and

a second valve is operable to obstruct fluid flow through the another fluid line.

65. The microfluidic device of claim 64 wherein the first valve and the second valve are actuated using a common source.

66. A method of operating a microfluidic device having an assay chamber, a sample chamber, and a harvesting port, the method comprising:

closing a fluid line between the assay chamber and the sample chamber;

flowing a sample into the sample chamber via a sample input line;

flowing an assay into the assay chamber via an assay input line;

opening the fluid line between the assay chamber and the sample chamber;

combining at least a portion of the sample and at least a portion of the assay to form a mixture;

reacting the mixture to form a reaction product;

closing the fluid line between the assay chamber and the sample chamber;

flowing a harvesting reagent from the harvesting port to the sample chamber; and

removing the reaction product from the microfluidic device.

IDT Exhibit 1020.2744

67. The method of claim 66 further comprising thermocycling the mixture.

68. The method of claim 66 wherein removing the reaction product from the microfluidic device comprises flowing the reaction product through at least a portion of the sample input line to a sample input port.

69. The method of claim 68 wherein flowing the reaction product through at least a portion of the sample input line comprises performing dilation pumping.

70. The method of claim 69 wherein dilation pumping comprises a fluid flow rate of less than or equal to 10 μl per hour.

71. The method of claim 69 wherein performing dilation pumping comprises:

a) closing a first valve disposed between the sample chamber and the sample input port;

b) opening a second valve disposed between the harvesting port and the sample chamber;

c) closing the second valve;

d) opening the first valve; and

repeating steps (a) through (d) a predetermined number of times.

72. The method of claim 70 wherein the fluid flow rate is less than or equal to 5 μl per hour.

73. The method of claim 72 wherein the fluid flow rate is less than or equal to 1 μl per hour.

74. The method of claim 66 wherein combining at least a portion of the sample and at least a portion of the assay to form a mixture comprises a free interface diffusion process.

75. The method of claim 66 wherein removing the reaction product comprises a fluid flow rate of less than or equal to 10 microliters per hour.

IDT Exhibit 1020.2745

76. The method of claim 75 wherein the fluid flow rate is less than or equal to 5 microliters per hour.

77. The method of claim 76 wherein the fluid flow rate is less than or equal to 2 microliters per hour.

78. The method of claim 77 wherein the fluid flow rate is less than or equal to 1 microliters per hour.

79. The method of claim 66 wherein removing the reaction product comprises removing at least 95% of the reaction product from the microfluidic device.

80. The method of claim 66 wherein flowing the sample into the sample chamber via the sample input line comprises applying pressure to a sample input port to pressurize the sample input line.

81. The method of claim 66 wherein flowing the assay into the assay chamber via the assay input line comprises applying pressure to an assay input port to pressurize the assay input line.

82. A method of preparing reaction products, the method comprising:

providing M samples;

providing N assays;

mixing the M samples and N assays to form MxN pairwise combinations, each of the MxN pairwise combinations being contained in a closed volume;

forming MxN reaction products from the MxN pairwise combinations; and

recovering the MxN reaction products.

83. The method of claim 82 wherein the M samples are contained in M sets of first chambers and the N assays are contained in N sets of second chambers.

IDT Exhibit 1020.2746

84. The method of claim 83 wherein the first chambers containing one of the M samples are associated with a set of second chambers, each of the set of second chambers containing one of the N assays.

85. The method of claim 84 wherein mixing the M samples and N assays comprises opening fluid lines to provide for fluid communication between the firsrt chambers containing one of the M samples and each of the set of second chambers containing one of the N assays.

86. The method of claim 82 wherein the first chambers are characterized by a volume of less than or equal to 100 nl.

87. The method of claim 86 wherein the volume is less than or equal to 40 nl.

88. The method of claim 82 wherein the second chambers are characterized by a volume of less than or equal to 10 nl.

89. The method of claim 88 wherein the volume is less than or equal to 2 nl.

90. The method of claim 82 wherein the closed volume is characterized by a volume of less than or equal to 100 nl.

91. The method of claim 82 further comprising thermocycling the MxN pairwise combinations.

92. The method of claim 82 wherein the M samples are provided at M sample ports of a microfluidic device.

93. The method of claim 91 wherein the MxN samples are recovered at the M sample ports of the microfluidic device.

94. The method of claim 82 wherein forming the MxN pairwise combinations is performed concurrently.

95. The method of claim 82 wherein forming the MxN pairwise combinations is performed sequentially.

IDT Exhibit 1020.2747

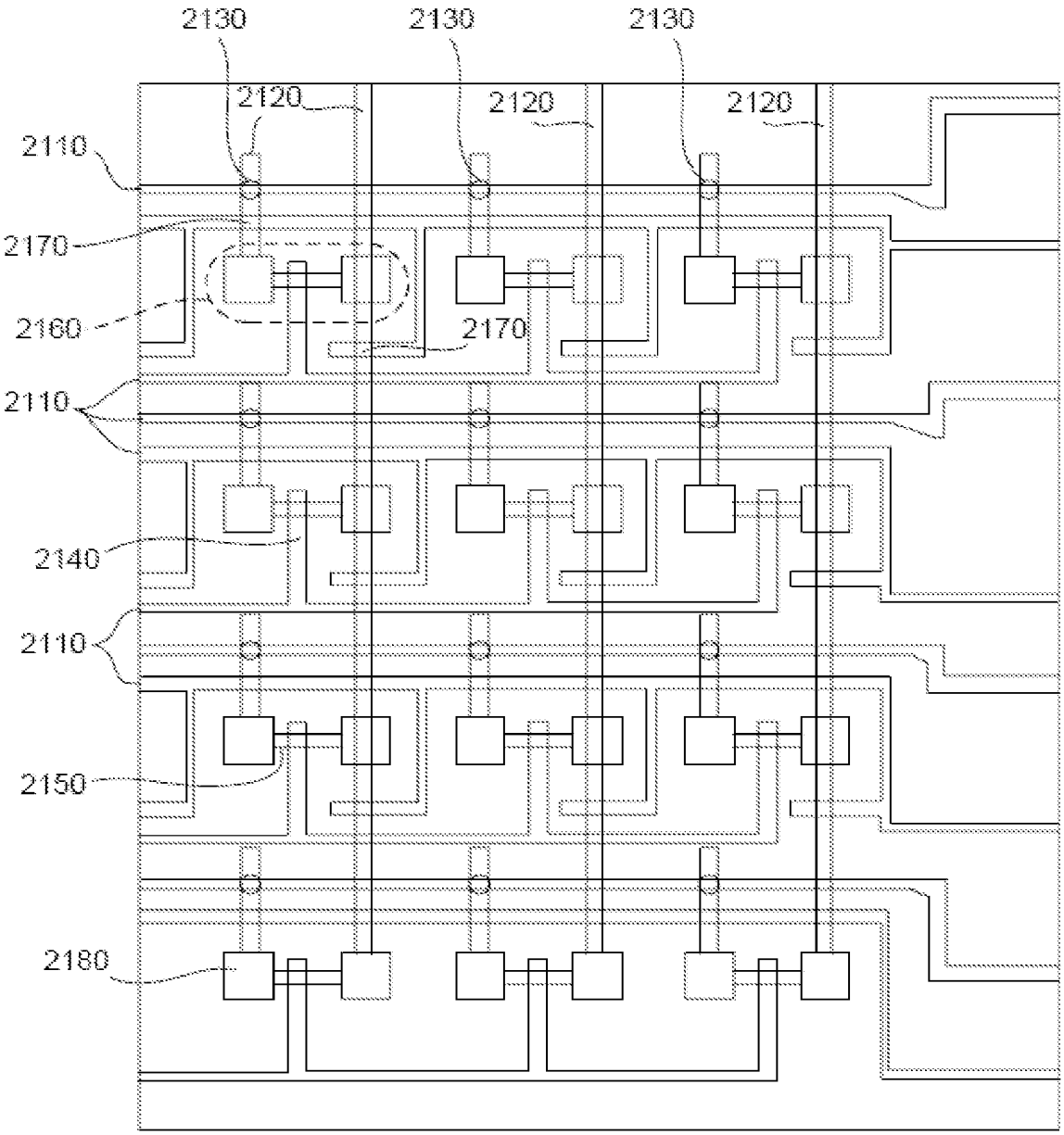

2130
2130
2130
2120
2120
2120
2110
2170
2160
2170
2110
2140
2110
2150
2180

Fig. 1

IDT Exhibit 1020.2749

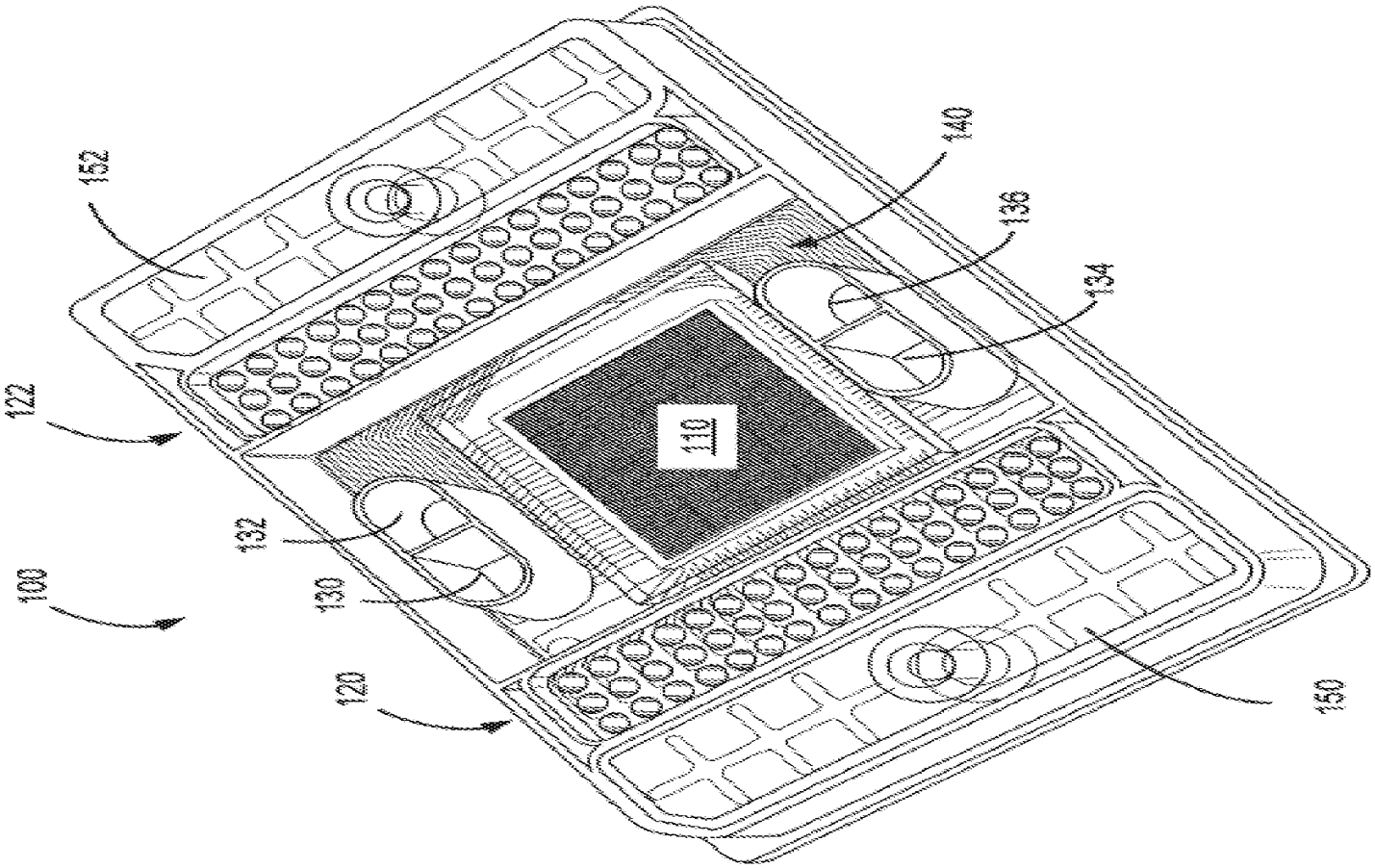


Fig. 2

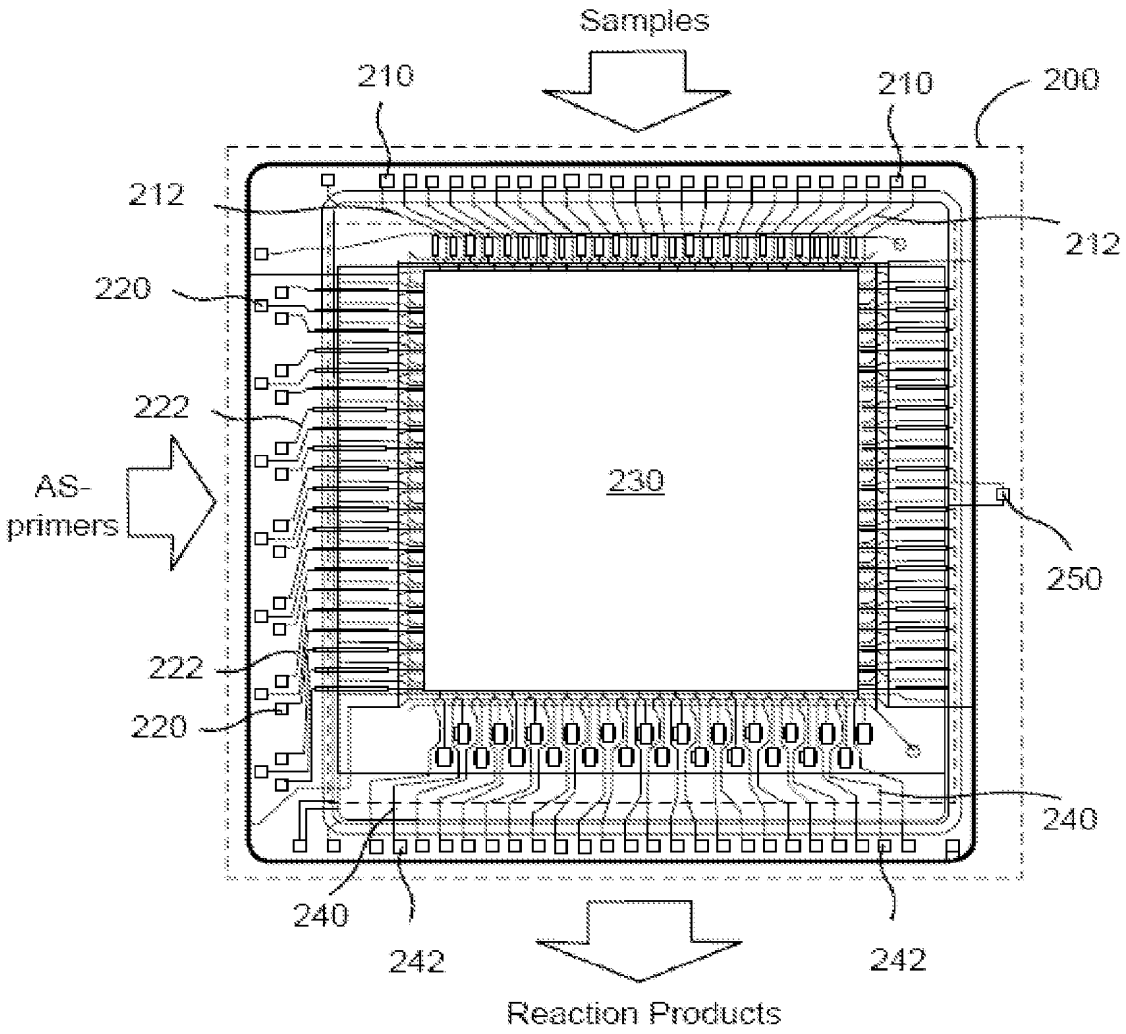

Samples
210
210
200
212
212
220
222
AS-
primers
230
250
222
220
240
240
242
242
Reaction Products

Fig. 3

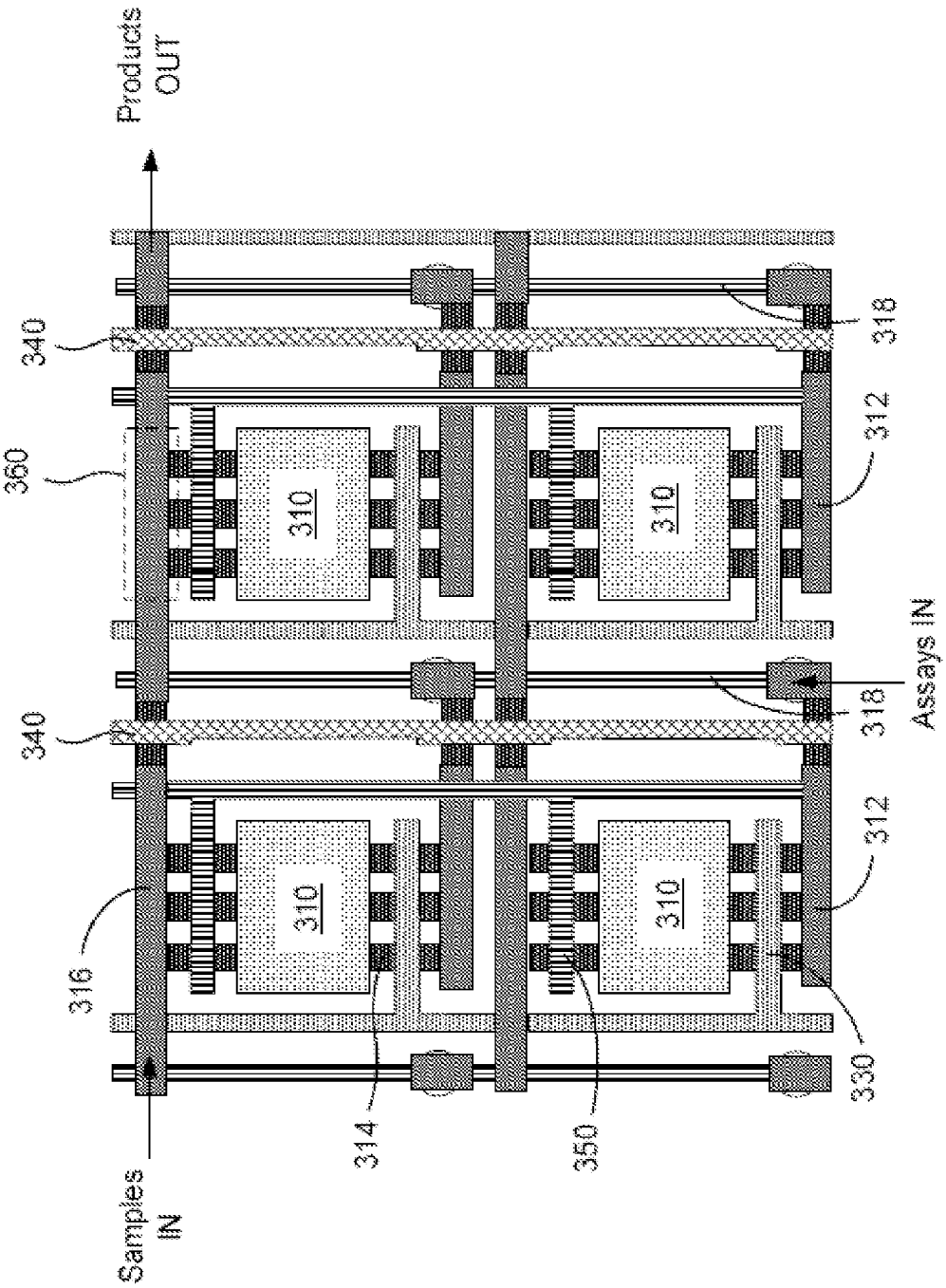

Products OUT
340
360
310
312
318
Assays IN
316
314
350
330
Samples IN

Fig. 4

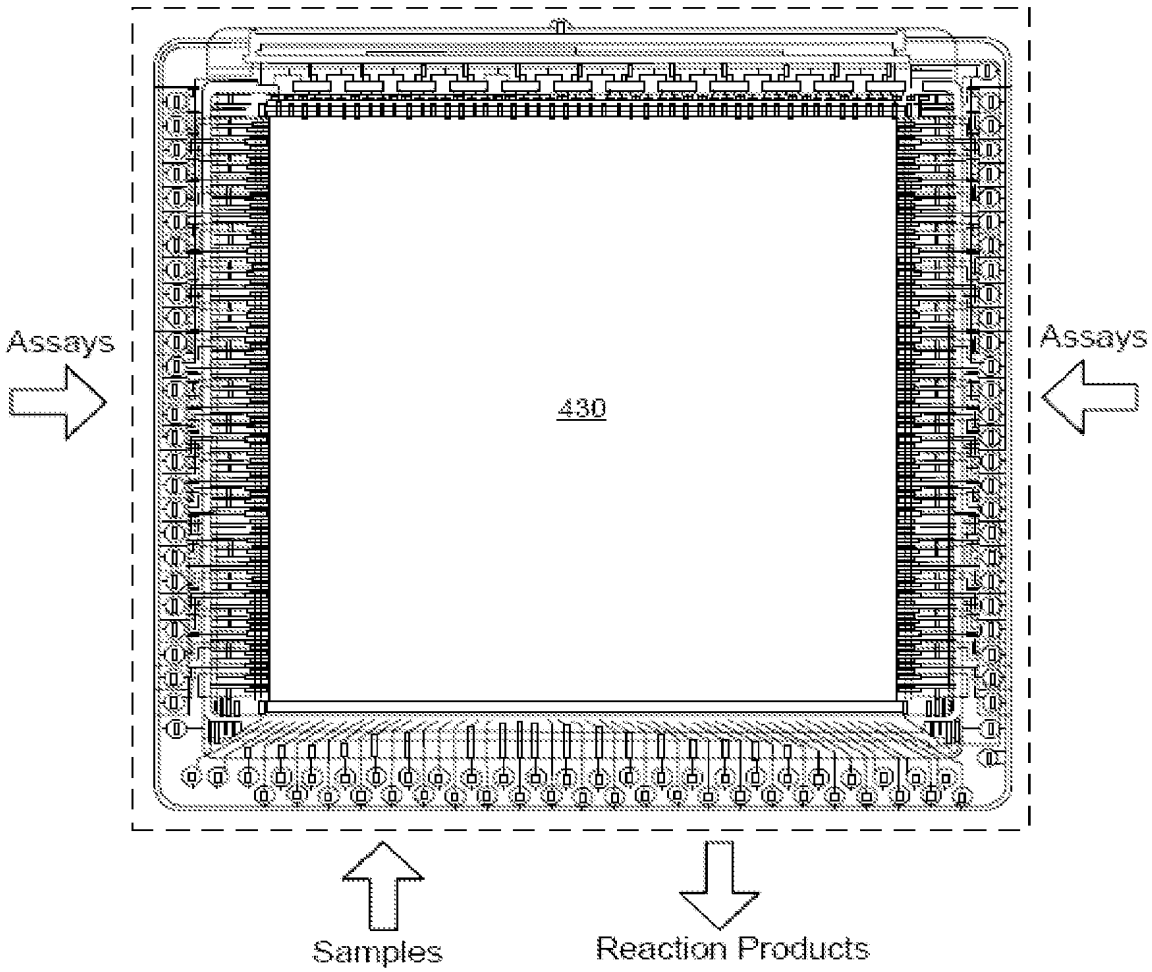


Fig. 5A

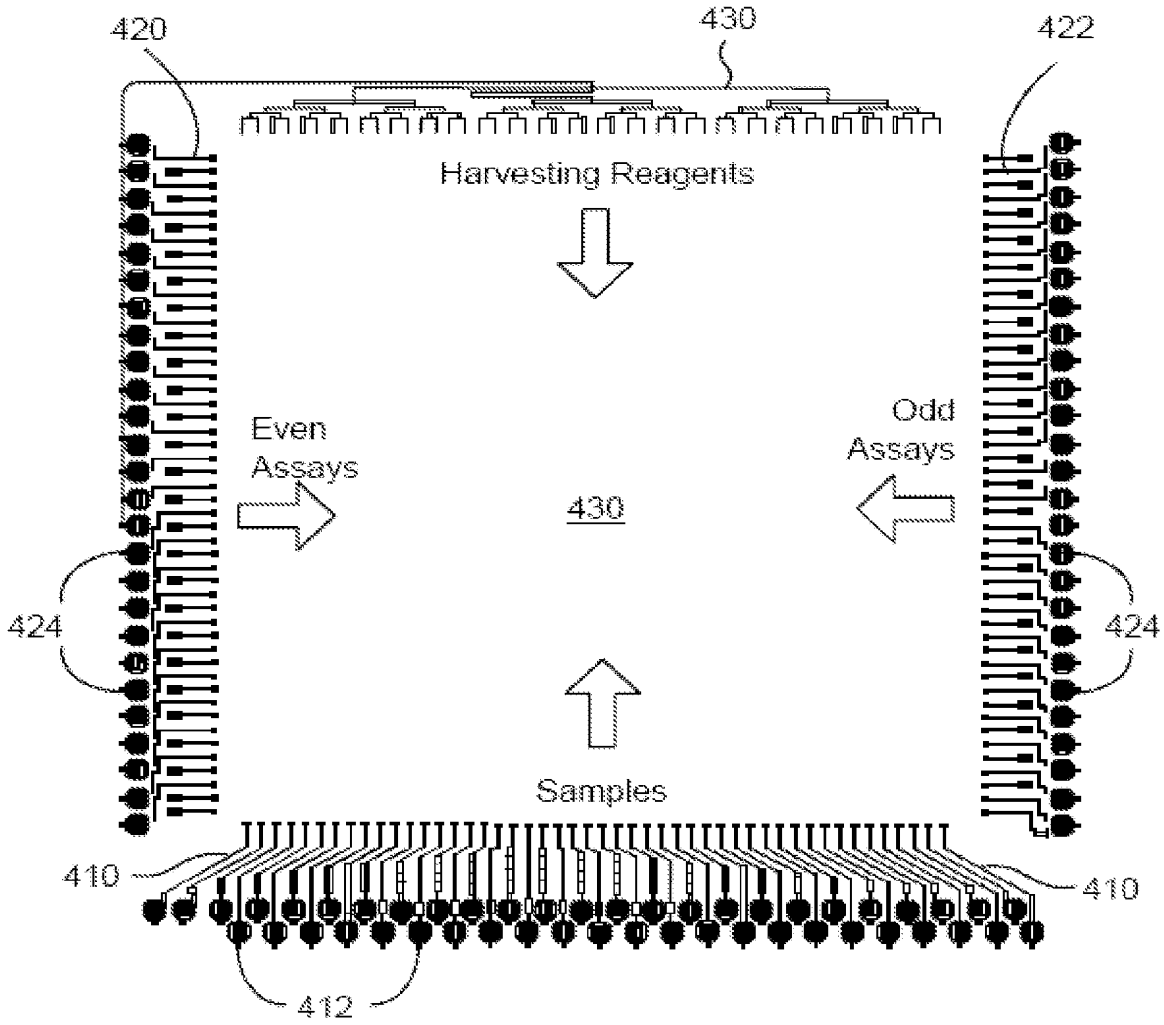

420
430
422
Harvesting Reagents
Even
Assays
Odd
Assays
430
424
424
Samples
410
410
412

Fig. 5B

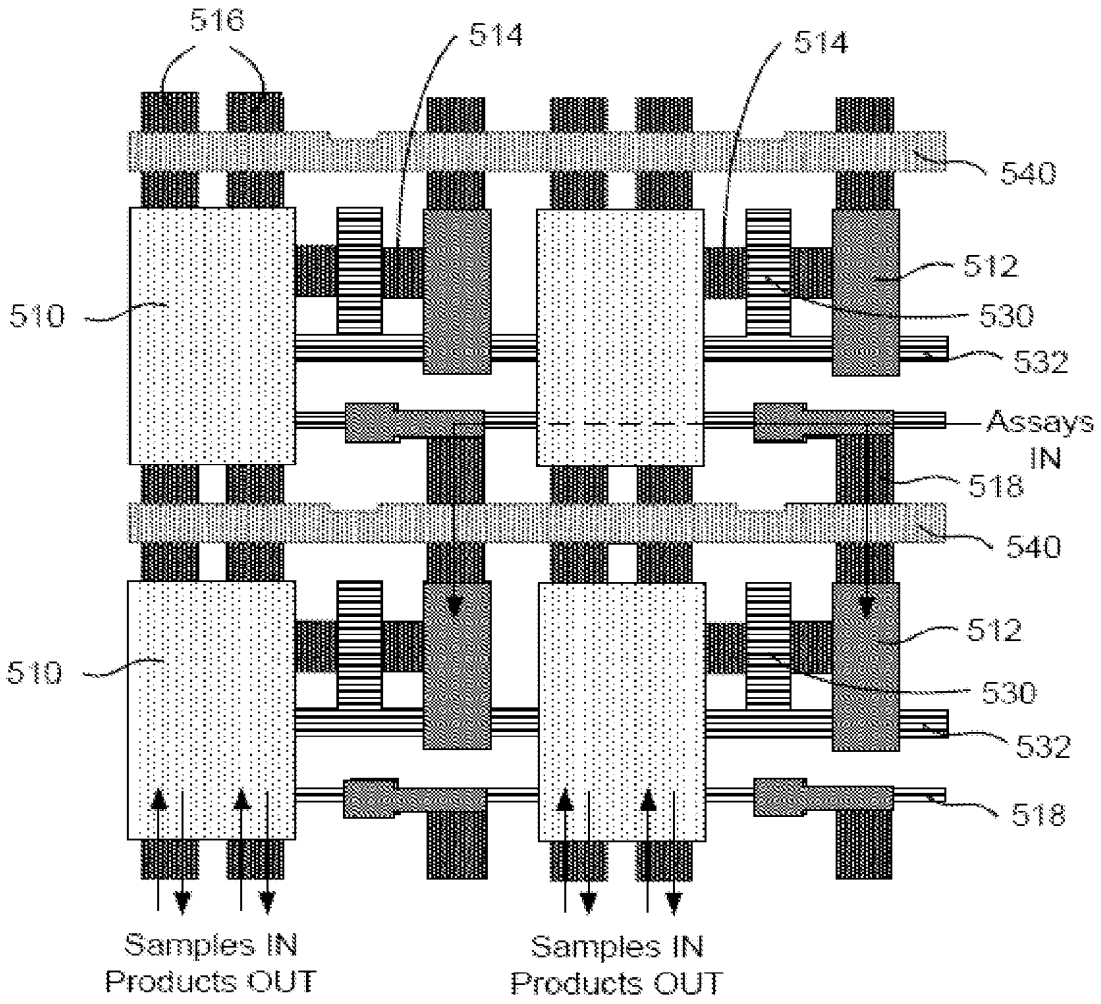

516
514
514
540
512
530
532
510
Assays IN
518
540
512
530
532
510
518
Samples IN
Products OUT
Samples IN
Products OUT

Fig. 6

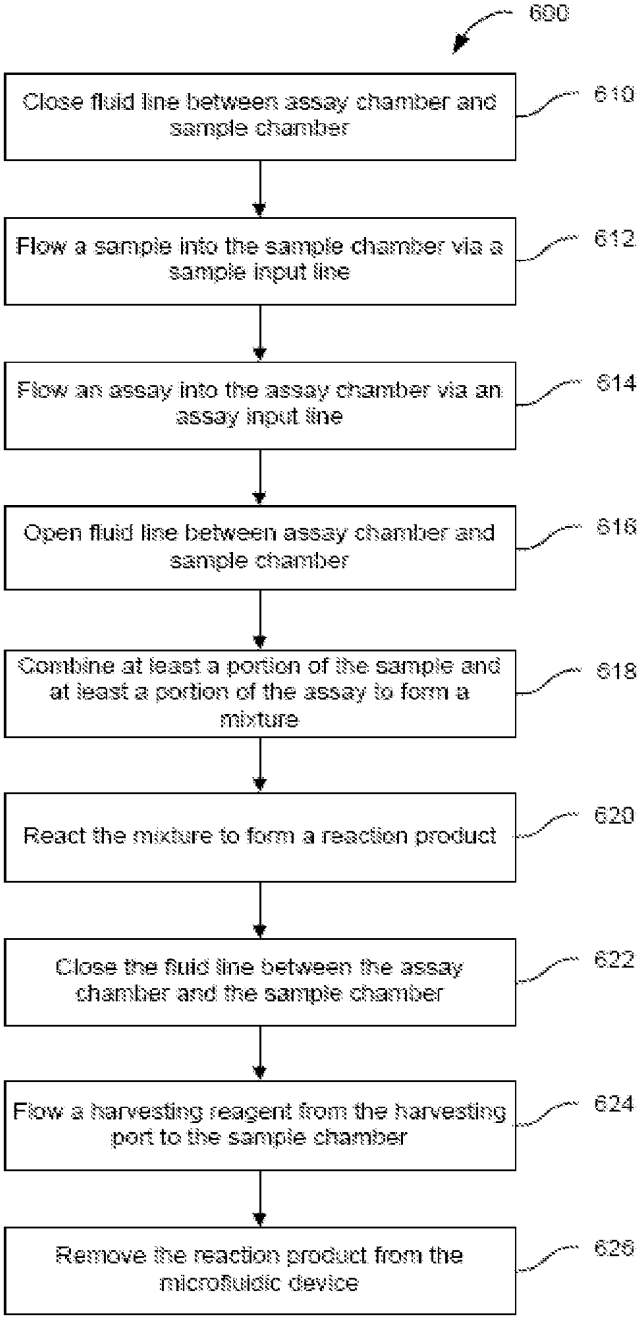

600
Close fluid line between assay chamber and sample chamber
610
Flow a sample into the sample chamber via a sample input line
612
Flow an assay into the assay chamber via an assay input line
614
Open fluid line between assay chamber and sample chamber
616
Combine at least a portion of the sample and at least a portion of the assay to form a mixture
618
React the mixture to form a reaction product
620
Close the fluid line between the assay chamber and the sample chamber
622
Flow a harvesting reagent from the harvesting port to the sample chamber
624
Remove the reaction product from the microfluidic device
626

Fig. 7

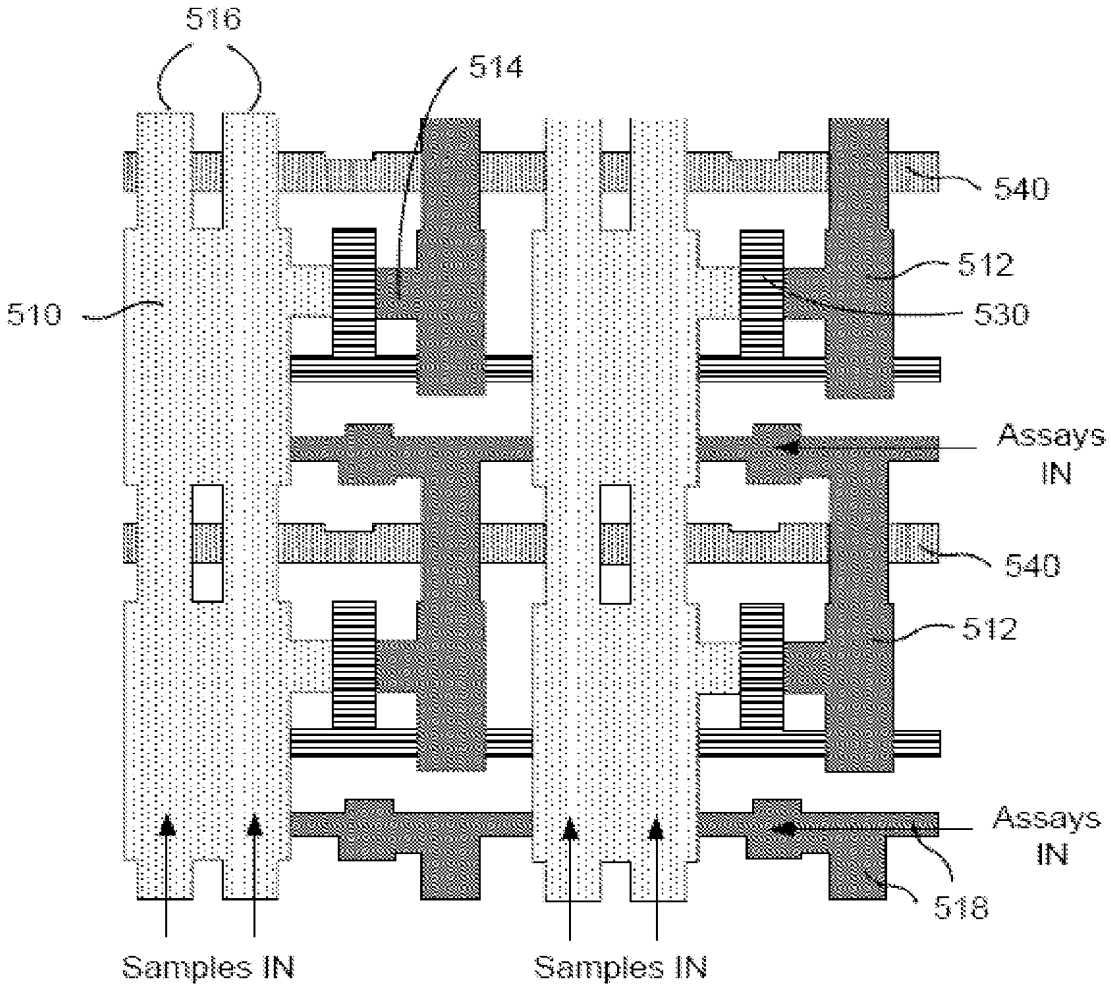

516
514
540
512
510
530
Assays IN
540
512
Assays IN
518
Samples IN
Samples IN

Fig. 8A

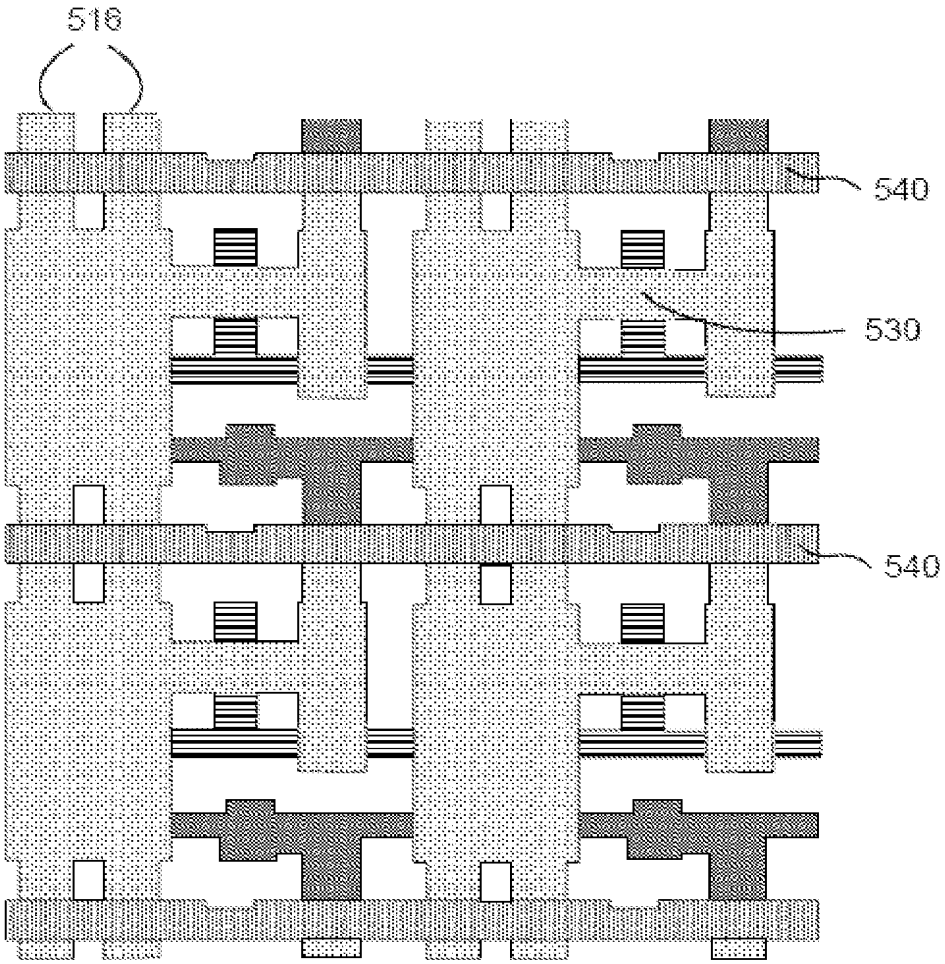


Fig. 8B

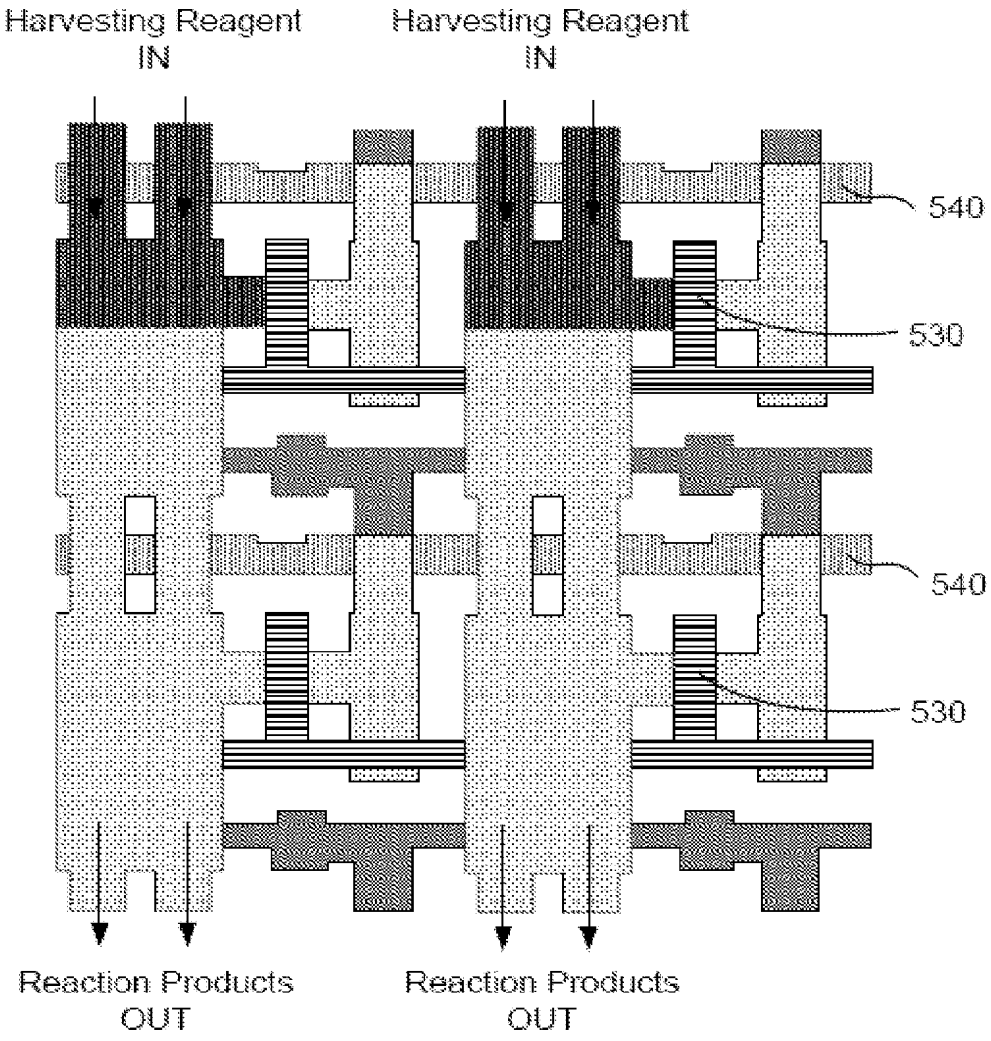

Harvesting Reagent IN
Harvesting Reagent IN
540
530
540
530
Reaction Products OUT
Reaction Products OUT

Fig. 8C

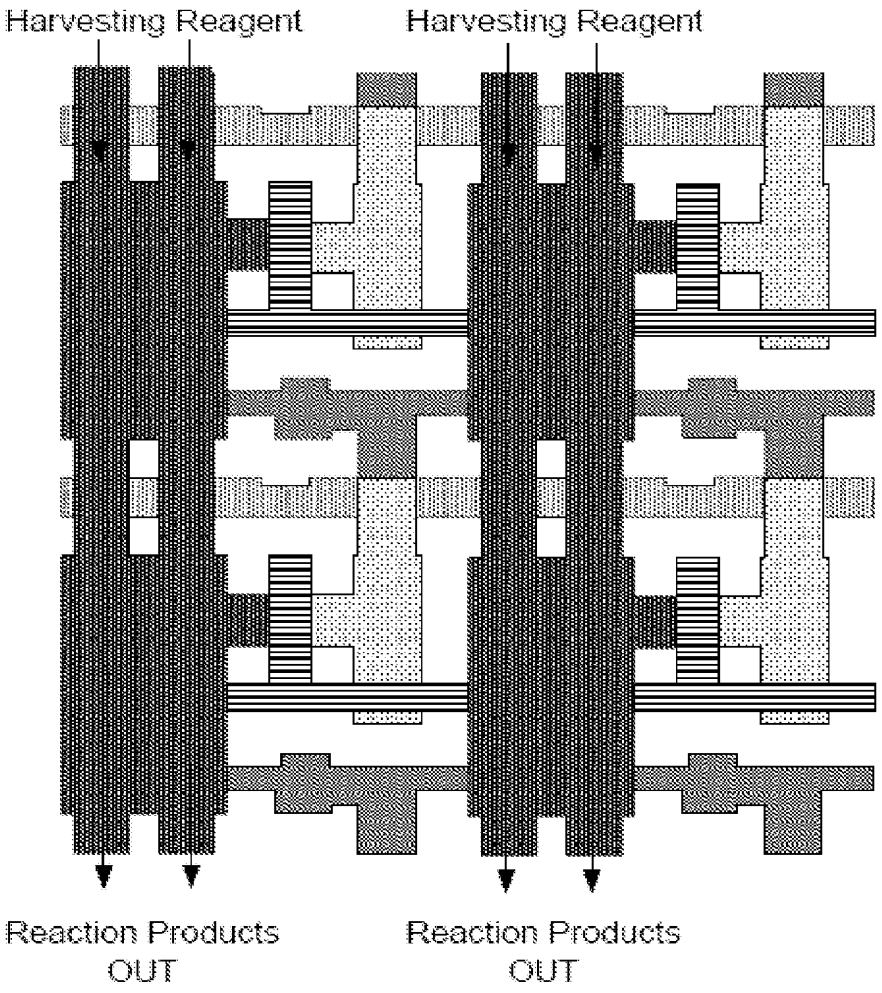

Harvesting Reagent
Harvesting Reagent
Reaction Products OUT
Reaction Products OUT

Fig. 8D

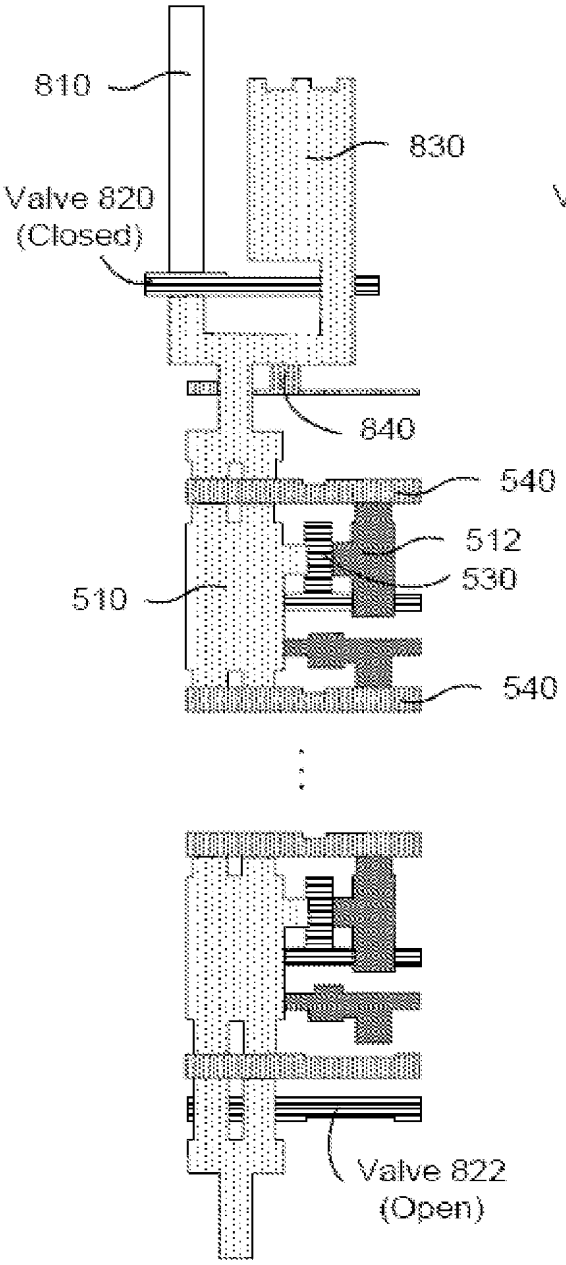

810
830
Valve 820
(Closed)
840
540
512
530
510
540
Valve 822
(Open)

Fig. 9A

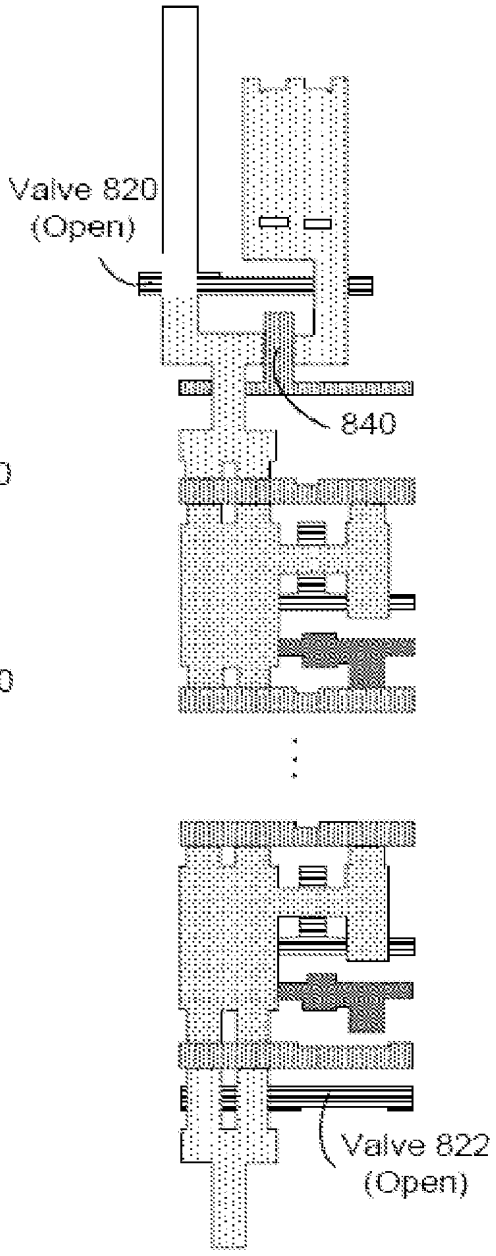

Valve 820
(Open)
840
Valve 822
(Open)

Fig. 9B

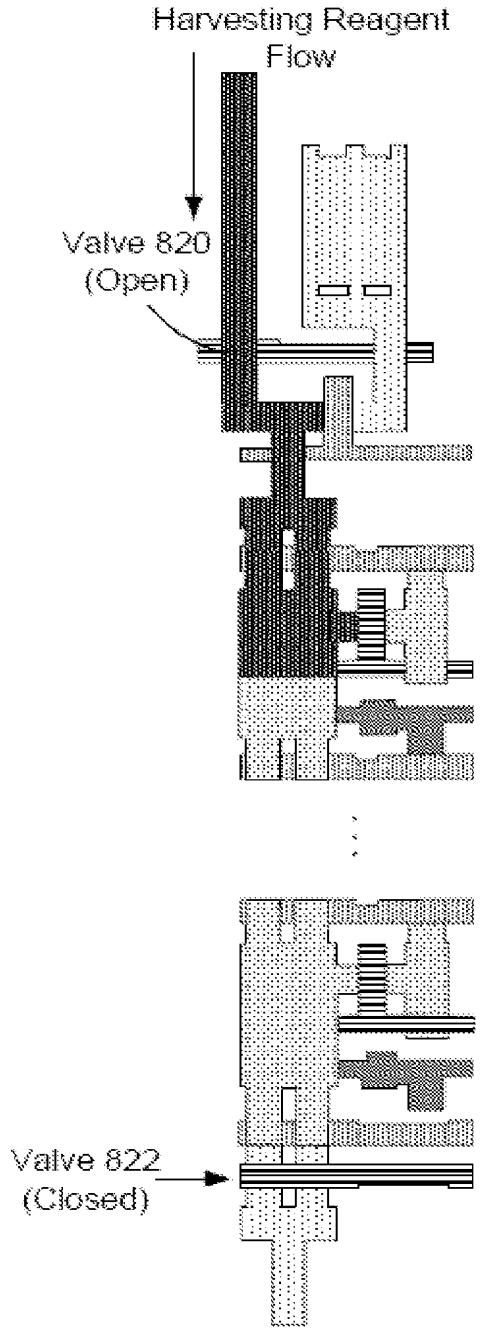

Harvesting Reagent Flow
Valve 820 (Open)
Valve 822 (Closed)

Fig. 9C

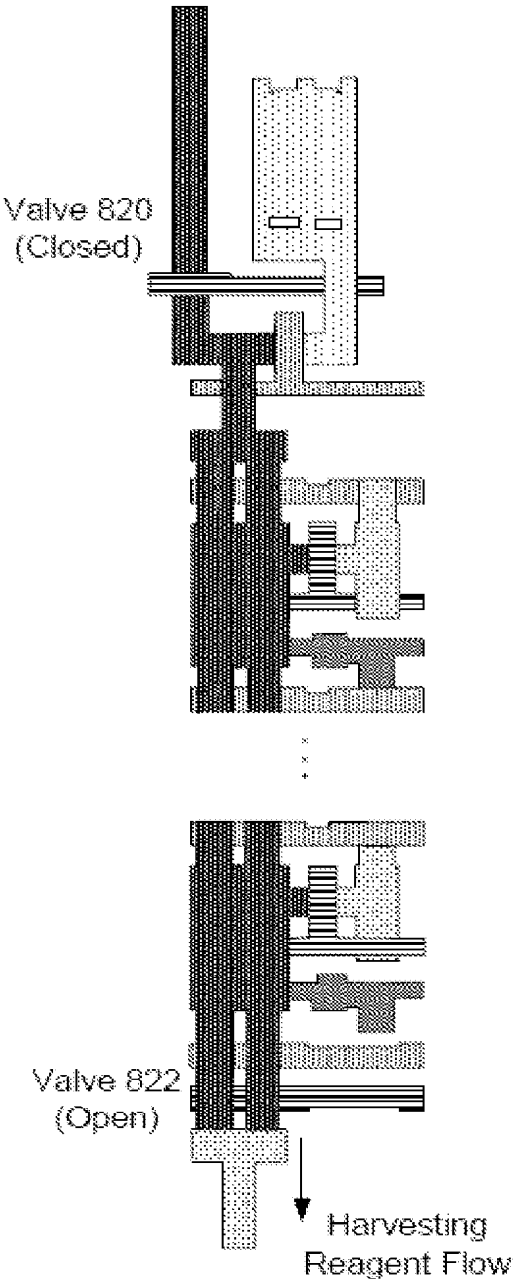

Valve 820 (Closed)
Valve 822 (Open)
Harvesting Reagent Flow

Fig. 9D

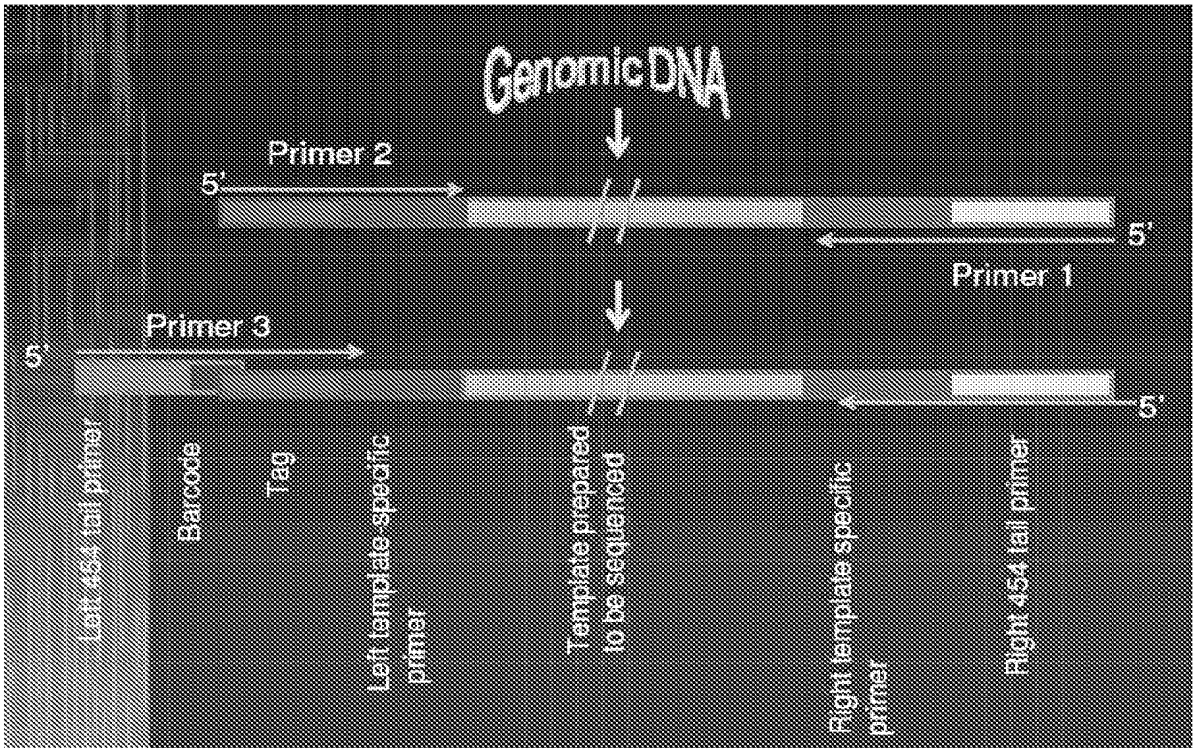

Genomic DNA
Primer 2
5'
5'
Primer 1
Primer 3
5'
5'
Left 454 tail primer
Barcode
Tag
Left template specific primer
Template prepared to be sequenced
Right template specific primer
Right 454 tail primer

Fig. 10

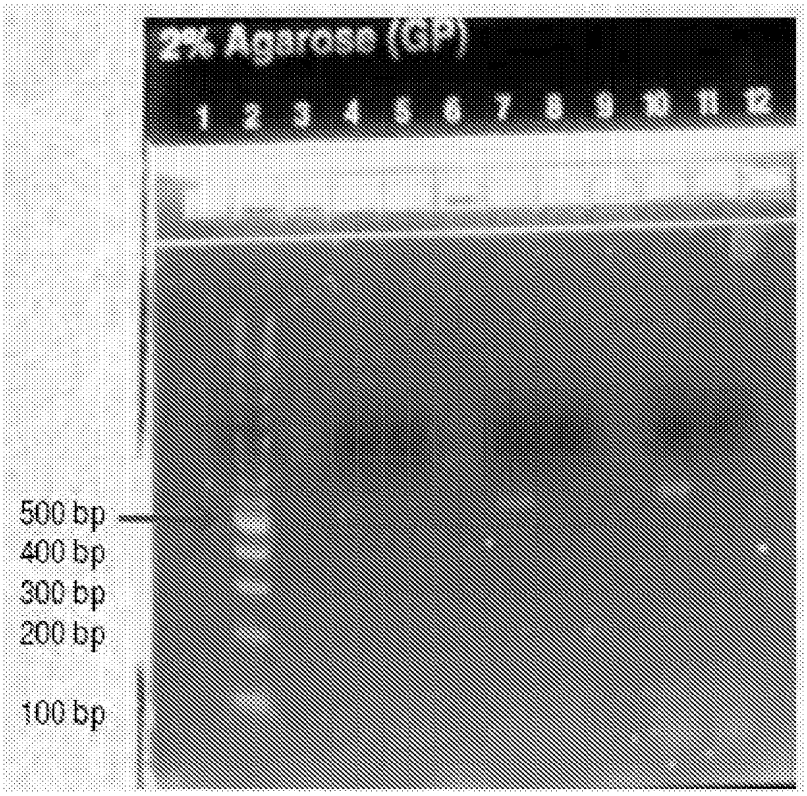

2% Agarose (GP)
1 2 3 4 5 6 7 8 9 10 11 12
500 bp
400 bp
300 bp
200 bp
100 bp

Fig. 11

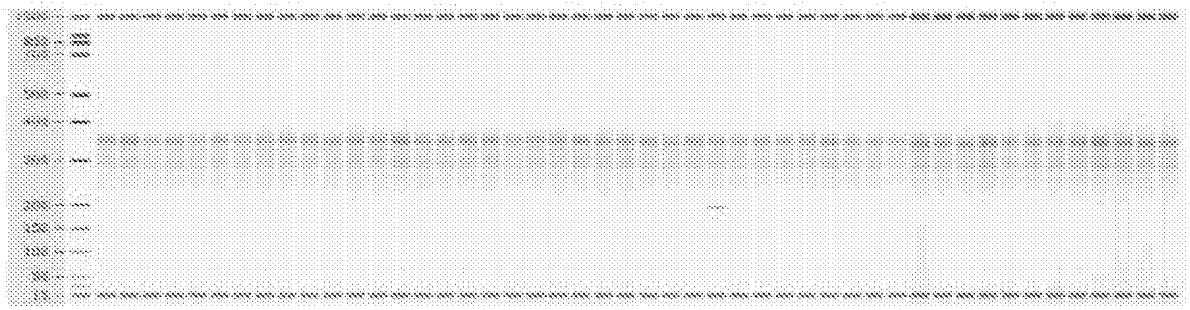
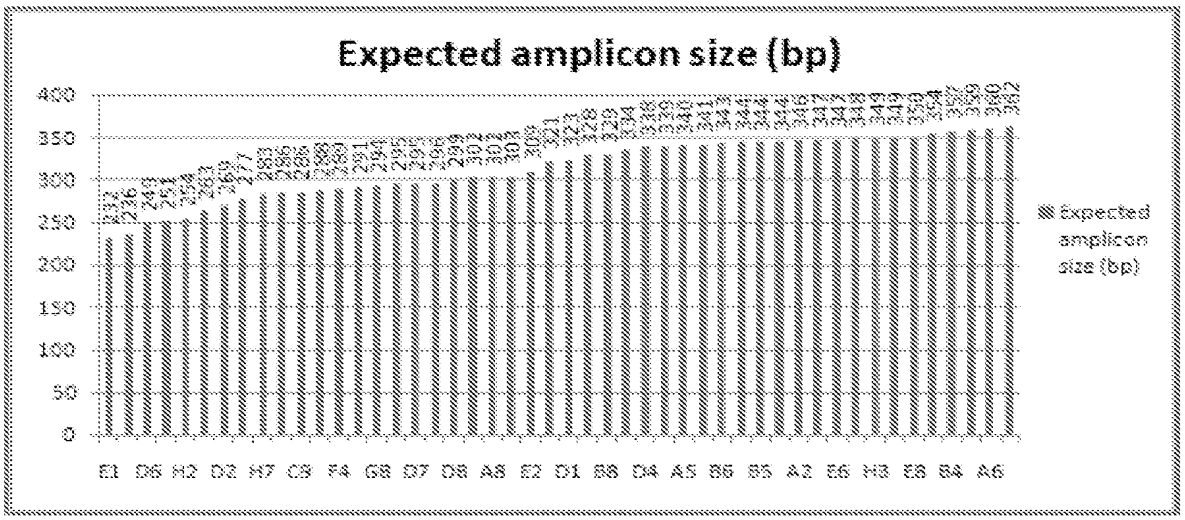


Fig. 12

A


Expected amplicon size (bp)
400
350
300
250
200
150
100
50
0
232 236 248 251 254 262 269 277 283 285 286 288 289 291 294 295 296 299 302 303 303 306 321 323 328 329 334 336 339 340 341 342 344 344 346 347 347 348 349 350 354 357 359 360 362
E1 D6 H2 D2 H7 C9 F4 G8 D7 D8 A8 E2 D1 B8 D4 A5 B6 B5 A2 E6 H3 E8 B4 A6
Expected amplicon size (bp)

B

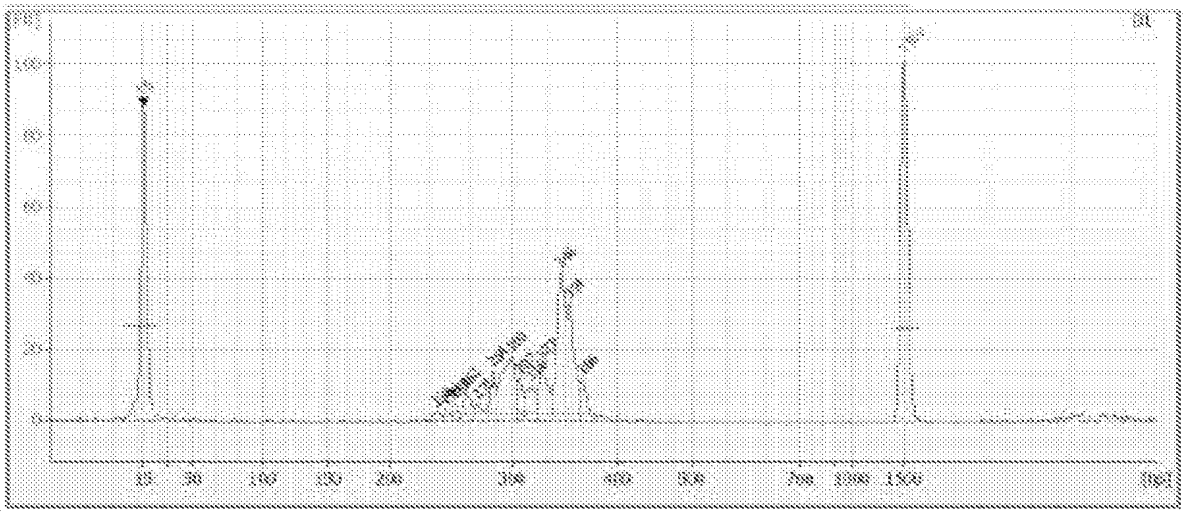
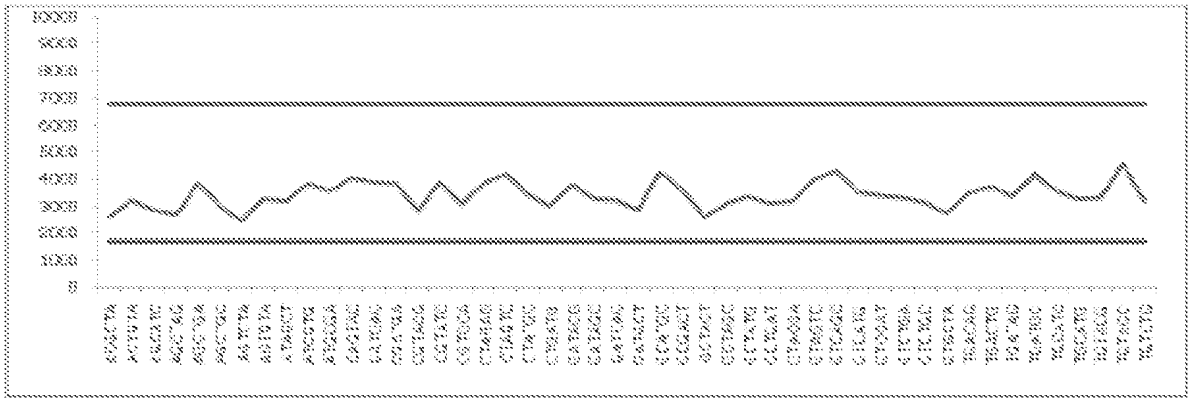


Fig. 13A-B

A



B

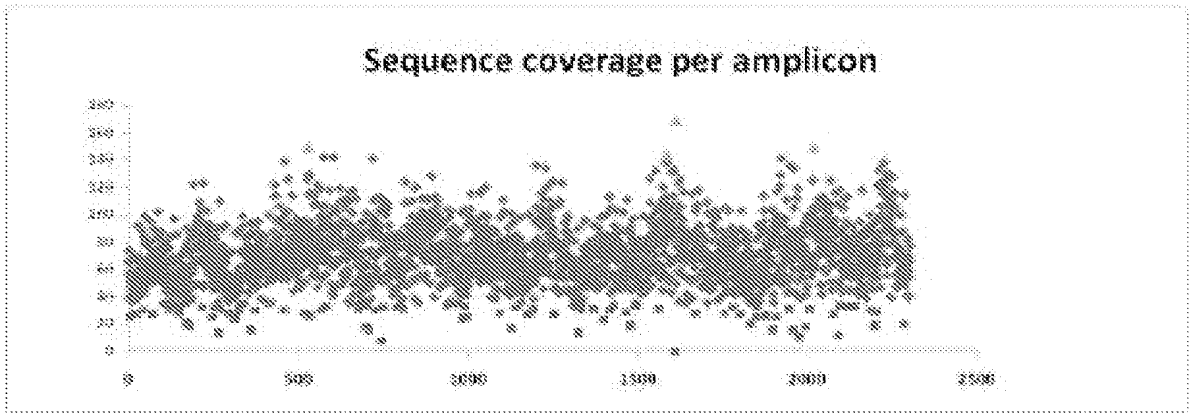


Fig. 14A-B

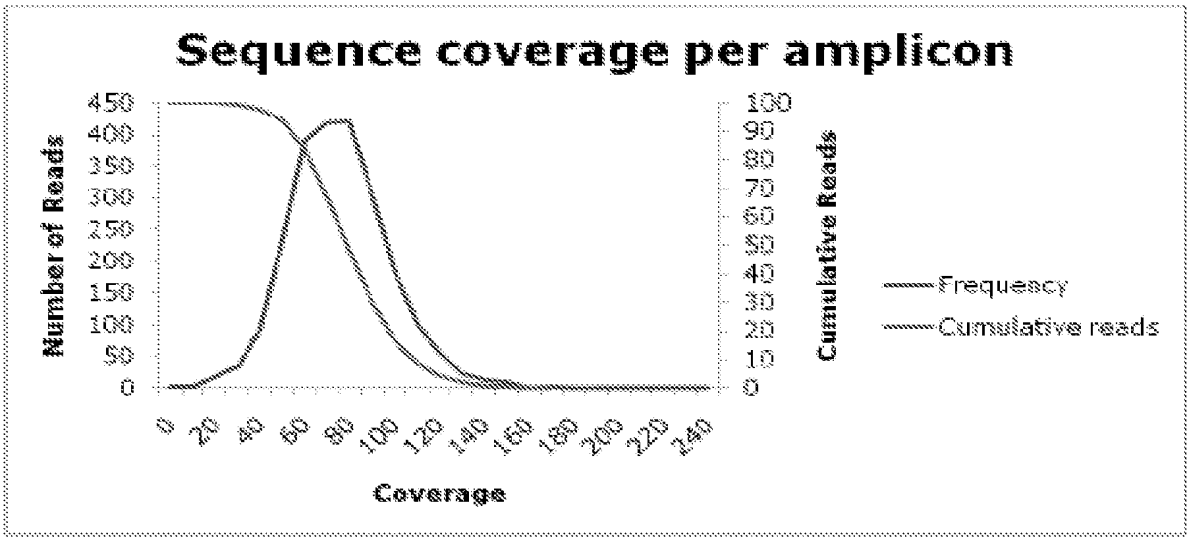
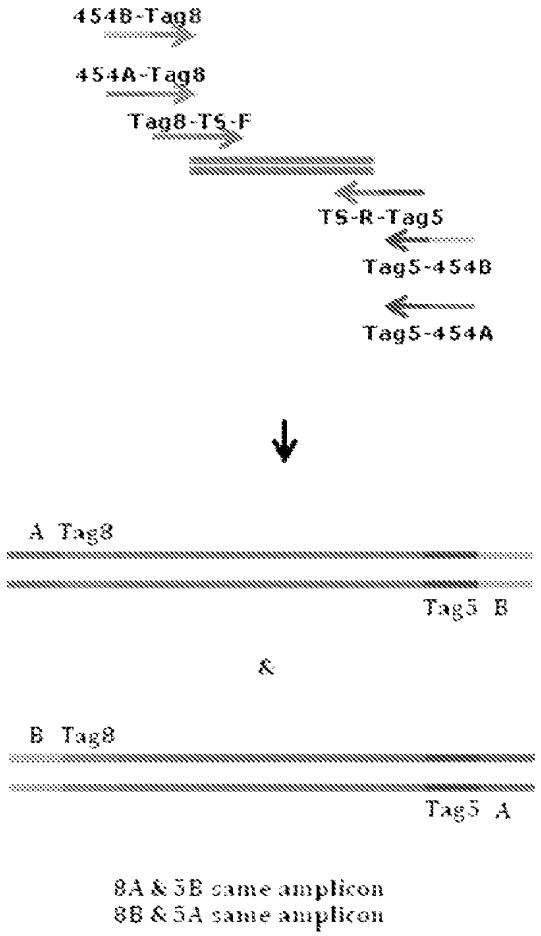

Sequence coverage per amplicon
Number of Reads
Cumulative Reads
Coverage
Frequency
Cumulative reads
0
20
40
60
80
100
120
140
160
180
200
220
240

Fig. 14C

A


454B-Tag8
454A-Tag8
Tag8-TS-F
TS-R-Tag5
Tag5-454B
Tag5-454A
A Tag8
Tag5 B
&
B Tag8
Tag5 A
8A & 5B same amplicon
8B & 5A same amplicon

B

Fig. 15A-B

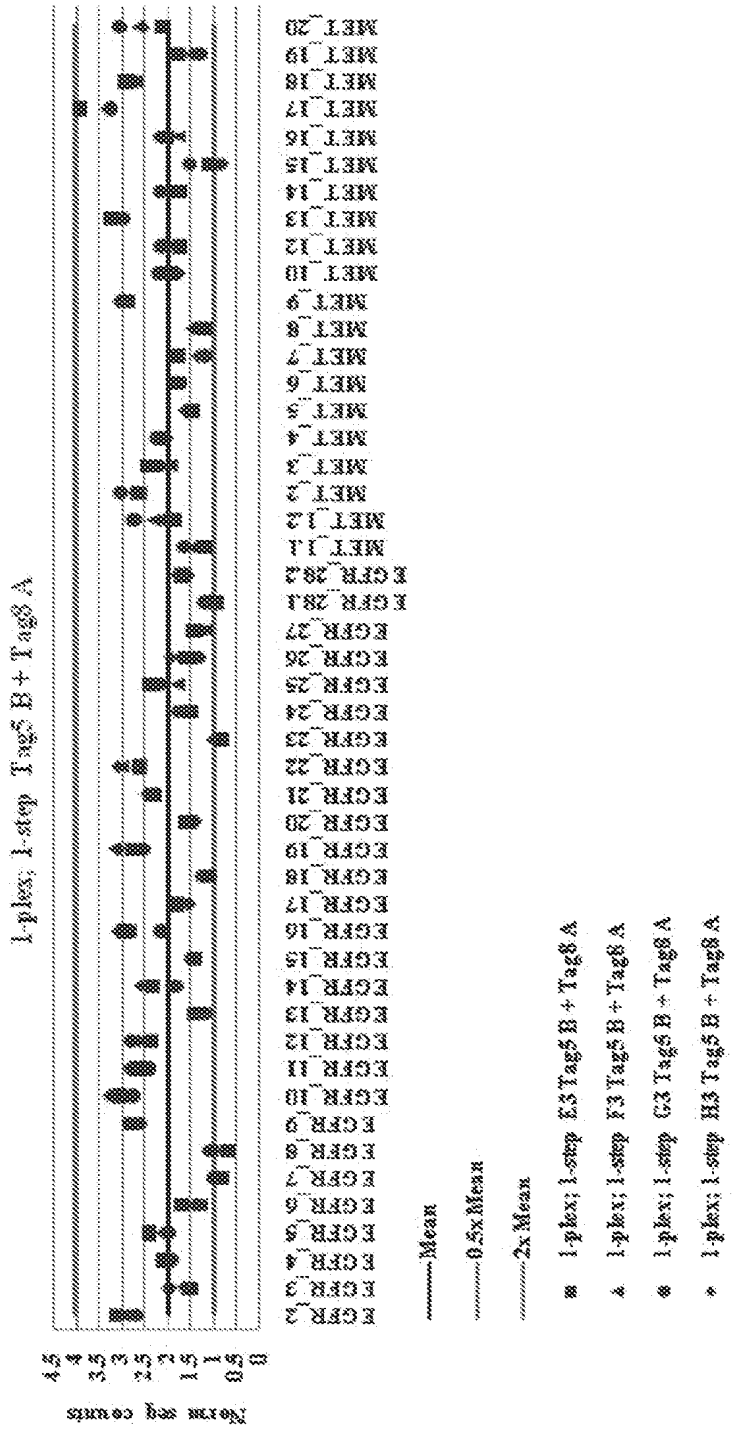

1-plex; 1-step Tag5 B + Tag8 A
Norm eq counts
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
EGFR_2
EGFR_3
EGFR_4
EGFR_5
EGFR_6
EGFR_7
EGFR_8
EGFR_9
EGFR_10
EGFR_11
EGFR_12
EGFR_13
EGFR_14
EGFR_15
EGFR_16
EGFR_17
EGFR_18
EGFR_19
EGFR_20
EGFR_21
EGFR_22
EGFR_23
EGFR_24
EGFR_25
EGFR_26
EGFR_27
EGFR_28.1
EGFR_28.2
MET_1.1
MET_1.2
MET_2
MET_3
MET_4
MET_5
MET_6
MET_7
MET_8
MET_9
MET_10
MET_12
MET_13
MET_14
MET_15
MET_16
MET_17
MET_18
MET_19
MET_20
Mean
0.5x Mean
2x Mean
1-plex; 1-step E3 Tag5 B + Tag8 A
1-plex; 1-step F3 Tag5 B + Tag8 A
1-plex; 1-step G3 Tag5 B + Tag8 A
1-plex; 1-step H3 Tag5 B + Tag8 A

Fig. 16A

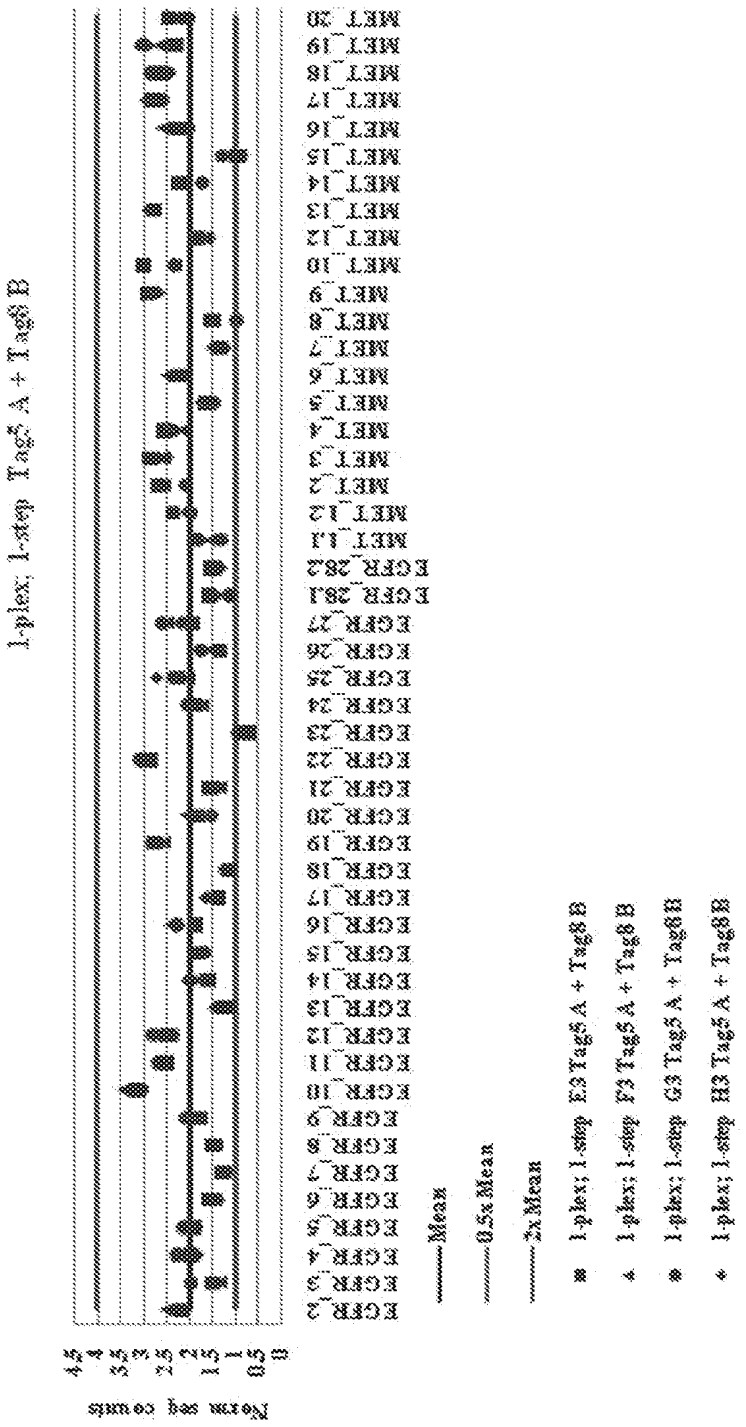

1-plex; 1-step Tag5 A + Tag8 B
Norm eq counts
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
EGFR_2
EGFR_3
EGFR_4
EGFR_5
EGFR_6
EGFR_7
EGFR_8
EGFR_9
EGFR_10
EGFR_11
EGFR_12
EGFR_13
EGFR_14
EGFR_15
EGFR_16
EGFR_17
EGFR_18
EGFR_19
EGFR_20
EGFR_21
EGFR_22
EGFR_23
EGFR_24
EGFR_25
EGFR_26
EGFR_27
EGFR_28J
EGFR_28Z
MET_1J
MET_1Z
MET_2
MET_3
MET_4
MET_5
MET_6
MET_7
MET_8
MET_9
MET_10
MET_12
MET_13
MET_14
MET_15
MET_16
MET_17
MET_18
MET_19
MET_20
Mean
0.5x Mean
2x Mean
1-plex; 1-step E3 Tag5 A + Tag8 B
1-plex; 1-step F3 Tag5 A + Tag8 B
1-plex; 1-step G3 Tag5 A + Tag8 B
1-plex; 1-step H3 Tag5 A + Tag8 B

Fig. 16B

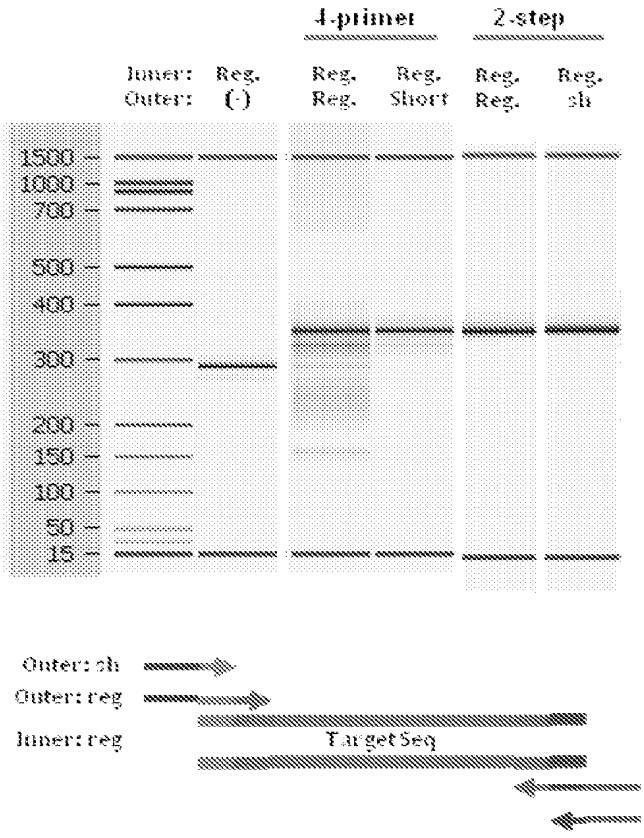

4-primer
2-step
Inner: Reg. Reg. Reg. Reg.
Outer: (-) Reg. Short Reg. sh
1500
1000
700
500
400
300
200
150
100
50
15
Outer: sh
Outer: reg
Inner: reg
Target Seq

Fig. 17

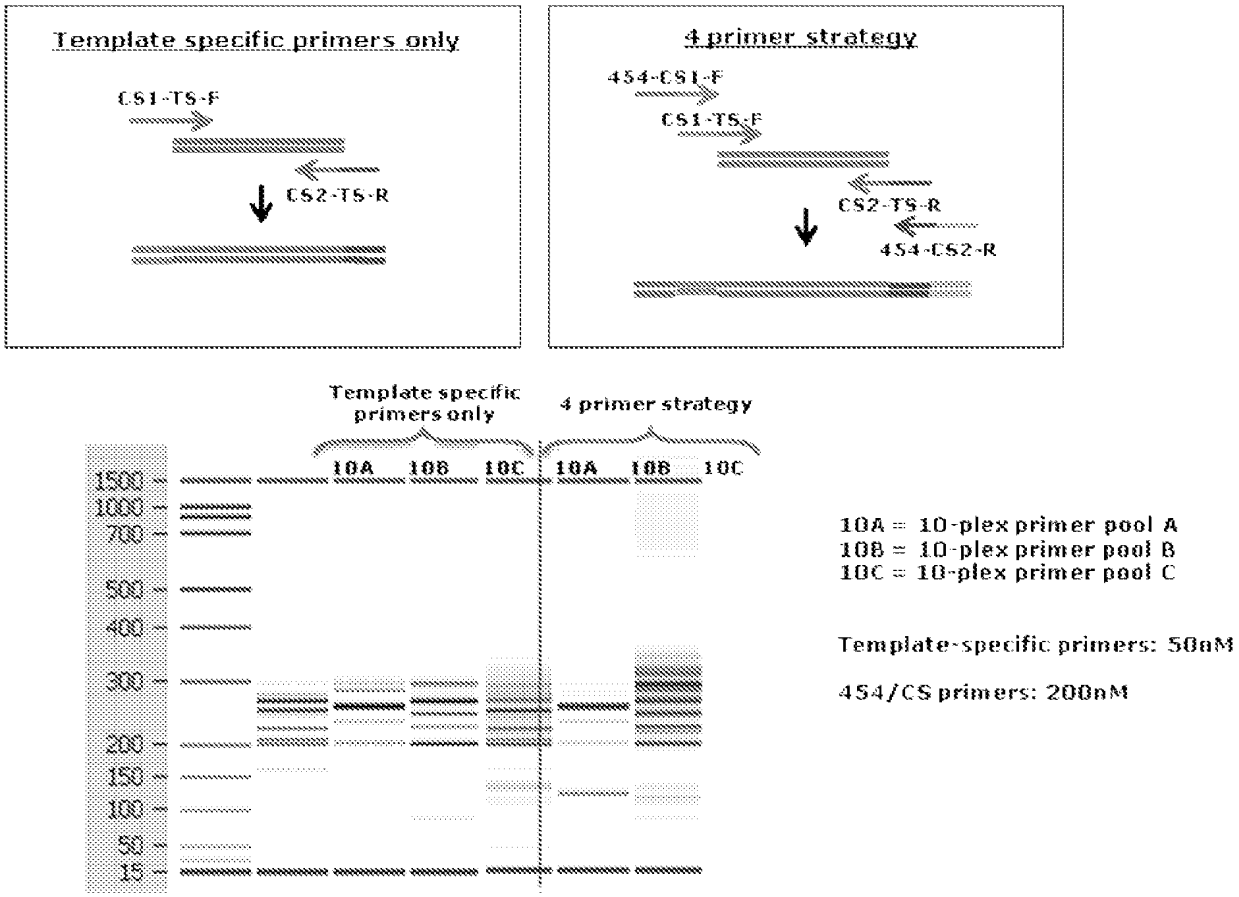

Template specific primers only
CS1-TS-F
CS2-TS-R
4 primer strategy
454-CS1-F
CS1-TS-F
CS2-TS-R
454-CS2-R
Template specific primers only
4 primer strategy
10A
10B
10C
10A
10B
10C
1500
1000
700
500
400
300
200
150
100
50
15
10A = 10-plex primer pool A
10B = 10-plex primer pool B
10C = 10-plex primer pool C
Template-specific primers: 50nM
454/CS primers: 200nM

Fig. 18

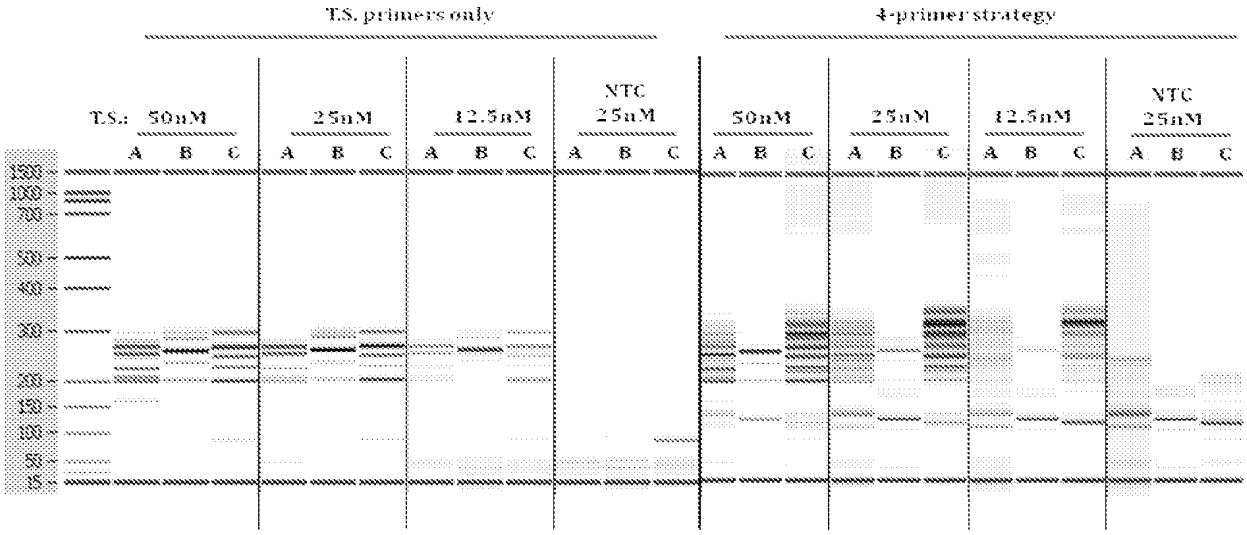

T.S. primers only
4-primer strategy
T.S.: 50nM
25nM
12.5nM
NTC 25nM
50nM
25nM
12.5nM
NTC 25nM
A B C
120
100
80
60
40
30
25
20
15
10
5

Fig. 19

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/29854

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/68 (2010.01)
USPC - 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC- 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC- 435/91.2

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest (PGPB,USPT,EPAB,JPAB): MAY ANDREW; CHEN PEILIN; WANG JUN; KAPER FIONA; ANDERSON MEGAN; BARCODE IDENTIFICATION, "ID", SEQUENCE, NUCLEOTIDE, PCR

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2009/0053719 A1 (Lo et al.) 26 February 2009 (26.02.2009) para [0008], [0012], [0024], [0019]	11-12, 16-18, 24, 26-33, 38-56
Y	US 4,683,202 A (Mullis) 28 July 1987 (28.07.1987) col 2, ln 26-40, 51-59; col 4, ln 51-58; fig 4-1; col 14, ln 8-13	1-18, 24, 26-56
Y	US 2007/0020640 A1 (McCloskey et al.) 25 January 2007 (25.01.2007) para [0008], [0002], [0084]	1-16, 34-37
Y	US 2008/0108063 A1 (Lucero et al.) 8 May 2008 (08.05.2008) para [0009]	28, 46
Y	US 5,066,584 A (Gyllensten et al.) 19 November 1991 (19.11.1991) col 3, ln 68 to col 4, ln 2	1-16, 34-37

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent 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 when the document is taken alone
"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

Date of the actual completion of the international search	Date of mailing of the international search report
9 August 2010 (09.08.2010)	30 AUG 2010
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/29854

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 19-23 and 25
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Group I: Claims 1-18, 24, 26-56 are directed to a method of amplifying a plurality of target nucleic acids in a plurality of samples.
Group II: Claims 57-81 are directed to a microfluidic device a method of using said device.
Group III: Claims 82-95 are directed to a method of preparing reaction products.

The inventions listed as Groups I - III do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is a method of amplifying a plurality of target nucleic acids in a plurality of samples. The special technical feature of Group I is not present in Group II or Group III.

The special technical feature of Group II is a microfluidic device. The special technical feature of Group II is not present in Group I or Group III.

********** See Supplemental Box to continue **********

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-18, 24, and 26-56

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

IDT Exhibit 1020.2775

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/29854

********** See Supplemental Box **********

In continuation of Box III: Observations where unity of invention is lacking:

The special technical feature of Group III is a method of preparing reaction products. The special technical feature of Group III is not present in Group I or Group II.

Accordingly, the inventions listed as Groups I - III lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

IDT Exhibit 1020.2776

Espacenet

Bibliographic data: WO2008005459 (A2) — 2008-01-10

METHODS FOR FRAGMENTATION AND LABELING OF NUCLEIC ACIDS

Inventor(s): KURN NURITH [US]; CHEN PENGCHIN [US] ± (KURN, NURITH, ; CHEN, PENGCHIN)

Applicant(s): NUGEN TECHNOLOGIES INC [US]; KURN NURITH [US]; CHEN PENGCHIN [US] ± (NUGEN TECHNOLOGIES, INC, ; KURN, NURITH, ; CHEN, PENGCHIN)

Classification: **- international:** ***C12Q1/68***
- cooperative: **C12Q1/6806**; **C12Q1/6811** more

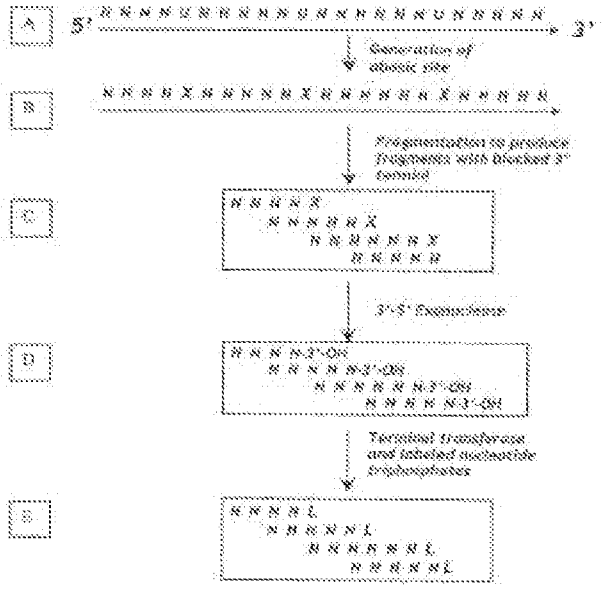
Application number: WO2007US15409 20070702 Global Dossier

Priority number (s): US20060817890P 20060630

Also published as: WO2008005459 (A3) CA2656315 (A1) EP2038429 (A2) EP2038429 (A4) EP2038429 (B1) more

Abstract of WO2008005459 (A2)

The invention provides methods, compositions, and kits for fragmentation and labeling of nucleic acids. More particularly, the invention relates to methods for fragmentation of nucleic acids to produce fragments with 3' end hydroxyl groups within a desired size range. In methods of the invention, nucleic acids are fragmented at abasic sites to produce fragments with blocked 3' ends. The 3' ends are unblocked to produce polynucleotide fragments with hydroxyl groups at their 3' ends. Methods, kits, and compositions for carrying out fragmentation of a polynucleotide template in a single reaction mixture to



IDT Exhibit 1020.2777

yield fragments with 3'-hydroxi ends within the desired size range are disclosed.

IDT Exhibit 1020.2778

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 January 2008 (10.01.2008)

PCT

(10) International Publication Number
WO 2008/005459 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2007/015409

(22) International Filing Date: 2 July 2007 (02.07.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/817,890 30 June 2006 (30.06.2006) US

(71) Applicant *(for all designated States except US)*: **NUGEN TECHNOLOGIES, INC.** [US/US]; 821 Industrial Road, Unit A, San Carlos, CA 94070 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KURN, Nurith** [IL/US]; 2876 Ramona Street, Palo Alto, CA 94306 (US). **CHEN, Pengchin** [US/US]; 1073 Belvedere Lane, San Jose, CA 95129 (US).

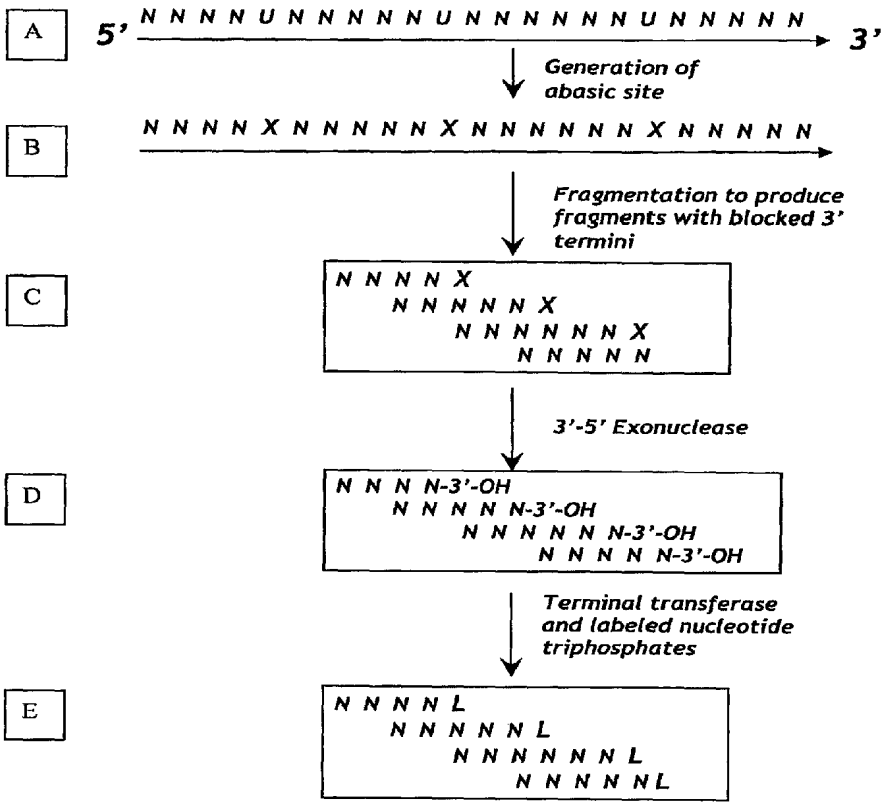
(74) Agent: **REAMEY, Robert, H.**; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA 94304 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, 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, 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, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHODS FOR FRAGMENTATION AND LABELING OF NUCLEIC ACIDS



(57) Abstract: The invention provides methods, compositions, and kits for fragmentation and labeling of nucleic acids. More particularly, the invention relates to methods for fragmentation of nucleic acids to produce fragments with 3' end hydroxyl groups within a desired size range. In methods of the invention, nucleic acids are fragmented at abasic sites to produce fragments with blocked 3' ends. The 3' ends are unblocked to produce polynucleotide fragments with hydroxyl groups at their 3' ends. Methods, kits, and compositions for carrying out fragmentation of a polynucleotide template in a single reaction mixture to yield fragments with 3'-hydroxl ends within the desired size range are disclosed.

WO 2008/005459 A2

IDT Exhibit 1020.2779

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHODS FOR FRAGMENTATION AND LABELING OF NUCLEIC ACIDS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application No. 60/817,890, filed June 30, 2006, which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] The invention relates to methods for fragmentation of nucleic acids, in particular to generate fragments with 3' end hydroxyl groups, and methods for labeling the fragments.

BACKGROUND

[0003] Nucleic acid fragmentation and labeling has gained importance in the field of nucleic acid analysis. Fragmented and labeled nucleic acids are useful for efficient hybridization based nucleic acid analysis, particularly when hybridizing to immobilized probes, such as in multiplex detection using microarrays or bead based hybridization assays. Fragmentation of nucleic acids to generate fragments having hydroxyl groups at their 3' termini is useful since a 3' hydroxyl group permits extension with a template-dependent polymerase or labeling with a template-independent polymerase, which permits introduction of a label. The extension of fragmented, hybridized nucleic acid along a nucleic acid probe or template molecule by target directed nucleic acid synthesis is useful for both nucleic acid analysis and the generation of recombinant nucleic acids.

[0004] A method for controlled nucleic acid fragmentation has been previously described which is based on the incorporation of non-canonical nucleotides into a polynucleotide strand which is synthesized *in vitro*, followed by generation of

IDT Exhibit 1020.2781

an abasic site at the site of incorporation of a non-canonical nucleotide, permitting subsequent fragmentation of the synthesized polynucleotide and/or labeling at the abasic site. (U.S. Patent Application No. 2004/0005614; PCT Application No. WO 04/011665). The size distribution of the fragmented products may be controlled by the level of incorporation of non-canonical nucleotides. The level of incorporation, and subsequent nucleic acid fragment size, may be adjusted to provide a suitable fragment size for the particular downstream use of the fragmented products.

[0005] A process utilizing dUTP as the non-canonical nucleotide during DNA synthesis, uracil N-glycosylase ("UNG") as the enzyme which removes the base portions of the non-canonical nucleotides to generate abasic sites, and cleavage of abasic sites with a polyamine, such as N, N'-dimethylethylenediamine, as described in U.S. Application No. 2004/0005614 and PCT Application No. WO 04/011665, results in the generation of fragmented polynucleotides with modified ("blocked") 3' ends that are capable of reacting with an aldehyde-reactive reagent. This process is useful for generation of fragmented and labeled nucleic acids suitable for analysis, for example, on a microarray, or fragmented nucleic acid targets suitable for immobilization.

[0006] The process of controlled fragmentation and labeling of single stranded nucleic acids is difficult to achieve using other methods, such as non-specific digestion with an enzyme such as DNase or chemical nucleotide modification, since these reactions cannot be carried out to completion without complete, or nearly complete, degradation of the single stranded nucleic acid to be fragmented and labeled. Double stranded DNA may be digested with restriction endonucleases to generate fragments of defined size distribution. Such enzymes are specific for a defined sequence content, both with respect to composition and length of a recognition site, and afford higher or lower frequency of cleavage depending on the restriction enzyme used. However, restriction endonucleases are specific for double stranded DNA and this approach does not apply to fragmentation of single stranded nucleic acid molecules.

[0007] A desirable feature for fragmented nucleic acids is the presence of a hydroxyl group at the 3' end, which may serve as a substrate for polynucleotide

IDT Exhibit 1020.2782

synthesis with a polymerase via template-dependent or template-independent extension, or ligation with another polynucleotide. Such manipulations may facilitate introduction of a label. For example, a label may be introduced by extending from the 3' end with terminal transferase (a template-independent polymerase) and a labeled nucleotide. The method described above, in which nucleic acid molecules are cleaved at an abasic site with a polyamine, results in nucleic acids with "blocked" 3' ends, precluding use of the fragments as primers for polynucleotide synthesis or substrates for labeling at a 3' hydroxyl group.

[0008] New methods for the efficient fragmentation of nucleic acids to generate fragments comprising 3' end hydroxyl groups, and of defined size distribution, are desirable.

BRIEF SUMMARY OF THE INVENTION

[0009] The invention provides methods, compositions, and kits for fragmenting, or fragmenting and labeling, a polynucleotide.

[0010] In one aspect, the invention features the generation and/or use of a polynucleotide with abasic sites distributed such that cleavage at the abasic sites generates fragments of a desired size range.

[0011] In one aspect, the invention provides a method for fragmenting a polynucleotide, said method comprising: (a) chemically cleaving a phosphodiester backbone of a polynucleotide comprising an abasic site at or near the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; and (b) contacting the polynucleotide fragment with an enzyme capable of unblocking (*i.e.*, which unblocks) the blocked 3' end of said fragment, whereby a polynucleotide fragment comprising a 3' end hydroxyl group is generated. In some embodiments, steps (a) and (b) are performed simultaneously in the same reaction mixture. In some embodiments, fragments within a desired size range are generated.

[0012] In some embodiments, the phosphodiester backbone is cleaved with a polyamine to generate a polynucleotide fragment with a blocked 3' end. In one embodiment, the polyamine is N, N'-dimethylethylenediamine (DMED).

IDT Exhibit 1020.2783

[0013] In some embodiments, the enzyme capable of unblocking the blocked 3' end comprises a 3' to 5' exonuclease activity. In one embodiment, the exonuclease activity is a non-processive exonuclease activity. In one embodiment, the enzyme that comprises an exonuclease activity does not comprise an endonuclease activity. In one embodiment, the enzyme that comprises an exonuclease activity also comprises an endonuclease activity, and contacting the polynucleotide fragment with the enzyme is under conditions in which the endonuclease activity is minimized or absent. In one embodiment, the enzyme comprising a 3' exonuclease activity is selected from the group consisting of endonuclease 4, exonuclease T, and apurinic/apyrimidinic endonuclease (APE 1).

[0014] In some embodiments, the method further comprises extending the polynucleotide fragment from the 3' hydroxyl group with a template independent polymerase and a labeled nucleotide, whereby a polynucleotide fragment labeled at the 3' end is generated. In one embodiment, the template independent polymerase is terminal deoxynucleotidyl transferase (TdT). In one embodiment, the labeled nucleotide is a biotinylated nucleotide. In various embodiments, the biotinylated nucleotide is selected from the group consisting of a biotinylated nucleotide triphosphate (NTP), a biotinylated deoxynucleotide triphosphate (dNTP), and a biotinylated dideoxynucleotide triphosphate (ddNTP). In some embodiments, the biotinylated nucleotide is selected from the group consisting of biotin 2',3'-dideoxy-UTP and biotin 2',3'-dideoxy-CTP. In one embodiment, the labeled nucleotide comprises a fluorophore. In one embodiment, a mixture of labeled and unlabeled nucleotides is used for labeling the polynucleotide fragment. As used herein, the term "nucleotide" encompasses nucleotide analogs, which are known in the art. The term "labeled nucleotide" encompasses labeled nucleotide analogs.

[0015] In some embodiments, the polynucleotide comprising an abasic site is generated by: (i) synthesizing a polynucleotide from a polynucleotide template in the presence of a non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated; and (ii) cleaving a base portion of the non-canonical nucleotide from the synthesized polynucleotide with an enzyme capable of cleaving (*i.e.*, which cleaves) the base portion of the non-canonical nucleotide,

IDT Exhibit 1020.2784

whereby an abasic site is generated. In some embodiments, the method involves synthesizing the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range. In various embodiments, the non-canonical nucleotide is selected from the group consisting of dUTP, dITP, and 5-OH-Me-dCTP. In one embodiment, the enzyme capable of cleaving a base portion of the non-canonical nucleotide is an N-glycosylase. In some embodiments, the N-glycosylase is selected from the group consisting of Uracil N-Glycosylase (UNG), hypoxanthine-N-Glycosylase, and hydroxy-methyl cytosine-N-glycosylase. In one embodiment, the non-canonical nucleotide is dUTP and the enzyme capable of cleaving a base portion of the non-canonical nucleotide is UNG. In one embodiment, the non-canonical nucleotide is dUTP, the enzyme capable of cleaving a base portion of the non-canonical nucleotide is UNG, and the phosphodiester backbone is cleaved with DMED. In one embodiment, the polynucleotide comprising a non-canonical nucleotide is synthesized using a primer comprising a non-canonical nucleotide. In one embodiment, the polynucleotide comprising a non-canonical nucleotide is synthesized in the presence of two or more different non-canonical nucleotides, whereby a polynucleotide comprising two or more different non-canonical nucleotides is synthesized. In one embodiment, polynucleotides comprising a non-canonical nucleotides are synthesized from two or more different polynucleotide templates. In some embodiments, the polynucleotide comprising an abasic site is generated by non-enzymatically converting a canonical or non-canonical nucleotide in a polynucleotide into an abasic site. Exemplary non-enzymatic methods for generating an abasic site include depurination or depyrimidination of a nucleotide using an acidic pH, an oxidizing agent, an alkylating agent, and any two or more of the foregoing.

[0016] In some embodiments, the polynucleotide comprising an abasic site is generated by cleaving a base portion of a methylated nucleotide with an agent capable of cleaving (*i.e.*, which cleaves) a base portion of the methylated nucleotide to create an abasic site, whereby an abasic site is generated. In some embodiments, the method includes cleaving a base portion of a methylated nucleotide in a polynucleotide with

IDT Exhibit 1020.2785

an agent capable of cleaving a base portion of the methylated nucleotide to create an abasic site, whereby the polynucleotide comprising an abasic site is generated.

[0017] In some embodiments, the polynucleotide comprising an abasic site is generated by cleaving a base portion of a canonical nucleotide with an agent capable of cleaving (*i.e.*, which cleaves) a base portion of the canonical nucleotide to create an abasic site, whereby an abasic site is generated. In some embodiments, the method includes cleaving a base portion of a canonical nucleotide in a polynucleotide with an agent capable of cleaving a base portion of the canonical nucleotide to create an abasic site, whereby the polynucleotide comprising an abasic site is generated. In one embodiment, the canonical nucleotide is cytosine and the agent capable of cleaving a base portion of the canonical nucleotide comprises cytosine deaminase in conjunction with UNG.

[0018] In some embodiments, the polynucleotide comprising an abasic site is synthesized from a polynucleotide template comprising DNA or RNA. In various embodiments, the polynucleotide template is selected from the group consisting of RNA, mRNA, cDNA, and genomic DNA. In one embodiment, the polynucleotide comprising an abasic site is single stranded. In one embodiment, the polynucleotide comprising an abasic site is double stranded.

[0019] In some embodiments of the methods of the invention, the polynucleotide to be fragmented (*i.e.*, the polynucleotide comprising an abasic site) is synthesized using a method comprising: (a) extending a composite primer in a complex comprising: (i) a polynucleotide template; and (ii) the composite primer, said composite primer comprising an RNA portion and a 3' DNA portion, wherein the polynucleotide template is hybridized to the composite primer; and (b) cleaving RNA of the annealed composite primer with an enzyme that cleaves RNA from an RNA/DNA hybrid such that another composite primer hybridizes to the template and repeats primer extension and strand displacement, whereby multiple copies of the complementary sequence of the polynucleotide template are produced. In one embodiment, the complex of part (a) comprises: (i) a complex of first and second primer extension products, wherein the first primer extension product is produced by extension of a first primer hybridized to a target RNA with at least one enzyme

IDT Exhibit 1020.2786

comprising RNA-dependent DNA polymerase activity, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; wherein RNA in the complex of first and second primer extension products is cleaved with at least one enzyme that cleaves RNA from an RNA/DNA hybrid such that a composite primer hybridizes to the second primer extension product; and (ii) the composite primer.

[0020] In some embodiments of the methods of the invention, the polynucleotide to be fragmented (*i.e.*, the polynucleotide comprising an abasic site) is synthesized by an amplification method selected from the group consisting of polymerase chain reaction (PCR), strand displacement amplification (SDA), multiple displacement amplification (MDA), rolling circle amplification (RCA), single primer isothermal amplification (SPIA), and Ribo-SPIA. In some embodiments, the polynucleotide is synthesized by a method selected from the group consisting of reverse transcription, primer extension, limited primer extension, replication, and nick translation. In one embodiment, the polynucleotide is synthesized using a labeled primer.

[0021] In one embodiment, the invention provides a method for fragmenting and labeling a polynucleotide, said method comprising: (a) synthesizing a polynucleotide from a polynucleotide template in the presence of a non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated; (b) cleaving a base portion of the non-canonical nucleotide from the synthesized polynucleotide with an enzyme capable of cleaving (*i.e.*, which cleaves) the base portion of the non-canonical nucleotide, whereby an abasic site is generated; (c) cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site at or near the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; (d) contacting the polynucleotide fragment with an enzyme capable of unblocking (*i.e.*, which unblocks) the blocked 3' end of said fragment, whereby a polynucleotide fragment comprising a 3' hydroxyl group is generated; and (e) contacting the polynucleotide fragment comprising a 3' hydroxyl group with an enzyme capable of extending (*i.e.*, which extends) the polynucleotide fragment from the 3' end and a labeled nucleotide, whereby a labeled polynucleotide fragment is generated. In some embodiments, steps (b), (c), and (d) are performed

IDT Exhibit 1020.2787

simultaneously. In some embodiments, steps (b), (c), and (d) are performed simultaneously in the same reaction mixture. In some embodiments, steps (b) and (c) are performed simultaneously. In some embodiments, steps (b) and (c) are performed simultaneously in the same reaction mixture. In some embodiments, steps (c) and (d) are performed simultaneously. In some embodiments, steps (c) and (d) are performed simultaneously in the same reaction mixture. In some embodiments, the method involves synthesizing the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range. In some embodiments, labeled fragments within a desired size range are generated.

[0022] In one embodiment, the invention provides a method for fragmenting a polynucleotide, said method comprising: (a) incubating a reaction mixture, said reaction mixture comprising: (i) a polynucleotide template; and (ii) a non-canonical nucleotide; wherein the incubation is under conditions that permit synthesis of a polynucleotide comprising the non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated; (b) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide comprising the non-canonical nucleotide; and (ii) an enzyme capable of cleaving (*i.e.*, which cleaves) a base portion of the non-canonical nucleotide, wherein the incubation is under conditions that permit cleavage of the base portion of the non-canonical nucleotide, whereby a polynucleotide comprising an abasic site is generated; (c) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide comprising the abasic site; and (ii) an agent capable of chemically cleaving (*i.e.*, which cleaves) the phosphodiester backbone of the polynucleotide comprising the abasic site at or near the abasic site, wherein the incubation is under conditions that permit cleavage of the phosphodiester backbone of the polynucleotide at or near the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; and (d) incubating a reaction mixture, said reaction mixture comprising: (i) the fragment of the polynucleotide comprising a blocked 3' end; and (ii) an enzyme capable of unblocking (*i.e.*, which unblocks) the blocked 3' end, whereby a polynucleotide

IDT Exhibit 1020.2788

fragment comprising a 3' hydroxyl group is generated. In some embodiments, the method further comprises: (e) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide fragment comprising a 3' hydroxyl group; and (ii) an agent capable of extending (*i.e.*, which extends) the fragment from the 3' hydroxyl group; and (iii) a labeled nucleotide, wherein the incubation is under conditions that permit extension of the polynucleotide fragment from the 3' hydroxyl group, whereby a labeled polynucleotide fragment is generated. In one embodiment, the agent capable of extending the polynucleotide fragment from the 3' hydroxyl group is TdT, wherein the polynucleotide fragment is labeled at the 3' hydroxyl group with a labeled nucleotide. In some embodiments, steps (b), (c), and (d) are performed simultaneously in the same reaction mixture. In some embodiments, the incubation is under conditions that permit synthesis of the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range. In some embodiments, fragments within a desired size range are generated.

[0023] In one embodiment, the invention provides a method for generating a polynucleotide fragment with a 3' end hydroxyl group, comprising contacting a polynucleotide fragment with a blocked 3' end with an enzyme capable of unblocking (*i.e.*, which unblocks) the blocked 3' end, wherein the polynucleotide fragment with a blocked 3' end is generated by cleaving a polynucleotide fragment comprising an abasic site at or near the abasic site.

[0024] In another aspect, the invention provides, a method of characterizing a polynucleotide template of interest, comprising analyzing a labeled polynucleotide fragment produced by a method as described herein. In one embodiment, the method comprises (a) generating a labeled polynucleotide fragment by a method as described herein; and (b) analyzing the labeled polynucleotide fragment. In one embodiment, analyzing the labeled polynucleotide fragment comprises determining amount of said products, whereby the amount of the polynucleotide template present in a sample is quantified. In one embodiment, analyzing the labeled polynucleotide fragment comprises contacting the labeled polynucleotide fragment with at least one probe. In

IDT Exhibit 1020.2789

one embodiment, the at least one probe is provided as a microarray. In one embodiment, the microarray is a high density polynucleotide microarray. In one embodiment, the microarray is a high density oligonucleotide microarray. In some embodiments, the microarray comprises at least one probe immobilized on a substrate fabricated from a material selected from the group consisting of paper, glass, ceramic, plastic, polypropylene, polystyrene, nylon, polyacrylamide, nitrocellulose, silicon, and optical fiber. In one embodiment, the at least one probe is immobilized on the substrate in a two-dimensional configuration or a three-dimensional configuration comprising pins, rods, fibers, tapes, threads, beads, particles, microtiter wells, capillaries, and cylinders.

[0025] In another aspect, the invention provides a method of determining gene expression profile in a sample, said method comprising determining the amount of labeled polynucleotide fragment from at least one polynucleotide fragment produced by a method as described herein, wherein the amount is indicative of amount of a polynucleotide template from which the polynucleotide fragment was generated in a sample, whereby a gene expression profile is determined. In one embodiment, the method comprises (a) generating a labeled polynucleotide fragment from at least one polynucleotide template in the sample using a method as described herein; and (b) determining amount of labeled polynucleotide fragment from a polynucleotide template, wherein said amount is indicative of amount of the polynucleotide template in the sample, whereby the a gene expression profile in the sample is determined. In one embodiment, the polynucleotide template is RNA or mRNA. In one embodiment, the amounts of a plurality of polynucleotide fragments derived from a plurality of polynucleotide templates in a sample is determined.

[0026] In another aspect, the invention provides a method of generating hybridization probes, comprising generating a labeled polynucleotide fragment using a method as described herein.

[0027] In another aspect, the invention provides a method for nucleic acid hybridization, comprising hybridizing a labeled polynucleotide fragment with at least one probe, wherein the labeled polynucleotide fragment is generated using a method as described herein. In one embodiment, the method comprises (a) generating a

IDT Exhibit 1020.2790

labeled polynucleotide fragment using a method as described herein; and (b) hybridizing the labeled polynucleotide fragment with at least one probe.

[0028] In another aspect, the invention provides a method for comparative hybridization, comprising comparing hybridization of a first population of labeled polynucleotide fragments prepared using a method as described herein to at least one probe with hybridization of a second population of labeled polynucleotide to the at least one probe. In one embodiment, the method comprises (a) preparing a first population of labeled polynucleotides fragments from a first template polynucleotide sample using a method as described herein; and (b) comparing hybridization of the first population to at least one probe with hybridization of a second population of labeled polynucleotide. In one embodiment, the first population and second population comprise detectably different labels. In one embodiment, the second population of labeled polynucleotides are prepared from a second polynucleotide sample using a method as described herein. In one embodiment, comparing comprises determining amount of said products, whereby the amount of the first and second polynucleotide templates is quantified. In one embodiment, the first and/or second template polynucleotides from which the first and/or second populations of labeled polynucleotides are prepared comprise genomic DNA.

[0029] In another aspect, the invention provides a method for detecting presence or absence of a mutation in a template, comprising analyzing a labeled polynucleotide fragment prepared by a method as described herein, whereby presence of absence of a mutation is detected. In one embodiment, the method comprises (a) generating a labeled polynucleotide fragment by a method as described herein; and (b) analyzing the labeled polynucleotide fragment, whereby presence or absence of a mutation is detected. In one embodiment, the labeled polynucleotide fragment is compared to a reference template. In various embodiments, the mutation is selected from the group consisting of a base substitution, a base insertion, a base deletion, and a single nucleotide polymorphism.

[0030] In another aspect, the invention provides a composition comprising: (a) an agent capable of cleaving (*i.e.*, which cleaves) a base portion of a nucleotide to generate an abasic site in a polynucleotide; (b) an agent capable of cleaving (*i.e.*,

IDT Exhibit 1020.2791

which cleaves) a phosphodiester backbone at or near an abasic site to produce a polynucleotide fragment with a blocked 3' end; and (c) an enzyme capable of unblocking (*i.e.*, which unblocks) a blocked 3' end to generate a polynucleotide comprising a 3' hydroxyl group. In one embodiment, (a) is an N-glycosylase, (b) is a polyamine, and (c) is an enzyme comprising a 3' exonuclease activity. In one embodiment, (a) is UNG, (b) is DMED, and (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1.

[0031] In another aspect, the invention provides a kit comprising: (a) an agent capable of cleaving (*i.e.*, which cleaves) a base portion of a nucleotide to generate an abasic site in a polynucleotide; (b) an agent capable of cleaving (*i.e.*, which cleaves) a phosphodiester backbone at or near an abasic site to produce a polynucleotide fragment with a blocked 3' end; and (c) an enzyme capable of unblocking (*i.e.*, which unblocks) a blocked 3' end to generate a polynucleotide comprising a 3' hydroxyl group. In one embodiment, the kit further comprises: (d) an agent capable of labeling (*i.e.*, which labels) a 3' hydroxyl group of a polynucleotide. In one embodiment, (a) is an N-glycosylase, (b) is a polyamine, (c) is an enzyme comprising a 3' exonuclease activity; and (d) is a template independent polymerase. In one embodiment, (a) is UNG, (b) is DMED, (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1; and (d) is TdT. In some embodiments, the kit further comprises a non-canonical nucleotide. In some embodiments, the kit further comprises a non-canonical nucleotide and an enzyme capable of synthesizing a polynucleotide comprising the non-canonical nucleotide. In one embodiment, the non-canonical nucleotide is dUTP and the agent capable of cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide is UNG.

[0032] In some embodiments, the kit further comprises: (e) a labeled nucleotide. In one embodiment, (a) is an N-glycosylase, (b) is a polyamine, (c) is an enzyme comprising a 3' exonuclease activity; (d) is a template independent polymerase; and (e) is a biotinylated nucleotide. In some embodiments, (e) is selected from the group consisting of a biotinylated nucleotide triphosphate (NTP), a biotinylated deoxynucleotide triphosphate (dNTP), and a biotinylated dideoxynucleotide triphosphate (ddNTP). In one embodiment, (a) is UNG, (b) is

IDT Exhibit 1020.2792

DMED, (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1; (d) is TdT, and (e) is selected from the group consisting of biotin 2',3'-dideoxy-UTP and biotin 2',3'-dideoxy-CTP.

[0033] In some embodiments, a kit of the invention comprises, in addition to the components described above, a template dependent DNA polymerase; a composite primer, wherein the composite primer comprises a 5' RNA portion and a 3' DNA portion; and an agent capable of cleaving RNA from an RNA-DNA hybrid. In some embodiments, the RNA portion of the composite primer is 5' with respect to the 3' DNA portion, the 5' RNA portion is adjacent to the 3' DNA portion, the RNA portion of the composite primer consists of about 5 to about 50 nucleotides and the DNA portion of the composite primer consists of 1 to about 20 nucleotides. In one embodiment, the agent that cleaves RNA from an RNA-DNA hybrid is RNAse H.

[0034] Kits of the invention generally comprise packaging, and may comprise instructions for use in a method for polynucleotide fragmentation, or polynucleotide fragmentation and labeling, as described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] **Figure 1** schematically depicts an embodiment of a nucleic acid fragmentation and labeling procedure as described herein. A single stranded nucleic acid comprising non-canonical nucleotide dU (**A**) is used to generate a nucleic acid with abasic sites (**B**). Non-canonical nucleotides are denoted as "U" in nucleic acid A and canonical nucleotides are denoted as "N." Abasic sites are denoted as "X" in nucleic acid B. The nucleic acid is fragmented at the abasic sites to generate fragments with blocked 3' termini (**C**). A 3'-5' exonuclease is used to unblock the 3' termini of fragments C, thereby generating nucleic acid fragments with 3' hydroxyl groups (**D**). The 3' hydroxyl groups serve as substrates for end labeling using terminal transferase and labeled nucleotide triphosphates, thereby producing labeled nucleic acid fragments (**E**). The end labels are denoted as "L" in fragments E.

[0036] **Figure 2** shows data from an experiment comparing the ability of various 3' exonucleases to unblock blocked 3' ends of polynucleotide fragments, as described in Example 1.

IDT Exhibit 1020.2793

[0037] **Figure 3** shows data from an experiment demonstrating the ability of terminal deoxynucleotidyl transferase (TdT) to label 3' hydroxyl groups of polynucleotide fragments which after blocked 3' ends were unblocked by treatment with apurinic/apyrimidinic exonuclease 1 (APE 1), as described in Example 2.

DETAILED DESCRIPTION

Methods for generating fragmented polynucleotides with 3' end hydroxyl groups

[0038] The invention provides novel methods, compositions, and kits for fragmenting polynucleotides to generated polynucleotide fragments with hydroxyl groups at their 3' ends. The methods, compositions, and kits of the invention are useful for fragmenting and labeling polynucleotides. These methods are suitable for, for example, generation of labeled polynucleotide fragments, for use as hybridization probes, or generation of polynucleotide fragments which may be hybridized to a polynucleotide template and extended with a template dependent polymerase.

[0039] In methods of the invention, a polynucleotide is cleaved at or near an abasic site present in the polynucleotide. The abasic site may be prepared by cleavage of a base portion of a non-canonical nucleotide present in the polynucleotide, cleavage of a base portion of a canonical nucleotide present in the polynucleotide, or cleavage of a base portion of a methylated nucleotide present in the polynucleotide. For methods in which an abasic site is generated by cleavage of a base portion of a non-canonical nucleotide, the spacing of non-canonical nucleotides in the polynucleotide to be fragmented and labeled, relates to and determines the size of fragments and intensity of labeling. This feature permits control of fragment size and/or site of labeling by use of conditions permitting controlled incorporation of non-canonical nucleotide, for example, during synthesis of the polynucleotide comprising the non-canonical nucleotide from a polynucleotide template.

[0040] Cleavage at or near an abasic site is generally effected with an agent or under conditions which do not generate a 3' hydroxyl group, *i.e.*, conditions that generate a "blocked" 3' end that contains a non-nucleotide moiety, such as an

IDT Exhibit 1020.2794

aldehyde group (*e.g.*, the sugar residue remaining behind after generation of the abasic site), at the 3' end.

[0041] Generally, a chemical fragmentation method is used, wherein a polynucleotide fragment comprising a blocked 3' end is generated. In some embodiments, chemical fragmentation at or near an abasic site is effected with a polyamine, such as, for example, N,N'-dimethylethylenediamine (DMED), which produces polynucleotide fragments with blocked 3' ends. Cleavage with a polyamine most commonly employs a β- or β-γ elimination mechanism. β-elimination results in cleavage of the 3'-phosphodiester bond and a remnant of the sugar moiety (an aldehyde) is attached to the 3' end. Depending on the cleavage agent used, other modifications of the 3' end are also possible, such as a 3'-phosphoglycolate group. Blocked 3' ends may be unblocked by digestion with an enzyme capable of removing the attached blocking moiety, such as a 3' to 5' exonuclease, preferably a non-processive exonuclease, thereby producing polynucleotide fragments with a 3' end hydroxyl group. In some embodiments, an enzyme that comprises an exonuclease activity and does not comprise endonuclease activity is used. In some embodiments, an enzyme that comprises both exonuclease and endonuclease activities is used. In one embodiment, an enzyme that comprises both exonuclease and endonuclease activities is used under conditions in which the endonuclease activity is substantially minimized or absent. In some embodiments, endonuclease 4, exonuclease T, or the 3' to 5' exonuclease activity of the apurinic/apyrimidinic endonuclease (APE 1) is used.

[0042] In some embodiments, polynucleotide fragments with unblocked 3' ends are labeled with an agent capable of labeling at 3' hydroxyl groups of polynucleotides. In some embodiments, a template independent polymerase is used for labeling. In one embodiment, the template independent polymerase is terminal deoxynucleotidyl transferase (TdT), an enzyme which is capable of attaching one or more nucleotides (*i.e.*, labeled nucleotides, unlabeled nucleotides, or a mixture of labeled and unlabeled nucleotides) at a polynucleotide 3' end hydroxyl group by extension of the polynucleotide from the 3' end. "Labeled" or "detectable" nucleotide or polynucleotide, as used herein, refers to a nucleotide (or nucleotide analog thereof) or polynucleotide that is directly or indirectly detectable. A

IDT Exhibit 1020.2795

nucleotide may comprise a directly-detectable label such as, for example, a fluorophore (*e.g.*, cy dyes, alexa dyes, fluorescein, etc.), an enzyme, a chromophore , or a radiolabel, or the nucleotide may comprise an indirectly-detectable label such as a hapten which is detectable by binding of a labeled second member of a specific binding pair, such as, for example, biotin/avidin or streptavidin, antigen/antibody, etc., and the label attached to the second member of the binding pair may be, for example, a fluorophore, an enzyme, a chromophore, or a radiolabel, or the second member of the binding pair may be attached to a detectable particle. In one embodiment, the nucleotide is a biotinylated nucleotide, such as, for example, biotin 2',3'-dideoxy-UTP, and is detectable by binding of labeled avidin or streptavidin.

[0043] In some embodiments, a polynucleotide fragment with an unblocked 3' end, produced as described herein, is hybridized to a polynucleotide template and extended from the 3' hydroxyl group with a template-dependent polymerase, using labeled or unlabeled nucleotides, or a mixture of labeled or unlabeled nucleotides. When labeled polynucleotides are incorporated, a labeled polynucleotide extension product is produced.

[0044] In some embodiments, a polynucleotide fragment with an unblocked 3' end, produced as described herein, is ligated to another polynucleotide with a ligase enzyme. If the polynucleotide to which the polynucleotide fragment is ligated is labeled, a labeled ligation product is produced.

[0045] In some embodiments, a polynucleotide fragment with an unblocked 3' end, produced as described herein, is "tailed" using a template-independent polymerase, wherein a "tail" of nucleotides, *i.e.*, labeled or unlabeled nucleotides or a mixture thereof, is added at the 3' end of the fragment.

Generation of abasic sites in polynucleotides comprising non-canonical nucleotides

[0046] In one aspect, the invention provides methods for fragmenting and labeling a polynucleotide comprising an abasic site produced by incorporation of a non-canonical nucleotide. The methods generally comprise generation of a polynucleotide comprising a non-canonical nucleotide, cleavage of a base portion of

IDT Exhibit 1020.2796

the non-canonical nucleotide present in the polynucleotide with an agent (such as an enzyme) capable of cleaving a base portion of the non-canonical nucleotide (whereby an abasic site is generated); chemical cleavage of the phosphodiester backbone at or near the abasic site with an agent or under conditions that do not generate a 3' hydroxyl group, *i.e.*, conditions that generate a polynucleotide fragment having a blocked 3' end; digestion of fragments with an enzyme capable of generating a 3' end with a 3' hydroxyl group from a blocked 3' end; and optionally labeling at the 3' hydroxyl group with an enzyme or agent capable of attaching a label to the 3' hydroxyl group, whereby labeled polynucleotide fragments are generated.

[0047] The methods of fragmenting and labeling a polynucleotide generally comprise synthesis of a polynucleotide comprising a non-canonical nucleotide from a polynucleotide template in the presence of a non-canonical nucleotide, whereby a polynucleotide comprising a non-canonical nucleotide(s) is generated.

[0048] Non-canonical nucleotides are known in the art and any suitable non-canonical nucleotide can be used. In some embodiments, two or more different non-canonical nucleotides are used, such that a polynucleotide comprising two or more non-canonical nucleotides is generated. Methods for synthesizing polynucleotides from a polynucleotide template are known in the art and described herein, and any suitable method can be used in the methods of the invention. In some embodiments, synthesis of the polynucleotide comprising the non-canonical nucleotides comprises SPIA™ (single primer isothermal amplification; *see* Kurn, U.S. Patent No. 6,251,639 and 6,692,918), Ribo-SPIA™ (*see* Kurn, U.S. Patent No. 6,946,251), PCR, primer extension, reverse transcription, strand displacement amplification (SDA), multiple displacement amplification (MDA), rolling circle amplification (RCA), nick translation based DNA synthesis, DNA replication, and the like. The polynucleotide that is synthesized can be single stranded, double-stranded or partially double stranded, and either or both strands can comprise a non-canonical nucleotide. In some embodiments, the polynucleotide that is synthesized comprises a cDNA. The polynucleotide template (from which the polynucleotide comprising a non-canonical nucleotide is synthesized) is any template from which one desires to produce polynucleotide fragments or labeled fragments thereof. In some embodiments, the

IDT Exhibit 1020.2797

template comprises RNA, mRNA, genomic DNA, cDNA, or synthetic DNA. In other embodiments, the template comprises a cDNA library, a subtractive hybridization library, or a genomic library. In one embodiment, the polynucleotide comprising the non-canonical nucleotide is synthesized using limited and/or controlled incorporation of the non-canonical nucleotide, which results in generation of a polynucleotide with a frequency or proportion of non-canonical nucleotides such that, labeled fragments of a desired size (or size range) are generated (following production of an abasic site, cleavage of the phosphodiester backbone at or near an abasic site with an agent or under conditions in which a 3' end comprising a hydroxyl group is not produced (*i.e.*, production of a fragments with blocked hydroxyl groups), generation of hydroxyl groups at the 3' ends on the polynucleotide fragments (*i.e.*, by unblocking the blocked 3' ends, for example, with an enzyme comprising a 3' to 5' exonuclease activity), and labeling of the polynucleotide fragments using polymerase extension from the 3' hydroxyl groups in the presence of a detectable nucleotide or ligation to a detectable polynucleotide (*e.g.*, extension at the 3' hydroxyl groups with a template independent polymerase such as terminal transferase to incorporate one or more detectable nucleotides at the 3' ends).

[0049] In some embodiments, a labeled primer is used during synthesis of the polynucleotide comprising a non-canonical nucleotide. In other embodiments, a primer comprising a non-canonical nucleotide (such as dUTP) is used during synthesis of the polynucleotide comprising a non-canonical nucleotide. In other embodiments, the primer is a composite primer, said composite primer comprising a RNA portion and a 3' DNA portion.

[0050] It is understood that a polynucleotide comprising a non-canonical nucleotide can be a multiplicity (from small to very large) of different polynucleotide molecules. Such populations can be related in sequence *(e.g.*, members of a gene family or superfamily) or extremely diverse in sequence (*e.g.*, generated from all mRNA, generated from all genomic DNA, etc.). Polynucleotides can also correspond to single sequences (which can be part or all of a known gene, for example, a coding region, genomic portion, etc.).

IDT Exhibit 1020.2798

[0051] A base portion of the non-canonical nucleotide is cleaved by an agent (such as an enzyme) capable of cleaving a base portion of a non-canonical nucleotide. Such agents are known in the art and described herein. In one embodiment, the agent capable of specifically cleaving a base portion of a non-canonical nucleotide is an N-glycosylase. In another embodiment, the agent is Uracil N-Glycosylase (interchangeably termed "UNG" or "uracil DNA glyosylase").

Generation of abasic sites in polynucleotides comprising methylated nucleotides

[0052] In one aspect, the invention provides methods for fragmenting and labeling a polynucleotide comprising an abasic site produced by cleaving a base portion of a methylated nucleotide with an agent capable of cleaving a base portion of the methylated nucleotide to create an abasic site, whereby an abasic site is generate. The methods generally comprise cleavage of a base portion of a methylated nucleotide with an agent (such as an enzyme) capable of cleaving a base portion of a methylated nucleotide (whereby an abasic site is generated); chemical cleavage of the phosphodiester backbone at or near the abasic site with an agent or under conditions that do not generate a 3' hydroxyl group, *i.e.*, conditions that generate a polynucleotide fragment with a blocked 3' end; digestion of fragments with an enzyme capable of generating a polynucleotide containing a hydroxyl group at the 3' end; and optionally labeling at the 3' end with an enzyme or agent capable of attaching a label to the 3' hydroxyl group, whereby labeled polynucleotide fragments are generated (3' end labeling).

Generation of abasic sites in polynucleotides by cleaving base portions of canonical nucleotides

[0053] In one aspect, the invention provides methods for fragmenting and labeling a polynucleotide comprising an abasic site produced by cleaving a base portion of a canonical nucleotide with an agent capable of cleaving a base portion of the canonical nucleotide to create an abasic site, whereby an abasic site is generated. The methods generally comprise cleavage of a base portion of a canonical nucleotide

IDT Exhibit 1020.2799

with an agent (such as an enzyme) capable of cleaving a base portion of a methylated nucleotide (whereby an abasic site is generated); chemical cleavage of the phosphodiester backbone at or near the abasic site with an agent or under conditions that do not generate a 3' hydroxyl group, *i.e.*, conditions that generate a polynucleotide fragment with a blocked 3' end; digestion of fragments with an enzyme capable of generating a polynucleotide containing a hydroxyl group at the 3' end; and optionally labeling at the 3' hydroxyl group with an enzyme or agent capable of attaching a label to the 3' hydroxyl group, whereby labeled polynucleotide fragments are generated.

Fragmentation at abasic sites to produce polynucleotide fragments with blocked 3' ends

[0054] In methods of the invention, the phosphodiester backbone of a polynucleotide comprising an abasic site is cleaved at or near the abasic site by an agent capable of cleaving the phosphodiester backbone at or near an abasic site, such that two or more fragments are produced. As used herein, "cleaving the backbone or phosphodiester backbone" is also termed "fragmentation" or "fragmenting." Fragmentation of the polynucleotide comprising an abasic site is conducted with an agent or under conditions in which polynucleotides comprising hydroxyl groups at their 3' ends are substantially not produced. Generally, a chemical fragmentation agent is used, producing polynucleotide fragments with blocked 3' ends. In some embodiments, a polyamine, such as DMED, is used for fragmentation.

[0055] Generally, cleavage occurs 3' to the abasic site (*e.g.*, cleavage between the deoxyribose ring and 3'-phosphate group of the abasic residue and the deoxyribose ring of the adjacent nucleotide, generating a free 5' phosphate group on the deoxyribose ring of the adjacent nucleotide), such that an abasic site is located at the 3' end of the resulting fragment. In still other embodiments, more complex forms of cleavage are possible, for example, cleavage such that cleavage of the phosphodiester backbone and cleavage of a portion of the abasic nucleotide results. Selection of reaction conditions also permits control of the degree, level or completeness of the fragmentation reactions. In some embodiments, reaction conditions can be selected

IDT Exhibit 1020.2800

such that the cleavage reaction is performed in the presence of a large excess of reagents and allowed to run to completion with minimal concern about excessive cleavage of the polynucleotide (i.e., while retaining a desired fragment size, which may be determined by spacing of the incorporated non-canonical nucleotide, during the synthesis step, above). In other embodiments, reaction conditions are selected such that fragmentation is not complete (in the sense that the backbone at some abasic sites remains uncleaved (unfragmented), such that polynucleotide fragments comprising more than one abasic site are generated. Such fragments comprise internal (unfragmented) abasic sites.

Unblocking of blocked 3' ends

[0056] The polynucleotide fragments generated by cleavage of the phosphodiester backbone are contacted with an agent, such as an enzyme comprising an exonuclease activity, that is capable of generating a polynucleotide comprising a hydroxyl group at the 3' end from a polynucleotide comprising a blocked 3' end. Preferably, a non-processive exonuclease is used, such as, for example, APE 1. The resulting polynucleotide fragments have unblocked 3' ends, *i.e.*, comprising a hydroxyl group at the 3' end.

[0057] In some embodiments, a polynucleotide fragment comprising a 3' end hydroxyl group, produced as described herein, is extended with a template independent or template dependent polymerase. In one embodiment, the polynucleotide fragment may be extended with a template independent polymerase, such as a terminal transferase, to incorporate one or more labeled nucleotide residue (or nucleotide analogs thereof), one or more unlabeled nucleotide residue (or nucleotide analogs thereof) or a mixture of labeled and unlabeled nucleotide residues (or nucleotide analogs thereof), at the 3' end. In another embodiment, the polynucleotide fragment is hybridized to a polynucleotide template and extended by a template dependent polymerase. In another embodiment, the polynucleotide fragment is ligated to another polynucleotide with an enzyme comprising a ligase enzyme.

IDT Exhibit 1020.2801

Labeling of polynucleotide fragments at 3' hydroxyl groups

[0058] Agents capable of labeling at a 3' hydroxyl group of a polynucleotide are known in the art. For example, a template independent polymerase, such as TdT, can be used to attach one or more labeled nucleotides at the 3' hydroxyl group (*i.e.*, extend from the 3' end). In some embodiments, the detectable moiety (label) is directly or indirectly detectable. In some embodiments, the detectable signal is amplified. In some embodiments, the detectable moiety comprises an organic molecule. In other embodiments, the detectable moiety comprises an antibody. In other embodiments, the detectable signal is fluorescent. In other embodiments, the detectable signal is enzymatically generated. In one embodiment, the fragments are labeled by template independent extension with TdT, using a labeled nucleotide triphosphate (or labeled nucleotide analog thereof). In one embodiment, the labeled nucleotide is a biotinylated nucleotide, such as, for example, biotin 2',3'-dideoxy-UTP, biotin-dUTP, or biotin-UTP. Other labeled nucleotides, *e.g.*, dNTPs or NTPs, or terminator nucleotides such as 2',3'-dideoxy-NTPs), or combinations thereof, as well as combinations of labeled and unlabeled nucleotides or dideoxy-nucleotides, may also be used.

[0059] The methods of the invention include methods of using polynucleotide fragments and labeled polynucleotides produced by the methods of the invention (so-called "applications"). The invention provides methods to characterize (for example, detect presence or absence of and/or quantify) a sequence of interest by analyzing fragmented and labeled products by detection/quantification methods such as those based on array technologies or solution phase technologies. In some embodiments, the invention provides methods of detecting the presence or absence of mutations.

[0060] In other embodiments, the invention provides methods of producing a hybridization probe; hybridization using the hybridization probes; detection using the hybridization probes; characterizing and/or quantitating nucleic acid; preparing a subtractive hybridization probe; comparative genomic hybridization; and determining a gene expression profile, using the fragmented nucleic acids generated by the methods of the invention.

IDT Exhibit 1020.2802

General Techniques

[0061] 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).

[0062] Primers, oligonucleotides and polynucleotides employed in the invention can be generated using standard techniques known in the art.

Definitions

[0063] A "template sequence," or "template nucleic acid" or "template" as used herein, is a polynucleotide comprising a sequence of interest, for which synthesis of a complement comprising a non-canonical nucleotide is desired. The template sequence may be known or not known, in terms of its actual sequence. In some instances, the terms "target," "template," and variations thereof, are used interchangeably.

[0064] "Polynucleotide," or "nucleic acid," as used interchangeably herein, refer to polymers of nucleotides of any length, and include DNA. The nucleotides can be deoxyribonucleotides, modified nucleotides or bases, and/or their analogs, or any substrate that can be incorporated into a polymer by DNA polymerase. Nucleotides include canonical and non-canonical nucleotides and a polynucleotide can comprise canonical and non-canonical nucleotides. A polynucleotide may comprise modified (altered) nucleotides, such as, for example, modification to the nucleotide structure and or modification to the phosphodiester backbone. As discussed herein modified nucleotide can be canonical nucleotide or non-canonical (cleavable) nucleotides. It is understood, however, that modified nucleotides that are not non-canonical (cleavable) nucleotide under the reaction conditions used in the methods of the invention, if

IDT Exhibit 1020.2803

present, generally should not affect the ability of the polynucleotide to undergo cleavage of a base portion of non-canonical nucleotide, such that an abasic site is generated, and/or cleavage of a phosphodiester backbone at an abasic site, such that fragments are generated, and/or immobilization of a polynucleotide (or fragment thereof) to a substrate, as described herein. If present, modification to the nucleotide structure, such as methylated nucleotides 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, carbamates, 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). It is understood that internucleotide modifications may, e.g., alter the efficiency and/or kinetics of cleavage of the phosphodiester backbone (as when, for example a phosphodiester backbone is cleaved at an abasic site, as described herein). 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. 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, α-anomeric sugars, epimeric sugars such as arabinose, xyloses or lyxoses, pyranose sugars, furanose sugars,

IDT Exhibit 1020.2804

sedoheptuloses, acyclic analogs and abasic nucleoside analogs. 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_2 ("amidate"), P(O)R, P(O)OR', CO or CH_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. Not all linkages in a polynucleotide need be identical. The preceding description applies to all polynucleotides referred to herein, including DNA. It is understood, however, that modified nucleotides and/or internucleotide linkages and/or, if present, generally should not affect the ability of the polynucleotide to undergo cleavage of a base portion of a non-canonical nucleotide, such that an abasic site is generated, and/or the ability of a polynucleotide to undergo cleavage of a phosphodiester backbone at an abasic site, such that fragments are generated, and/or the ability of a polynucleotide to be immobilized at an abasic site (such as an abasic site at an end of a polynucleotide and/or an abasic site that is not at an end of a polynucleotide) to a surface, as described herein.

[0065] "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 terms "oligonucleotide" and "polynucleotide" are not mutually exclusive. The description above for polynucleotides is equally and fully applicable to oligonucleotides.

[0066] A "primer," as used herein, refers to a nucleotide sequence (a polynucleotide), generally with a free 3'-OH group, that hybridizes with a template sequence (such as a template 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 an RNA template created following RNase cleavage of a template RNA-DNA complex) that is hybridized to a sequence in the template itself (for example, as a hairpin loop),

IDT Exhibit 1020.2805

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.

[0067] A "complex" is an assembly of components. A complex may or may not be stable and may be directly or indirectly detected. For example, as is described herein, given certain components of a reaction, and the type of product(s) of the reaction, existence of a complex can be inferred. For purposes of this invention, a complex is generally an intermediate with respect to the final polynucleotide fragments, labeled polynucleotide, labeled polynucleotide fragments, and/or immobilized polynucleotide or fragment thereof.

[0068] A "fragment" of a polynucleotide or oligonucleotide is a contiguous sequence of 2 or more bases. In other embodiments, a fragment (also termed "region" or "portion") is any of about 3, about 5, about 10, about 15, about 20, about 25, about 30 about 35 about 40, about 50, about 65, about 75, about 85, about 100, about 125, about 150, about 175, about 200, about 225, about 250, about 300, about 350, about 400, about 450, about 500, about 550, about 600, about 650 or more nucleotides in length. In some embodiments, the fragments can be at least about 3, about 5, about 10, about 15, about 20, about 25, about 30 about 35 about 40, about 50, about 65, about 75, about 85, about 100, about 125, about 150, about 175, about 200, about 225, about 250, about 300, about 350, about 400, about 450, about 500, about 550, about 600, about 650 or more nucleotides in length. In other embodiments, the fragments can be less than about 3, about 5, about 10, about 15, about 20, about 25, about 30 about 35 about 40, about 50, about 65, about 75, about 85, about 100, about 125, about 150, about 175, about 200, about 225, about 250, about 300, about 350, about 400, about 450, about 500, about 550, about 600, about 650 or more nucleotides in length. In some embodiment, these fragment lengths represent an average size in the population of fragments generated using the methods of the invention.

[0069] A "reaction mixture" is an assemblage of components, which, under suitable conditions, react to form a complex (which may be an intermediate) and/or a product(s).

IDT Exhibit 1020.2806

[0070] "A", "an" and "the", and the like, unless otherwise indicated include plural forms. "A" fragment means one or more fragments. "A" non-canonical nucleotide means one or more non-canonical nucleotides.

[0071] Conditions that "allow" or "permit" an event to occur or conditions that are "suitable" for an event to occur, such as polynucleotide synthesis, cleavage of a base portion of a non-canonical nucleotide, cleavage of a phosphodiester backbone at an abasic site, and the like, or "suitable" conditions are conditions that do not prevent such events from occurring. Thus, these conditions permit, enhance, facilitate, and/or are conducive to the event. Such conditions, known in the art and described herein, depend upon, for example, the nature of the polynucleotide sequence, temperature, and buffer conditions. These conditions also depend on what event is desired, such as polynucleotide synthesis, cleavage of a base portion of a non-canonical nucleotide, cleavage of a phosphodiester backbone at an abasic site, labeling an abasic site, immobilizing a polynucleotide fragment or a polynucleotide, etc.

[0072] "Microarray" and "array," as used interchangeably herein, comprise a surface with an array, preferably ordered array, of putative binding (e.g., by hybridization) sites for a biochemical sample (target) which often has undetermined characteristics. In a preferred embodiment, a microarray refers to an assembly of distinct polynucleotide or oligonucleotide probes immobilized at defined positions on a substrate. Arrays are formed on substrates fabricated with materials such as paper, glass, plastic (e.g., polypropylene, nylon, polystyrene), polyacrylamide, nitrocellulose, silicon and other metals, optical fiber or any other suitable solid or semi-solid support, and configured in a planar (e.g., glass plates, silicon chips) or three-dimensional (e.g., pins, fibers, beads, particles, microtiter wells, capillaries) configuration. Probes forming the arrays may be attached to the substrate by any number of ways including (i) *in situ* synthesis (e.g., high-density oligonucleotide arrays) using photolithographic techniques (see, Fodor et al., *Science* (1991), 251:767-773; Pease et al., *Proc. Natl. Acad. Sci. U.S.A.* (1994), 91:5022-5026; Lockhart et al., *Nature Biotechnology* (1996), 14:1675; U.S. Pat. Nos. 5,578,832; 5,556,752; and 5,510,270); (ii) spotting/printing at medium to low-density (e.g., cDNA probes) on glass, nylon or

IDT Exhibit 1020.2807

nitrocellulose (Schena et al, *Science* (1995), 270:467-470, DeRisi et al, *Nature Genetics* (1996), 14:457-460; Shalon et al., *Genome Res.* (1996), 6:639-645; and Schena et al., *Proc. Natl. Acad. Sci. U.S.A.* (1995), 93:10539-11286); (iii) by masking (Maskos and Southern, *Nuc. Acids. Res.* (1992), 20:1679-1684) and (iv) by dot-blotting on a nylon or nitrocellulose hybridization membrane (see, e.g., Sambrook et al., Eds., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., Vol. 1-3, Cold Spring Harbor Laboratory (Cold Spring Harbor, N.Y.)). Probes may also be noncovalently immobilized on the substrate by hybridization to anchors, by means of magnetic beads, or in a fluid phase such as in microtiter wells or capillaries. The probe molecules are generally nucleic acids such as DNA, RNA, PNA, and cDNA but may also include proteins, polypeptides, oligosaccharides, cells, tissues and any permutations thereof which can specifically bind the target molecules.

[0073] The term "3'" generally refers to a region or position in a polynucleotide or oligonucleotide 3' (downstream) from another region or position in the same polynucleotide or oligonucleotide.

[0074] The term "5'" generally refers to a region or position in a polynucleotide or oligonucleotide 5' (upstream) from another region or position in the same polynucleotide or oligonucleotide.

[0075] The term "3'-DNA portion," "3'-DNA region," "3'-RNA portion," and "3'-RNA region," refer to the portion or region of a polynucleotide or oligonucleotide located towards the 3' end of the polynucleotide or oligonucleotide, and may or may not include the 3' most nucleotide(s) or moieties attached to the 3' most nucleotide of the same polynucleotide or oligonucleotide. The 3' most nucleotide(s) can be preferably from about 1 to about 50, more preferably from about 10 to about 40, even more preferably from about 20 to about 30 nucleotides.

[0076] As used herein, "canonical" nucleotide means a nucleotide comprising one the four common nucleic acid bases adenine, cytosine, guanine and thymine that are commonly found in DNA. The term also encompasses the respective deoxyribonucleosides, deoxyribonucleotides or 2'-deoxyribonucleoside-5'-triphosphates that contain one of the four common nucleic acid bases adenine, cytosine, guanine and thymine (though as explained herein, the base can be a

IDT Exhibit 1020.2808

modified and/or altered base as discussed, for example, in the definition of polynucleotide). As used herein, the base portions of canonical nucleotides are generally not cleavable under the conditions used in the methods of the invention.

[0077] As used herein, "non-canonical nucleotide" (interchangeably called "non-canonical deoxyribonucleoside triphosphate") refers to a nucleotide comprising a base other than the four canonical bases. The term also encompasses the respective deoxyribonucleosides, deoxyribonucleotides or 2'-deoxyribonucleoside-5'-triphosphates that contain a base other than the four canonical bases. In the context of this invention, nucleotides containing uracil (such as dUTP), or the respective deoxyribonucleosides, deoxyribonucleotides or 2'-deoxyribonucleoside-5'-triphosphates, are a non-canonical nucleotides. As used herein, the base portions of non-canonical nucleotides are capable of being, generally, specifically or selectively cleaved (such that a nucleotide comprising an abasic site is created) under the reaction conditions used in the methods of the invention. As described herein, non-canonical nucleotides are generally also capable of being incorporated into a polynucleotide during synthesis of a polynucleotide (during e.g., primer extension and/or replication); capable of being generally, specifically or selectively cleaved by an agent that cleaves a base portion of a nucleotide, such that a polynucleotide comprising an abasic site is generated; comprise a suitable internucleotide connection (when incorporated into a polynucleotide) such that a phosphodiester backbone at an abasic site (i.e., the non-canonical nucleotide following cleavage of a base portion) is capable of being cleaved by an agent capable of such cleavage; capable of being labeled (following generation of an abasic site); and/or capable of immobilization to a surface (following generation of an abasic site), according to the methods described herein. It is understood that the non-canonical nucleotide may, but does not necessarily, require all of the features described above, depending on the particular method of the invention in which the non-canonical nucleotide is to be used. In some embodiments, non-canonical nucleotides are altered and/or modified nucleotides as described herein. Non-canonical nucleotide refers to a nucleotide that is incorporated into a polynucleotide as well as to a single nucleotide.

IDT Exhibit 1020.2809

[0078] The term "analyte" as used herein refers to a substance to be detected or assayed by the method of the present invention, for example, a compound whose properties, location, quantity and/or identity is desired to be characterized. Typical analytes may include, but are not limited to proteins, peptides, nucleic acid segments, cells, microorganisms and fragments and products thereof, organic molecules, inorganic molecules, or any substance for which immobilization sites for binding partner(s) can be developed. As this disclosure clearly conveys, an analyte is a substrate.

[0079] As used herein, an "abasic site" refers to the site of incorporation of the non-canonical nucleotide following treatment with an agent capable of effecting cleavage of a base portion of the non-canonical nucleotide. An abasic site (interchangeably termed "AP site") can comprise a hemiacetal ring, and lacks a base portion of the non-canonical nucleotide. As used herein, "abasic site" encompasses any chemical structure remaining following treatment of a canonical or non-canonical nucleotide (present in a polynucleotide chain) with an agent (e.g., an enzyme, or heat or basic conditions) capable of effecting cleavage of a base portion of a non-canonical nucleotide. Thus, an abasic site as used herein includes a modified sugar moiety attached to the 3' terminus of nicked polynucleotide, as when, for example, endonuclease III or OGG1 protein are used to cleave the base portion of the non-canonical nucleotide. *See, e.g.*, Kow, (2000) *Methods* 22, 164-169 (e.g., Figure 4).

[0080] As used herein, cleavage of a backbone (*e.g.*, phosphodiester backbone) "at" an abasic site means cleavage of the phosphodiester linkage 3' to the abasic site or 5' to the abasic site, or both. As the disclosure herein conveys, "at" an abasic site refers to proximate or near location (such as immediately 3', immediately 5'). In still other embodiments, more complex forms of cleavage are possible, for example, cleavage such that cleavage of the phosphodiester backbone and cleavage of (a portion of) the abasic nucleotide results.

[0081] As used herein, a "label" (interchangeably called a "detectable moiety") refers to a moiety that is associated or linked with a polynucleotide (interchangeably called "labeling"). The labeled polynucleotide may be directly or indirectly detected, generally through a detectable signal. The detectable moiety

IDT Exhibit 1020.2810

(label) can be attached (or associated) either directly or through a non-interfering linkage group with other moieties capable of specifically associating with one or more sites to be labeled. The detectable moiety (label) may be covalently or non-covalently associated as well as directly or indirectly associated.

[0082] As used herein, a "blocked 3' end" refers to a 3' end of a polynucleotide fragment that contains a moiety, such as an aldehyde group, *e.g.*, the sugar residue attached to the nucleotide base that was removed to create an abasic site, which is left behind at the 3' end of the polynucleotide as a result of cleaving a polynucleotide at the abasic site as described herein, rather than a hydroxyl group at the 3' end. "Unblocking" of the 3' end refers to removal of the non-nucleotide moiety, *e.g.*, an aldehyde group, resulting in a polynucleotide having a hydroxyl group at the 3' end.

[0083] The following are examples of the methods of the invention. It is understood that various other embodiments may be practiced, given the general description provided herein. For example, reference to using an agent capable of cleaving a base portion of a non-canonical nucleotide means that any of the agents capable of cleaving a base portion of the non-canonical nucleotide described herein may be used.

Methods for labeling and fragmenting nucleic acids

[0084] The invention provides methods for generating fragments of nucleic acids. The methods generally comprise cleavage of the phosphodiester backbone of a polynucleotide comprising an abasic site at or near the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; and contacting the polynucleotide fragment with an enzyme capable of unblocking the blocked 3' end of the fragment, whereby a polynucleotide fragment comprising a 3' hydroxyl group is generated. The polynucleotide comprising an abasic site may be produced by cleaving the base portion of a non-canonical nucleotide, a canonical nucleotide, or a methylated nucleotide. In some embodiments, the polynucleotide comprising an abasic site is cleaved at the abasic site with a polyamine, thereby producing a polynucleotide fragment comprising a blocked 3' end, wherein the blocked 3' end

IDT Exhibit 1020.2811

comprises a sugar aldehyde group from the nucleotide residue from which the abasic site was generated. In some embodiments, the polynucleotide fragment comprising a blocked 3' end is contacted with an enzyme comprising a 3' to 5' exonuclease activity, thereby unblocking the blocked 3' end (*e.g.*, removing a sugar aldehyde group or other non-nucleotide moiety from the 3' end) to produce a polynucleotide fragment comprising a 3' hydroxyl group.

[0085] In one embodiment, the invention provides a method for fragmenting a polynucleotide comprising a blocked 3' end, comprising unblocking the blocked 3' end with an agent capable of unblocking the blocked 3' end of a polynucleotide, for example, an enzyme comprising a 3'-5' exonuclease activity as described herein. The polynucleotide comprising a blocked 3' end may be produced, for example, by cleaving a polynucleotide comprising an abasic site at or near the abasic site with an agent, for example, a polyamine, capable of cleaving a polynucleotide at or near an abasic site to produce a polynucleotide fragment comprising a blocked 3' end.

[0086] In some embodiments, polynucleotide fragments comprising 3' hydroxyl groups produced as described herein are labeled. In one embodiment, the polynucleotide fragment is labeled by extension at the 3' end with a template independent polymerase, resulting in addition of one or more labeled nucleotides (or nucleotides thereof) to the 3' end of the fragment. In one embodiment, the polynucleotide fragment is labeled by hybridizing the polynucleotide fragment to a polynucleotide template, extending from the 3' end with a template dependent polymerase, and incorporating labeled nucleotides into the extended polynucleotide fragment. In one embodiment, the polynucleotide fragment is labeled by ligation at the 3' end to a labeled polynucleotide with a ligase enzyme.

[0087] In one embodiment, a polynucleotide fragment comprising a 3' end hydroxyl group, produced as described herein, is extended from the 3' end using a template independent polymerase to produce a tailed 3' end which may comprises labeled and/or unlabeled nucleotides. Tailing of the fragment comprising a 3' end hydroxyl group may comprise polymerizing a plurality of dNTPs or ribo-NTPs at the 3' end.

IDT Exhibit 1020.2812

[0088] For simplicity, individual steps of the labeling and fragmentation method are discussed below. It is understood, however, that the steps may be performed simultaneously and/or in varied order, as discussed herein.

Synthesis of a polynucleotide comprising a non-canonical nucleotide

[0089] In some embodiments, methods of the invention comprise generation of a polynucleotide comprising an abasic site by cleaving the base portion of a non-canonical nucleotide in a polynucleotide comprising a non-canonical nucleotide. The polynucleotide comprising a non-canonical nucleotide may be produced by synthesizing a polynucleotide from a template in the presence of at least one non-canonical nucleotide (interchangeably termed "non-canonical deoxyribonucleoside triphosphate"), whereby a polynucleotide comprising a non-canonical nucleotide is generated. The frequency of incorporation of non-canonical nucleotides into the polynucleotide relates to the size of fragment produced using the methods of the invention because the spacing between non-canonical nucleotides in the polynucleotide comprising a non-canonical nucleotide, along with the reaction conditions used, determines the approximate size of the fragments resulting from generation of an abasic site from the non-canonical nucleotide and cleavage of the backbone at the abasic site, as described herein.

[0090] Generally, the polynucleotide is DNA, though, as noted herein, the polynucleotide can comprise altered and/or modified nucleotides, internucleotide linkages, ribonucleotides, etc. As generally used herein, it is understood that "DNA" applies to polynucleotide embodiments.

[0091] Methods for synthesizing polynucleotides, e.g., single and double stranded DNA, from a template are well known in the art, and include, for example, single primer isothermal amplification (SPIA™), Ribo-SPIA™, PCR, reverse transcription, primer extension, limited primer extension, replication (including rolling circle replication), strand displacement amplification (SDA), nick translation, multiple displacement amplification (MDA), rolling circle amplification (RCA) and, *e.g.*, any method that results in synthesis of the complement of a template sequence such that at least one non-canonical nucleotide can be incorporated into a

IDT Exhibit 1020.2813

polynucleotide. *See, e.g.,* Kurn, U.S. Patent No. 6,251,639; Kurn, WO 02/00938; Kurn, U.S. Patent No. 6,946,251, Kurn, U.S. Patent No. 6,692,918 ; Mullis, U.S. Patent No. 4,582,877; Wallace, U.S. patent No. 6,027,923; U.S. Patent No. 5,508,178; 5,888,819; 6,004,744; 5,882,867; 5,710,028; 6,027,889; 6,004,745; 5,763,178; 5,011,769; *see also* Sambrook (1989) "Molecular Cloning: A Laboratory Manual", second edition; Ausebel (1987, and updates) "Current Protocols in Molecular Biology"; Mullis, (1994) "PCR: The Polymerase Chain Reaction". One or more methods known in the art can be used to generate a polynucleotide comprising a non-canonical nucleotide. It is understood that the polynucleotide comprising a non-canonical nucleotide can be single stranded or double stranded or partially double stranded, and that one or both strands of a double stranded polynucleotide can comprise a non-canonical nucleotide. For convenience, "DNA" is used herein to describe (and exemplify) a polynucleotide. Suitable methods include methods that result in one single- or double-stranded polynucleotide comprising a non-canonical nucleotide (for example, reverse transcription, production of double stranded cDNA, a single round of DNA replication), as well as methods that result in multiple single stranded or double stranded copies or copies of the complement of a template (for example, single primer isothermal amplification or Ribo-SPIA™ or PCR). In one embodiment, a single-stranded polynucleotide comprising a non-canonical nucleotide is synthesized using single primer isothermal amplification. *See* Kurn, U.S. Patent Nos. 6,251,639 and 6,692,918.

[0092] Generally, the polynucleotide comprising a non-canonical nucleotide is generated from a template in the presence of all four canonical nucleotides and at least one non-canonical nucleotide under reaction conditions suitable for synthesis of polynucleotides, including suitable enzymes and primers, if necessary. Reaction conditions and reagents, including primers, for synthesizing the polynucleotide comprising a non-canonical nucleotide are known in the art, and further discussed herein. As described herein, non-canonical nucleotides are generally capable of polymerization (i.e., are substrates for DNA polymerase), and capable of being rendered abasic following treatment with a suitable agent capable of generally, specifically or selectively cleaving a base portion of a non-canonical nucleotide.

IDT Exhibit 1020.2814

Suitable non-canonical nucleotides are well-known in the art, and include: deoxyuridine triphosphate (dUTP), deoxyinosine triphosphate (dITP), 5-hydroxymethyl deoxycytidine triphosphate (5-OH-Me-dCTP). *See, e.g.,* Jendrisak, U.S. Patent No. 6,190,865 B1; *Mol. Cell Probes* (1992) 251-6. Generally, in embodiments in which a polynucleotide comprising an non-canonical nucleotide serves as a template for further amplification (e.g., as when multiple copies of a double stranded polynucleotides comprising a non-canonical nucleotide are synthesized, e.g., by PCR amplification), a polynucleotide comprising a non-canonical nucleotide must be capable of serving as a template for further amplification.

[0093] It is understood that two or more different non-canonical nucleotides can be incorporated into the polynucleotide synthesized from the template by DNA polymerase, whereby a polynucleotide comprising at least two different non-canonical nucleotides is generated.

[0094] Conditions for limited and/or controlled incorporation of a non-canonical nucleotide are known in the art. *See, e.g.,* Jendrisak, U.S. Patent No. 6,190,865 B1; *Mol. Cell Probes* (1992) 251-6; *Anal. Biochem.* (1993) 211:164-9; *see also* Sambrook (1989) "Molecular Cloning: A Laboratory Manual", second edition; Ausebel (1987, and updates) "Current Protocols in Molecular Biology". The frequency (or spacing) of non-canonical nucleotides in the resulting polynucleotide comprising a non-canonical nucleotide, and thus the average size of fragments generated using the methods of the invention (i.e., following cleavage of a base portion of a non-canonical nucleotide, and cleavage of a phosphodiester backbone at a non-canonical nucleotide), is controlled by variables known in the art, including: frequency of nucleotide(s) corresponding to the non-canonical nucleotide(s) in the template (or other measures of nucleotide content of a sequence, such as average G-C content), ratio of canonical to non-canonical nucleotide present in the reaction mixture; ability of the polymerase to incorporate the non-canonical nucleotide, relative efficiency of incorporation of non-canonical nucleotide verses canonical nucleotide, and the like. It is understood that the average fragmentation size also relates to the reaction conditions used during fragmentation, as is further discussed herein. The reaction conditions can be empirically determined, for example, by

IDT Exhibit 1020.2815

assessing average fragment size generated using the methods of the invention taught herein. The level of labeling at an abasic site also relates to the frequency of incorporation of non-canonical nucleotides, as is further discussed herein.

[0095] Generally, a non-canonical base can be incorporated at about every 5, 10, 15, 20, 25, 30, 40, 50, 65, 75, 85, 100, 123, 150, 175, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650 or more nucleotides apart in the resulting polynucleotide comprising a non-canonical nucleotide. In one embodiment, the non-canonical nucleotide is incorporated about every 200 nucleotides, about every 100 nucleotide, or about every 50 nucleotide. In another embodiment, the non-canonical nucleotide is incorporated about every 50 to about 200 nucleotides. In some embodiments, a 1:5 ratio of dUTP and dTTP is used in the reaction mixture.

[0096] The polynucleotide template (along which the polynucleotide comprising a non-canonical nucleotide is synthesized) may be any template from which labeled polynucleotide fragments are desired to be produced. As is evident from the description herein, the labeled polynucleotide fragments are the complement of the sequence of the polynucleotide template. The template includes double-stranded, partially double-stranded, and single-stranded nucleic acids from any source in purified or unpurified form, which can be DNA (dsDNA and ssDNA) or RNA, including tRNA, mRNA, rRNA, 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. Obtaining and purifying nucleic acids use standard techniques in the art. RNAs can be obtained and purified using standard techniques in the art. A DNA template (including genomic DNA template) can be transcribed into RNA form, which can be achieved using methods disclosed in Kurn, U.S. Patent No. 6,251,639 B1, and by other techniques (such as expression systems) known in the art. RNA copies of genomic DNA would generally include untranscribed sequences generally not found in mRNA, such as introns, regulatory and control elements, etc. DNA copies of an RNA template can be synthesized using methods described in Kurn, U.S. Patent No. 6,946,251 or other techniques known in the art). Synthesis of polynucleotide comprising a non-canonical

IDT Exhibit 1020.2816

nucleotide from a DNA-RNA hybrid can be accomplished by denaturation of the hybrid to obtain a ssDNA and/or RNA, cleavage with an agent capable of cleaving RNA from an RNA/DNA hybrid, and other methods known in the art. The template 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. The template 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 each other. Therefore, the methods of the invention are useful not only for producing one specific polynucleotide comprising a non-canonical nucleotide, but also for producing simultaneously more than one different specific polynucleotides comprising a non-canonical nucleotide. The template DNA can be a sub-population of nucleic acids, for example, a subtractive hybridization probe, total genomic DNA, restriction fragments, a cDNA library, cDNA prepared from total mRNA, a cloned library, or amplification products of any of the templates described herein. In some cases, the initial step of the synthesis of the complement of a portion of a template nucleic acid sequence is template denaturation. The denaturation step may be thermal denaturation or any other method known in the art, such as alkali treatment.

[0097] For simplicity, the polynucleotide comprising a non-canonical nucleotide is described as a single nucleic acid. It is understood that the polynucleotide can be a single polynucleotide, or a population of polynucleotides (from a few to a multiplicity to a very large multiplicity of polynucleotides). It is further understood that a polynucleotide comprising a non-canonical nucleotide can be a multiplicity (from small to very large) of different polynucleotide molecules. Such populations can be related in sequence (e.g., member of a gene family or superfamily) or extremely diverse in sequence (e.g., generated from all mRNA, generated from all genomic DNA, etc.). Polynucleotides can also correspond to single sequence (which can be part or all of a known gene, for example a coding region, genomic portion, etc.). Methods, reagents, and reaction conditions for generating specific polynucleotide sequences and multiplicities of polynucleotide sequences are known in the art.

IDT Exhibit 1020.2817

[0098] Suitable methods of synthesis of a polynucleotide comprising a non-canonical nucleotide are generally template-dependent (in the sense that polynucleotide comprising a non-canonical nucleotide is synthesized along a polynucleotide template, as generally described herein). It is understood that non-canonical nucleotides can be incorporated into a polynucleotide as a result of template-independent methods. For example, one or more primer(s) can be designed to comprise one or more non-canonical nucleotides. *See, e.g.,* Richards, U.S. Patent Nos. 6,037,152, 5,427,929, and 5,876,976. As discussed herein, inclusion of at least one non-canonical nucleotide in a primer results in cleavage of a base-portion of a non-canonical nucleotide and labeling at the abasic site (i.e., following generation of an abasic site, as described herein), thus generating a polynucleotide fragment or a labeled polynucleotide fragment comprising a portion of the primer. Inclusion of a non-canonical nucleotide in a primer may be particularly suitable for methods such as single primer isothermal amplification. *See* Kurn, U.S. Patent No. 6,251,639 B1; Kurn, WO 02/00938; Kurn, U.S. Patent Publication No. 2003/0087251 A1. Non-canonical nucleotide(s) can also be added to a polynucleotide by template-independent methods such as tailing or ligation of a second polynucleotide comprising a non-canonical nucleotide. Methods for tailing and ligation are well-known in the art.

Cleavage of a base portion of a nucleotide to create an abasic site

[0099] In methods for fragmentation of polynucleotides as described herein, a polynucleotide comprising an abasic site is cleaved at or near the abasic site to generate a polynucleotide fragment with a blocked 3' end, and contacted with an enzyme capable of unblocking the blocked 3' end to generate a polynucleotide fragment with a 3' hydroxyl group. The polynucleotide comprising an abasic site may be generated by cleaving the base portion of a nucleotide to create an abasic site. In various embodiments, the nucleotide from which the base portion is cleaved to create an abasic site is a non-canonical nucleotide, a canonical nucleotide, or a methylated nucleotide.

IDT Exhibit 1020.2818

Cleaving a base portion of a non-canonical nucleotide to create an abasic site

[0100] A polynucleotide comprising a non-canonical nucleotide is treated with an agent, such as an enzyme, capable of generally, specifically, or selectively cleaving a base portion of the non-canonical deoxyribonucleoside to create an abasic site. As used herein, "abasic site" encompasses any chemical structure remaining following removal of a base portion (including the entire base) with an agent capable of cleaving a base portion of a nucleotide, *e.g.*, by treatment of a non-canonical nucleotide (present in a polynucleotide chain) with an agent (*e.g.*, an enzyme, acidic conditions, or a chemical reagent) capable of effecting cleavage of a base portion of a non-canonical nucleotide. In some embodiments, the agent (such as an enzyme) catalyzes hydrolysis of the bond between the base portion of the non-canonical nucleotide and a sugar in the non-canonical nucleotide to generate an abasic site comprising a hemiacetal ring and lacking the base (interchangeably called "AP" site), though other cleavage products are contemplated for use in the methods of the invention. Suitable agents and reaction conditions for cleavage of base portions of non-canonical nucleotides are known in the art, and include: N-glycosylases (also called "DNA glycosylases" or "glycosidases") including Uracil N-Glycosylase ("UNG"; specifically cleaves dUTP) (interchangeably termed "uracil DNA glyosylase"), hypoxanthine-N-Glycosylase, and hydroxy-methyl cytosine-N-glycosylase; 3-methyladenine DNA glycosylase, 3- or 7- methylguanine DNA glycosylase, hydroxymethyluracil DNA glycosylase; T4 endonuclease V. *See, e.g.*, Lindahl, *PNAS* (1974) 71(9):3649-3653; Jendrisak, U.S. Patent No. 6,190,865 B1. In one embodiment, UNG is used to cleave a base portion of the non-canonical nucleotide.

[0101] Generally, cleavage of base portions of non-canonical nucleotides is general, specific or selective cleavage, in the sense that the agent (such as an enzyme) capable of cleaving a base portion of a non-canonical nucleotide generally, specifically or selectively cleaves the base portion of a particular non-canonical nucleotide, whereby greater than about 98%, about 95%, about 90%, about 85%, or about 80% of the base portions cleaved are base portions of non-canonical nucleotides. However, extent of cleavage can be less. Thus, reference to specific cleavage is exemplary. General, specific or selective cleavage is desirable for control

IDT Exhibit 1020.2819

of the fragment size in the methods of generating labeled polynucleotide fragments of the invention (i.e., the fragments generated by cleavage of the backbone at an abasic site). Generally, reaction conditions are selected such that the reaction in which the abasic site(s) are created can run to completion.

[0102] In some embodiments, the polynucleotide comprising a non-canonical nucleotide is purified following synthesis of the non-canonical polynucleotide (to eliminate, for example, residual free non-canonical nucleotides that are present in the reaction mixture). In other embodiments, there is no intermediate purification between the synthesis of the polynucleotide comprising the non-canonical nucleotide and subsequent steps (such as cleavage of a base portion of the non-canonical nucleotide and cleavage of a phosphodiester backbone at the abasic site).

[0103] As noted herein, for convenience, cleavage of a base portion of a non-canonical nucleotide (whereby an abasic site is generated) has been described as a separate step. It is understood that this step may be performed simultaneously with synthesis of the polynucleotide comprising a non-canonical nucleotide (as described above), cleavage of the backbone at an abasic site (fragmentation) and/or labeling at an abasic site.

[0104] It is understood that the choice of non-canonical nucleotide can dictate the choice of enzyme to be used to cleave the base portion of that non-canonical enzyme, to the extent that particular non-canonical nucleotides are recognized by particular enzymes that are capable of cleaving a base portion of the non-canonical nucleotide.

Cleaving a base portion of a canonical nucleotide to create an abasic site

[0105] In another aspect, the invention comprises use of an agent, such as an enzyme, that cleaves a base portion of a canonical nucleotide, to generate a polynucleotide comprising an abasic site. In some embodiments, the agent is not capable of cleaving a methylated nucleotide.

[0106] In some embodiments, the agent is an enzyme. In one embodiment, the enzyme is cytosine deaminase. See Sohail et al, NAR 2003, 31: 2990-94. Cytosine deaminase catalyzes the deamination of cytosine, such that dUTP is

IDT Exhibit 1020.2820

generated. Cleavage of a base portion of dUTP is necessary to generate the abasic site. Thus, the invention encompasses use of (a) an agent (such as cytosine deaminase) that modifies a nucleotide (such as dCTP), whereby dUTP is generated, in conjunction with (b) an agent (such as an enzyme, such as UNG) that cleaves a base portion of dUTP, whereby an abasic site is generated. Methods for cleaving a base portion of dUTP are known in the art. See, e.g., Lindahl, PNAS (1974) 71(9):3649-3653; Jendrisak, U.S. Patent No. 6,190,865 B1; U.S. Patent No. 5,035,996; U.S. Patent No. 5,418,149; Sartori et al (2002) EMBO J 21:3182-3191. As used herein, "in conjunction" encompasses simultaneous treatment (e.g., when cytosine deaminase and UNG cleavage occurs in the same reaction mixture) and/or treatment at different times (e.g., when cytosine deaminase and UNG treatment is conducted sequentially).

[0107] Generally, cleavage of the base portion of a canonical nucleotide is general, specific or selective cleavage (in the sense that the agent (such as an enzyme) capable of cleaving a base portion of a canonical nucleotide generally, specifically or selectively cleaves the base portion of a particular canonical nucleotide, whereby about any of 98%, 95%, 90%, 85%, or 80% of the base portions cleaved are base portions of canonical nucleotides. However, extent of cleavage can be less. Thus, reference to specific cleavage is exemplary.

[0108] It is understood that the frequency (or spacing) of abasic sites in the resulting polynucleotide comprising an abasic site (following cleavage of a base portion of a unmethylated nucleotide, and thus the average size of fragments generated using the methods of the invention (i.e., following cleavage of a phosphodiester backbone at an abasic site), is controlled by variables known in the art, including: frequency of the target canonical nucleotide(s) from which an abasic site will be generated in the polynucleotide (or other measures of nucleotide content of a sequence, such as average G-C content), length of the polynucleotide, and the reaction conditions used during generation of abasic site.

Cleaving a base portion of a methylated nucleotide to create an abasic site

[0109] In aspects involving cleavage of a base portion of a methylated nucleotide to generate a polynucleotide comprising an abasic site, the polynucleotide

IDT Exhibit 1020.2821

comprising a methylated nucleotide (in some embodiments, suspected of comprising a methylated nucleotide) is treated with an agent, such as an enzyme, capable of generally, specifically, or selectively cleaving a base portion of the methylated deoxyribonucleoside to create an abasic site. As used herein, "abasic site" encompasses any chemical structure remaining following removal of a base portion (including the entire base) of a methylated nucleotide with an agent capable of cleaving a base portion of a methylated nucleotide, e.g., by treatment of a methylated nucleotide (present in a polynucleotide chain) with an agent (e.g., an enzyme) capable of effecting cleavage of a base portion of a methylated nucleotide. In some embodiments, the agent (such as an enzyme) catalyzes hydrolysis of the bond between the base portion of the methylated nucleotide and a sugar in the methylated nucleotide to generate an abasic site comprising a hemiacetal ring and lacking the base (interchangeably called "AP" site), though other cleavage products are contemplated for use in the methods of the invention. Generally, the methods involving cleavage of a base portion of a methylated nucleotide are suitable for use with polynucleotides comprising a low frequency of methylated nucleotides (i.e., generally, not hypermethylated regions, such a CpG islands and the like), though other uses are contemplated.

[0110] The polynucleotide comprising (in some embodiments, suspected of comprising) a methylated nucleotide is treated with an agent, such as an enzyme, capable of generally, specifically, or selectively cleaving a base portion of the methylated deoxyribonucleoside to create an abasic site. As used herein, "abasic site" encompasses any chemical structure remaining following removal of a base portion (including the entire base) of a methylated nucleotide with an agent capable of cleaving a base portion of a methylated nucleotide, e.g., by treatment of a methylated nucleotide (present in a polynucleotide chain) with an agent (e.g., an enzyme, acidic conditions, or a chemical reagent) capable of effecting cleavage of a base portion of a methylated nucleotide. In some embodiments, the agent (such as an enzyme) catalyzes hydrolysis of the bond between the base portion of the methylated nucleotide and a sugar in the canonical nucleotide to generate an abasic site comprising a hemiacetal ring and lacking the base (interchangeably called "AP" site),

IDT Exhibit 1020.2822

though other cleavage products are contemplated for use in the methods of the invention.

[0111] In some embodiments, the methylated nucleotide is 5-methylcytosine. In some embodiments, 5-methylcytosine is present as a methylated CpG dinucleotide. The CpG dinucleotide may be fully or hemi-methylated. In other embodiments, the methylated nucleotide is 3-methyladenine. In other embodiments, the methylated nucleotide is 7-methyladenine and/or 3-methylguanine.

[0112] Suitable agents and reaction conditions for cleavage of base portions of methylated nucleotides are known in the art, and include: 5-methylcytosine DNA glycosylase (5-MCDG), which cleaves the base portion of 5-methylcytosine (5-MeC) from the DNA backbone (Wolffe et al., Proc. Nat. Acad. Sci. USA 96:5894-5896, 1999); 3-methyladenosine-DNA glycosylase I, which cleaves the base portion of 3-methyl adenosine from the DNA backbone (see, e.g. Hollis et al (2000) Mutation Res. 460: 201-210); and/or 3-methyladenosine DNA glycosylase II, which cleaves the base portion of 3-methyladenosine, 7-methylguanine, 7-methyladenosine, and/ 3-methylguanine from the DNA backbone. *See* McCarthy et al (1984) EMBO J. 3:545-550. Multifunctional and mono-functional forms of 5-MCDG have been described. *See* Zhu et al., Proc. Natl. Acad. Sci. USA 98:5031-6, 2001; Zhu et al., Nuc. Acid Res. 28:4157-4165, 2000; and Neddermann et al., J.B.C. 271:12767-74, 1996 (describing bifunctional 5-MCDG; Vairapandi & Duker, Oncogene 13:933-938, 1996; Vairapandi et al., J. Cell. Biochem. 79:249-260, 2000 (describing mono-functional enzyme comprising 5-MCDG activity). In some embodiments, 5-MCDG preferentially cleaves fully methylated polynucleotide sites (e.g., CpG dinucleotides), and in other embodiments, 5-MCDG preferentially cleaves a hemi-methylated polynucleotide. For example, mono-functional human 5-methylcytosine DNA glycosylase cleaves DNA specifically at fully methylated CpG sites, and is relatively inactive on hemimethylated DNA (Vairapandi & Duker, supra; Vairapandi et al., supra). By contrast, chick embryo 5-methylcytosine-DNA glycosylase has greater activity directed to hemimethylated methylation sites. In some embodiments, the activity of 5-MCDG is potentiated (increased or enhanced) with accessory factors, such as recombinant CpG-rich RNA, ATP, RNA helicase enzyme, and proliferating

IDT Exhibit 1020.2823

cell nuclear antigen (PCNA). See U.S. Patent Publication No. 20020197639 A1. One or more agents may be used. In some embodiments, the one or more agents cleave a base portion of the same methylated nucleotide. In other embodiments, the one or more agents cleave a base portion of different methylated nucleotides. Treatment with two or more agents may be sequential or simultaneous.

[0113] As is evident, in some embodiments, dUTP is generated as an intermediate and cleavage of a base portion of dUTP is necessary to generate the abasic site. Methods for cleaving a base portion of dUTP are known in the art. *See, e.g.*, Lindahl, PNAS (1974) 71(9):3649-3653; Jendrisak, U.S. Patent No. 6,190,865 B1; U.S. Patent No. 5,035,996; U.S. Patent No. 5,418,149; Sartori et al (2002) EMBO J 21:3182-3191. Thus, in some embodiments, an agent that cleaves a base portion of a methylated nucleotides (such as an enzyme, such as 5-MCDG) is used in conjunction with UNG to generate an abasic site from the methylated nucleotide. As used herein, "in conjunction" encompasses simultaneous treatment (e.g., when 5-MCDG and UNG cleavage occurs in the same reaction mixture) and/or treatment at different times (e.g., when 5-MCDG and UNG treatment is conducted sequentially).

[0114] In some embodiments, cleavage of the base portion of the methylated nucleotides is general, specific or selective cleavage (in the sense that the agent (such as an enzyme) capable of cleaving a base portion of a methylated nucleotide generally, specifically or selectively cleaves the base portion of a particular methylated nucleotide), whereby about any of 98%, 95%, 90%, 85%, or 80% of the base portions cleaved are base portions of methylated nucleotides. However, extent of cleavage can be less. Thus, reference to specific cleavage is exemplary.

[0115] It is understood that the frequency (or spacing) of abasic sites in the resulting polynucleotide comprising an abasic site (following cleavage of a base portion of a methylated nucleotide, and thus the average size of fragments generated using the methods of the invention (i.e., following cleavage of a phosphodiester backbone at an abasic site), is controlled by variables known in the art, including: frequency of methylated nucleotide(s) in the polynucleotide (or other measures of nucleotide content of a sequence, such as average G-C content), length of the

IDT Exhibit 1020.2824

polynucleotide comprising a methylated nucleotide, and the reaction conditions used during generation of abasic site, as is further discussed herein.

Cleaving the backbone at or near the abasic site to generate a polynucleotide fragment with a blocked 3' end

[0116] The backbone of the polynucleotide comprising an abasic site is cleaved at or near the abasic site with an agent that generates a polynucleotide fragment with a blocked 3' end. It is understood that cleavage of the base portion of a nucleotide to create an abasic site and cleavage of the polynucleotide backbone can be performed simultaneously. For convenience, however, these reactions are described as separate steps.

[0117] Following generation of an abasic site by cleavage of the base portion of a nucleotide, for example, a non-canonical nucleotide present in the polynucleotide, the backbone of the polynucleotide is cleaved at or near the abasic site, for example, the site of incorporation of a non-canonical nucleotide (also termed the abasic site, following cleavage of the base portion of the non-canonical nucleotide), with an agent capable of effecting cleavage of the backbone at the abasic site to generate a polynucleotide fragment comprising a blocked 3' end. Cleavage of the polynucleotide backbone (also termed "fragmentation") results in at least two fragments (depending on the number of abasic sites present in the polynucleotide comprising an abasic site, and the extent of cleavage), one of which does not comprise a blocked 3' end.

[0118] Suitable agents (for example, an enzyme, a chemical and/or reaction conditions such as heat) capable of cleavage of the backbone at an abasic site to generate a polynucleotide fragment with a blocked 3' end are well known in the art, and include: heat treatment and/or chemical treatment (including basic conditions, acidic conditions, alkylating conditions, or amine mediated cleavage of abasic sites, (*see e.g.*, McHugh and Knowland, *Nucl. Acids Res.* (1995) 23(10):1664-1670; *Bioorgan. Med. Chem* (1991) 7:2351; Sugiyama, *Chem. Res. Toxicol.* (1994) 7: 673-83; Horn, *Nucl. Acids. Res.*, (1988) 16:11559-71). As used herein, "agent" encompasses reaction conditions such as heat. In another embodiment, cleavage is

IDT Exhibit 1020.2825

with a polyamine, such as N, N'-dimethylethylenediamine (DMED). *See, e.g.* McHugh and Knowland, *supra.*

[0119] Generally, cleavage is between the nucleotide immediately 3' to the abasic residue and the abasic residue. As is well known in the art, cleavage can be 3' to the abasic site (e.g., cleavage between the deoxyribose ring and 3' -phosphate group of the abasic residue and the deoxyribose ring of the adjacent nucleotide, generating a free 5' phosphate group on the deoxyribose ring of the adjacent nucleotide), such that an abasic site is located at the 3' end of the resulting fragment. Treatment under basic conditions or with amines (such as N, N'-dimethylethylenediamine) results in cleavage of the phosphodiester backbone immediately 3' to the abasic site to produce a polynucleotide fragment with a blocked 3' end. In addition, more complex forms of cleavage are also possible, for example, cleavage such that cleavage of the phosphodiester backbone and cleavage of (a portion of) the abasic nucleotide results. For example, under certain conditions, cleavage using chemical treatment and/or thermal treatment may comprise a β-elimination step which results in cleavage of a bond between the abasic site deoxyribose ring and its 3' phosphate, generating a reactive α,β-unsaturated aldehyde which can be labeled or can undergo further cleavage and cyclization reactions. *See, e.g.* Sugiyama, *Chem. Res. Toxicol.* (1994) 7: 673-83; Horn, *Nucl. Acids. Res.*, (1988) 16:11559-71. It is understood that more than one method of cleavage can be used, including two or more different methods which result in multiple, different types of cleavage products comprising blocked 3' ends.

[0120] Generally, cleavage of the backbone at an abasic site is general, specific or selective cleavage, whereby greater than about 98%, about 95%, about 90% , about 85%, or about 80% of the cleavage is at an abasic site. However, extent of cleavage can be less. Thus, reference to specific cleavage is exemplary. General, specific or selective cleavage is desirable for control of the fragment size in the methods of generating labeled polynucleotide fragments of the invention. In some embodiments, reaction conditions can be selected such that the cleavage reaction is performed in the presence of a large excess of reagents and allowed to run to completion with minimal concern about excessive cleavage of the polynucleotide (i.e., while retaining a desired fragment size, which may be determined by spacing of

IDT Exhibit 1020.2826

incorporated non-canonical nucleotides, during the synthesis step, above). In other embodiments, extent of cleavage can be less, such that polynucleotide fragments are generated comprising an abasic site at an end and an abasic site(s) within or internal to the polynucleotide fragment (i.e., not at an end).

[0121] As noted herein, in embodiments in which an abasic site is generated by cleavage of a base portion of a non-canonical nucleotide in a polynucleotide synthesized in the presence of a non-canonical nucleotide, the frequency of incorporation of non-canonical nucleotides into the polynucleotide relates to the size of fragment produced using the methods of the invention because the spacing between non-canonical nucleotides in the polynucleotide comprising a non-canonical nucleotide, as well as the reaction conditions selected, determines the approximate size of the resulting fragments (following cleavage of a base portion of a non-canonical nucleotide, whereby an abasic site is generated, and cleavage of the backbone at the abasic site as described herein).

[0122] In methods of the invention, suitable fragment sizes are generally about 5, 10, 15, 20, 25, 30, 40, 50, 65, 75, 85, 100, 123, 150, 175, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650 or more nucleotides in length. In some embodiments, the fragment is about 200 nucleotides, about 100 nucleotides, or about 50 nucleotides in length. In another embodiment, the size of a population of fragments is about 50 to 200 nucleotides. It is understood that the fragment size is approximate, particularly when populations of fragments are generated, because the incorporation of a non-canonical nucleotide (which relates to the fragment size following cleavage) will vary from template to template, and also between copies of the same template. Thus, fragments generated from same starting material (such as a single polynucleotide template) may have different (and/or overlapping) sequence, while still having the same approximate size or size range.

[0123] Following cleavage of the polynucleotide backbone at the abasic site, every fragment will comprise one abasic site (if cleavage is completely efficient), except for the 3'-most fragment, which will lack an abasic site. All other fragments will comprise a 3' abasic site (a blocked 3' end).

IDT Exhibit 1020.2827

Unblocking of the blocked 3' end to generate a polynucleotide fragment with a 3' hydroxyl group

[0124] A polynucleotide fragment comprising a blocked 3' end, prepared as described herein, is contacted with an enzyme capable of unblocking the blocked 3' end, whereby a polynucleotide fragment comprising a 3' hydroxyl group is generated. In some embodiments, the enzyme capable of unblocking the blocked 3' end comprises a 3' to 5' exonuclease activity, generally a non-processive activity. In some embodiments, the enzyme that comprises a 3' to 5' exonuclease activity does not comprise an endonuclease activity, although an enzyme comprising an endonuclease activity may also be used, under conditions in which the endonuclease activity is minimized or absent. In some embodiments, the enzyme comprising a 3' to 5' exonuclease activity is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1.

[0125] Various 3'-5' exonucleases may be used for removal of the 3' blocking group following fragmentation of a polynucleotide comprising an abasic site with an agent, such as a polyamine, for example, DMED, that produces a fragment with a blocked 3' end. A review of 3'-5' exonucleases is presented in Shevelev et al. (2002) *Nature Reviews Molecular Cell Biology* 3:367-376. Many DNA repair related exonucleases have been discovered in recent years in addition to APE1 and homologous AP nucleases in other eukaryotes and prokaryotes. Of special interest are the TREX1 and two 3'-5' exonucleases capable of unblocking a blocked polynucleotide 3' terminus, as described in Mazur et al., *J. Biol. Chem.* 276(20):17033-17029. Another effective exonuclease is APE2, which efficiently removes 3' blocking groups from polynucleotide 3' termini, as described in Burkovics et al. (2006) *Nucleic Acids Res.* 34(9):2508-2515.

[0126] In some embodiments, an important feature of the 3'-5' exonuclease is the ability to remove a blocking group at the 3' terminus of a single stranded polynucleotide (such as a single stranded amplification product). In some embodiments, the exonuclease is non-processive. An example of such a non-processive exonuclease is human TREX2 3'-5' exonuclease, as described in Perrino et al., *J. Biol. Chem.* 280(15): 15212-15218.

IDT Exhibit 1020.2828

Polymerase extension or ligation of polynucleotide fragments with 3' hydroxyl groups

[0127] A polynucleotide fragment with an unblocked 3' end, prepared as described herein, may be extended from the hydroxyl group at the 3' end by a template independent or template dependent polymerase or may be ligated at the 3' end to another polynucleotide with a ligase enzyme.

[0128] In one embodiment, a polynucleotide fragment with an unblocked 3' end, prepared as described herein, is extended from the 3' hydroxyl group with a template independent polymerase, such as TdT, to incorporate one or more nucleotides, for example, one or more detectable nucleotides, at the 3' end of the polynucleotide fragment. In some embodiments, a labeled nucleotide is incorporated. In one embodiment, the labeled nucleotide is a biotinylated nucleotide, such as a biotinylated triphosphate (NTP), deoxynucleotide triphosphate (dNTP), or dideoxynucleotide triphosphate (ddNTP). For example, biotin 2'3'-dideoxy-UTP or biotin 2',3'-dideoxy-CTP may be incorporated at the 3' end of the polynucleotide fragment with TdT. In another embodiment, the labeled nucleotide comprises a fluorophore (*e.g.*, cy dyes, alexa dyes, fluorescein, and other fluorophores known in the art). In other embodiments, the incorporated detectable nucleotide comprises an enzyme, a chromophore, a radiolabel, or a hapten which is detectable by binding of a labeled second member of a binding pair, such as, for example, biotin/avidin or streptavidin, antigen/antibody, and the label attached to the second member of the binding pair may comprise, for example, a fluorophore, an enzyme, a chromophore, a radiolabel, or may be attached to a detectable particle. In one embodiment, the labeled nucleotide is a biotinylated nucleotide, for example, biotin 2',3'-dideoxy-UTP or 2',3'-dideoxy-CTP, and is detectable by binding of labeled avidin or streptavidin. In some embodiments, a polynucleotide fragment prepared in accordance with methods described herein is tailed with unlabeled nucleotides or a mixture of labeled and unlabeled nucleotides with a template independent polymerase such as TdT.

[0129] In another embodiment, a polynucleotide fragment with an unblocked 3' end, prepared as described herein, is hybridized to a polynucleotide template and extended from the 3' hydroxyl group with a template dependent polymerase. In one

IDT Exhibit 1020.2829

embodiment, the polynucleotide fragment with an unblocked 3' end is used as a primer to initiate synthesis of a polynucleotide complementary to the template. In one embodiment, the polynucleotide fragment is extended in the presence of one or more labeled nucleotides, such as one or more nucleotides attached to a member of a binding pair that is detectable by binding of a labeled second member of the binding pair, as described above, to produce a detectable polynucleotide.

[0130] In another embodiment, a polynucleotide fragment with an unblocked 3' end, prepared as described herein, is ligated at the 3' hydroxyl group to another polynucleotide using a ligase enzyme. In some embodiments, the polynucleotide to which the polynucleotide fragment is ligated comprises one or more detectable nucleotides as described above, resulting in a detectable ligated polynucleotide comprising the polynucleotide fragment with the unblocked 3' end and the polynucleotide to which the fragment was ligated.

[0131] As discussed above, a "label" can be directly detectable, or the label can be indirectly detectable, such as, for example, when the label is covalently or non-covalently associated with another moiety which is itself detectable. For example, biotin can be attached to nucleotide, which may be detected by binding to detectable avidin or streptavidin. In another example, an antibody (that can be detectably labeled) binds to a cognate antigen that is attached to a nucleotide. In some embodiments, the label comprises an organic molecule, a hapten, or a particle (such as a polystyrene bead). In some embodiments, the label is detected using antibody binding, biotin binding, or via fluorescence or enzyme activity. In some embodiments, the detectable signal is amplified.

[0132] In some embodiments, labeled polynucleotide fragments are produced which each comprise a single label, for example, incorporation of a detectable nucleotide at the 3' hydroxyl group with terminal transferase. This aspect is useful in quantitating level of hybridization, because signal is proportional to number of bound fragments, and does not relate to the length of the hybridizing fragment or the number of labels per fragment. Thus, hybridization intensity can generally be directly compared, regardless of fragment length. This offers an advantage over prior art methods in which nucleic acid fragments are labeled with multiple detectable

IDT Exhibit 1020.2830

moieties, e.g., incorporation of labeled nucleotides, and other methods of directly and indirectly detecting incorporated nucleotides. These methods generally result in multiple labels per hybridizing fragment, and thus are generally less suitable for quantitative applications. Multiple labels per nucleic acid can result in quenching, and potential interference with hybridization kinetics (due to the presence of multiple labeled moieties per nucleic acid).

[0133] Methods of signal detection are known in the art. Signal detection may be visual or utilize a suitable instrument appropriate to the particular label used, such as a spectrometer, fluorimeter, or microscope. For example, where the label is a radioisotope, detection can be achieved using, for example, a scintillation counter, or photographic film as in autoradiography. Where a fluorescent label is used, detection may be by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence, such as by microscopy, visual inspection or photographic film, fluorometer, CCD cameras, scanner and the like. Where enzymatic labels are used, detection may be by providing appropriate substrates for the enzyme and detecting the resulting reaction product. For example, many substrates of horseradish peroxidase, such as o-phenylenediamine, give colored products. Simple colorimetric labels can usually be detected by visual observation of the color associated with the label; for example, conjugated colloidal gold is often pink to reddish, and beads appear the color of the bead. Instruments suitable for high sensitivity detection are known in the art.

[0134] It is understood that the polynucleotide or polynucleotide fragments can be additionally labeled using other methods known in the art, such as incorporation of labeled nucleotide analogs during synthesis of a polynucleotide comprising a non-canonical nucleotide, from which a polynucleotide comprising an abasic site is generated. In addition, following cleavage of the phosphodiester backbone of the polynucleotide comprising an abasic site, the 3' most fragment will lack an abasic site, (in embodiments in which the fragmentation reaction goes to completion). However, if polynucleotide synthesis step requires primer(s), a labeled primer(s) can be used such that the resulting fragment comprising a primer is labeled. Suitable labels and methods of labeling primers are known. In addition, a primer

IDT Exhibit 1020.2831

comprising a non-canonical nucleotide can be used. Following generation of an abasic site, cleavage of the phosphodiester backbone at the abasic site, and labeling at the abasic site, the fragment comprising at least a portion of the primer will be labeled.

Reaction conditions and detection

[0135] Appropriate reaction media and conditions for carrying out the methods of the invention include those that permit cleavage of a polynucleotide comprising an abasic site with an agent capable of cleaving a polynucleotide to produce a polynucleotide fragment with a blocked 3' end, and unblocking of a blocked 3' end with an agent capable of unblocking the 3' end of a polynucleotide to produce a polynucleotide comprising a 3' end hydroxyl group.

[0136] Appropriate reaction media and conditions for carrying out the cleavage of the phosphodiester backbone at an abasic site according to the methods of the invention are those that permit cleavage of the phosphodiester backbone at or near an abasic site to produce a polynucleotide fragment comprising a blocked 3' end. Such media and conditions are known to persons of skill in the art, and are described in various publications, such as *Bioorgan. Med. Chem* (1991) 7:2351; Sugiyama, *Chem. Res. Toxicol.* (1994) 7: 673-83; Horn, *Nucl. Acids. Res.*, (1988) 16:11559-71); Lindahl, *PNAS* (1974) 71(9):3649-3653; Jendrisak, U.S. Patent No. 6,190,865 B1; Shida, *Nucleic Acids Res.* (1996) 24(22):4572-76; Srivastava, *J. Biol Chem.* (1998) 273(13):21203-209; Carey, *Biochem.* (1999) 38:16553-60; *Chem Res Toxicol* (1994) 7:673-683. A reaction mixture suitable for simultaneous UNG treatment and N, N'-dimethylethylenediamine treatment is described in Example 4 of U.S. Patent Application No. 2004/0005614.

[0137] In another example, nucleic acids containing abasic sites are heated in a buffer solution containing an amine, for example, 25mM Tris-HCl and 1-5 mM magnesium ions, for 10-30 minutes at 70°C to 95°C. Alternatively, 1.0 M piperidine (a base) is added to polynucleotide comprising an abasic site which has been precipitated with ethanol and vacuum dried. The solution is then heated for 30 minutes at 90°C and lyophilized to remove the piperidine. In another example,

IDT Exhibit 1020.2832

cleavage is effected by treatment with basic solution, e.g., 0.2 M sodium hydroxide at 37° for 15 minutes. *See* Nakamura (1998) *Cancer Res.* 58:222-225. In yet another example, incubation at 37C with 100 mM N, N'-dimethylethylenediamine acetate, pH 7.4 is used to cleave. *See* McHugh and Knowland, (1995) *Nucl. Acids Res.* 23(10) 1664-1670.

[0138] In some embodiments, nucleic acid synthesis is performed to produce the polynucleotide to be fragmented. Appropriate media and conditions are known to persons of skill in the art, and are described in various publications, such as U.S. Pat. Nos. 6,190,865; 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; PCT Pub. No. WO99/42618; *Mol. Cell Probes* (1992) 251-6; and *Anal. Biochem.* (1993) 211:164-9. 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^{2+} or Mn^{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 strand binding proteins (e.g., 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 (if any) can also be included. Any aspect of the methods of the invention can occur at the same or varying temperatures. The synthesis reactions (particularly,

IDT Exhibit 1020.2833

primer extension other than the first and second strand cDNA synthesis steps, and strand displacement) can be performed isothermally, which avoids the cumbersome thermocycling process. The synthesis reaction is carried out at a temperature that permits hybridization of the oligonucleotides (primer) of the invention to the template polynucleotide and primer extension products, 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. In some embodiments that include RNA transcription, the temperature for the transcription steps is lower than the temperature(s) for the preceding steps. In these embodiments, 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.

[0139] In embodiments in which the polynucleotide comprising an abasic site is produced from a polynucleotide comprising a non-canonical nucleotide, Nucleotides, including non-canonical nucleotides (or other nucleotide analogs), that can be employed for synthesis of the nucleic acid comprising a non-canonical nucleotide 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. The oligonucleotide components of the synthesis reactions of the invention are generally in excess of the number of template nucleic acid sequence to be replicated. 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}, 10^{12} times the amount of target nucleic acid. Composite 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.

[0140] Optionally, the polynucleotide comprising a non-canonical nucleotide can be treated with hydroxylamine (or any other suitable agent) to remove any aldehydes that may have formed spontaneously in the nucleic acid. *See, e.g.*, Makrogiorgos, WO00/39345.

[0141] For convenience, the synthesis of a polynucleotide comprising a non-canonical nucleotide, and the cleavage of a base portion of that polynucleotide by an

IDT Exhibit 1020.2834

enzyme capable of cleaving a base portion of the non-canonical nucleotide, and the cleavage of the phosphodiester backbone at the abasic site, are described as separate steps. It is understood that these steps may be performed simultaneously, except (generally) in the case when a polynucleotide comprising a non-canonical nucleotide must be capable of serving as a template for further amplification (as in exponential methods of amplification, e.g. PCR).

[0142] Appropriate reaction media and conditions for carrying out the cleavage of a base portion of a non-canonical nucleotide according to the methods of the invention are those that permit cleavage of a base portion of a non-canonical nucleotide. Such media and conditions are known to persons of skill in the art, and are described in various publications, such as Lindahl, *PNAS* (1974) 71(9):3649-3653; Jendrisak, U.S. Patent No. 6,190,865 B1; U.S. Patent No. 5,035,996; U.S. Patent No. 5,418,149. For example, buffer conditions can be as described above with respect to polynucleotide synthesis. In one embodiment, UNG (Epicentre Technologies, Madison WI) is added to a nucleic acid synthesis reaction mixture, and incubated at 37°C for 20 minutes. In one embodiment, the reaction conditions are the same for the synthesis of a polynucleotide comprising a non-canonical nucleotide and the cleavage of a base portion of the non-canonical nucleotide. In another embodiment, different reaction conditions are used for these reactions. In some embodiments, a chelating regent (e.g. EDTA) is added before or concurrently with UNG in order to prevent the polymerase from extending the ends of the cleavage products.

[0143] In some embodiments, some of the components for performing a method as described herein are added simultaneously at various timepoints. In one embodiment, components for cleaving a base portion of a nucleotide to generate a polynucleotide comprising an abasic site and components for cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site may be combined for simultaneous reaction. In one embodiment, components for cleaving a base portion of a nucleotide to generate a polynucleotide comprising an abasic site, components for cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site, and components for unblocking a blocked 3' end may be combined for simultaneous reaction. In one embodiment, components for cleaving

IDT Exhibit 1020.2835

the phosphodiester backbone of a polynucleotide comprising the abasic site, and components for unblocking a blocked 3' end may be combined for simultaneous reaction. Such components may be added in any order at appropriate timepoints. Such timepoints can be readily identified by a person of skill in the art. In these embodiments, the reaction conditions and components may be varied between the different reactions.

[0144] The fragmenting or fragmenting and labeling process 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. 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 (depleted) enzyme or other reagent. 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.

[0145] The reaction can be allowed to proceed without purification of intermediate reaction products. Alternatively, products can be purified at various timepoints, conditions for which can be readily identified by a person of skill in the art.

Compositions

[0146] The invention also provides compositions and kits used in the methods described herein. The compositions may comprise any component(s), reaction mixture(s) and/or intermediate(s) described herein, as well as any combination thereof. For example, in one embodiment, the invention provides a composition comprising an agent capable of cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide, an agent capable of cleaving a phosphodiester backbone at or near an abasic site to produce a polynucleotide fragment with a

IDT Exhibit 1020.2836

blocked 3' end, and an enzyme capable of unblocking a blocked 3' end to generate a polynucleotide comprising a 3' hydroxyl group. In one embodiment, the agent capable of cleaving a base portion of a nucleotide to generate an abasic site is an N-glycosylase, for example, UNG. In one embodiment, the agent capable of cleaving a phosphodiester backbone at or near the abasic site is a polyamine, for example, DMED. In one embodiment, the enzyme capable of unblocking a blocked 3' end comprises a 3'-5' exonuclease activity, preferably a non-processive exonuclease activity, for example, endonuclease 4, exonuclease T, or APE1. Compositions of the invention may also comprise buffers, co-factors, or other components for carrying out the reactions of the methods described herein.

[0147] The invention also provides a composition comprising a polynucleotide fragment produced by a method as described herein and a template independent polymerase, a template dependent polymerase, or a ligase. In one embodiment, the invention provides a composition comprising a polynucleotide fragment produced as described herein, a template independent polymerase, for example, TdT, and a labeled nucleotide, an unlabeled nucleotide, or a mixture or labeled and unlabeled nucleotides. In one embodiment, the invention provides a composition comprising a polynucleotide fragment produced as described herein, a template dependent polymerase, and a polynucleotide template to which the polynucleotide fragment is capable of hybridizing, optionally further comprising nucleotides for polymerization, *e.g.*, labeled, unlabeled, or a mixture of labeled and unlabeled nucleotides. In one embodiment, the composition comprises a complex comprising the polynucleotide fragment hybridized to a polynucleotide template. In one embodiment, the invention provides a composition comprising a polynucleotide fragment prepared as described herein, a ligase enzyme, and a polynucleotide to which the polynucleotide fragment is desired to be ligated.

[0148] The compositions are generally in lyophilized or in a suitable medium, such as aqueous form (if appropriate), preferably in a suitable buffer.

[0149] The invention also provides polynucleotide fragments and labeled polynucleotide fragments produced by any of the methods described herein, and compositions comprising such fragments. Accordingly, the invention provides a

IDT Exhibit 1020.2837

population of fragmented or fragmented and labeled polynucleotides, which are produced by any of the methods described herein (or compositions comprising the products).

[0150] The invention also provides reaction mixtures (or compositions comprising reaction mixtures) which contain various combinations of components described herein. Examples of reaction mixtures have been described. In one embodiment, the reaction mixture comprises a polynucleotide comprising an abasic site, and an agent (such as UNG) that is capable of cleaving a base portion from a non-canonical nucleotide, an agent (such as an amine, such as N, N'-dimethylethylenediamine) capable of cleaving the phosphodiester back at an abasic site, and an enzyme capable of unblocking a blocked 3' end of a polynucleotide. In one embodiment, the invention provides a reaction mixture comprising (a) UNG; (b) N, N'-dimethylethylenediamine; and (c) APE1.

Kits

[0151] The invention also provides kits for carrying out the methods of the invention. Accordingly, a variety of kits are provided in suitable packaging. The kits may contain instructions for carrying out any of the methods described herein for production of polynucleotide fragments or labeled polynucleotide fragments, or for one or more of the following applications using polynucleotide fragments prepared as described herein: methods of producing a hybridization probe; characterizing and/or quantitating nucleic acid; detecting a mutation; preparing a subtractive hybridization probe; detection (using a hybridization probe); and determining a gene expression profile, using the fragmented nucleic acids generated by the methods of the invention. The kits of the invention comprise one or more containers comprising any combination of the components described herein, and the following are examples of such kits.

[0152] In one embodiment, a kit of the invention comprises an agent capable of cleaving a base portion of a nucleotide to generate abasic site in a polynucleotide, an agent capable of cleaving a phosphodiester backbone of a polynucleotide at or near an abasic site to produce a polynucleotide fragment with a blocked 3' end, and an

IDT Exhibit 1020.2838

enzyme capable of unblocking the blocked 3' end to produce a polynucleotide fragment comprising a 3' end hydroxyl group. The kit may further comprise an agent capable of labeling the polynucleotide fragment comprising a 3' end hydroxyl group and/or a label to be incorporated into an extension or ligation product of the polynucleotide fragment. In one embodiment, the kit further comprises an enzyme capable of extending the 3' end of the polynucleotide fragment in a template independent or template dependent manner and optionally further comprises nucleotide substrates for such enzymes, either labeled or unlabeled or a mixture of labeled and unlabeled nucleotides. In one embodiment, the kit further comprises a ligase enzyme. In one embodiment, the kit comprises UNG, DMED, and APE1, and optionally further comprises TdT. In one embodiment, the kit further comprises components for synthesis of a polynucleotide to be fragmented, such as a primer, for example, a composite primer, and/or nucleotides, for example, canonical and/or non-canonical nucleotides.

[0153] Kits may also include one or more suitable buffers (as described herein) or any other necessary reagents for carrying out the reactions of the methods described herein. One or more reagents in the kit can be provided as a dry powder, usually lyophilized, including excipients, which on dissolution will provide for a reagent solution having the appropriate concentrations for performing any of the methods described herein. Each component can be packaged in separate containers or some components can be combined in one container where cross-reactivity and shelf life permit.

[0154] The kits of the invention may optionally include a set of instructions, generally written instructions, although electronic storage media (*e.g.*, magnetic diskette or optical disk) containing instructions are also acceptable, relating to the use of components of the methods of the invention for the intended methods of the invention, and/or, as appropriate, for using the products for purposes such as, for example preparing a hybridization probe, expression profiling, preparing a microarray, or characterizing a nucleic acid. The instructions included with the kit generally include information as to reagents (whether included or not in the kit) necessary for practicing the methods of the invention, instructions on how to use the

IDT Exhibit 1020.2839

kit, and/or appropriate reaction conditions. Instructions may be in the form of printed media, electronic media, or a reference to a website address where instructions may be obtained.

[0155] The component(s) of the kit may be packaged in any convenient, appropriate packaging. The components may be packaged separately, or in one or multiple combinations.

[0156] The relative amounts of the various components in the kits can be varied widely to provide for concentrations of the reagents that substantially optimize the reactions that need to occur to practice the methods disclosed herein and/or to further optimize the sensitivity of any assay.

Tailing of the fragmented polynucleotide and subsequent isothermal single primer amplification (SPIA™)

[0157] Fragmented polynucleotides may be tailed and subsequently amplified using the previously-described SPIA™ amplification procedure (U.S. Patent Nos. 6,251,639 and 6,692,918).

[0158] Single stranded cDNA may be produced in the presence of non-canonical nucleotides, for example, amplified from RNA or DNA (*e.g.*, amplified from RNA by the previously-described Ribo-SPIA™ method (U.S. Patent No. 6,946,251)) or amplified from DNA by the SPIA™ method (U.S. Patent Nos. 6,251,639 and 6,692,918), or generated without amplification by random priming or poly-A initiated reverse transcription using primers with or without a 5'-end tails (which can be RNA, for example, in a chimeric primer, or DNA). Single stranded cDNA is produced in the presence of non-canonical nucleotides and subsequently fragmented using a method of the invention as described herein. The fragmented single stranded DNA generated by the method of the invention comprises a 3'-hydroxyl terminus which can be extended by the template independent DNA synthesis using terminal transferase. Tailing of DNA using terminal transferase is well known in the art. The addition of homopolymeric tails, with deoxyribonucleotides has been previously described and is widely used for manipulation of cDNA (*see, e.g.,* G Deng et al. (1981) *Nucleic Acids Res.* 9(16): 4173–4188; Schmidt et al., *Nucleic Acids Res.*

IDT Exhibit 1020.2840

24(9) 1789-1791; Albuquerque-Silva et al., *Nucleic Acids Research* 26(13): 3314-3316; U.S. Patent No. 6,406,890). Tailing at the 3'-end of a polynucleotide comprising a 3' end hydroxyl group using terminal transferase is widely used in molecular biology and commercial reagents and kits for various manipulations, including cDNA cloning, are readily available.

[0159] It is desirable to generate cDNA copies of single stranded, fragmented cDNA for further analysis, especially for analysis based on hybridization to probes which are designed for hybridization to the second strand cDNA (hybridization to cDNA which is the same sense as mRNA). It is also desirable to linearly amplify second strand cDNA fragments. The generation of fragmented amplified cDNA comprising 3'-hydroxyl groups is described herein. Tailing of the fragmented amplified cDNA with terminal transferase can be achieved by methods described in the above references and well known in the art. It is desirable to limit the length of the tail, which can be achieved by controlling the reaction conditions, the amount of the incorporated dNTP or rNTP or the use of a mixture of dNTP in the presence of a terminator such as a ddNTP. Tailing of fragmented cDNA products (produced from a single mRNA or multiplicity of mRNA or total RNA) results in the generation of a whole representative population of fragmented cDNA (libraries) with a common consensus 3'-end sequence.

[0160] These products may be amplified using a first chimeric primer comprising a 3'-DNA portion that is hybridizable to the 3' tail sequence of the tailed fragmented cDNA and a 5'-RNA sequence which is not hybridizable to the fragmented cDNA products. DNA polymerase comprising both DNA-dependent RNA-dependent DNA polymerase activities may be used to extend the 3' ends of the fragmented and tailed cDNA along the hybridized chimeric primer and the 3'-end of the hybridized primer along the hybridized fragmented and tailed cDNA, to generate double stranded DNA products with an RNA/DNA heteroduplex at one end. The DNA-dependent DNA polymerase and RNA-dependent DNA polymerase may be activities of the same enzyme or of two different enzymes. Extension of the tailed fragmented polynucleotide along the first chimeric primer requires that the 3' end of the fragmented and tailed polynucleotide be unblocked. Therefore, it is desirable to

IDT Exhibit 1020.2841

tail the fragmented polynucleotide with nucleotides that do not serve as terminators. This requirement dictates the type of tailing reaction mixtures and conditions used for the tailing reaction. Amplification of this product may proceed by the addition of a second amplification chimeric primer that contains a sequence homologous to the sequence of the RNA portion of the first chimeric primer, RNase H, and a DNA polymerase with strand displacement activity, as previously described (U.S. Patent Nos. 6,251,239 and 6,692,918). Extension of the first primer along the fragmented polynucleotide is not essential to the process, and thus a 3'-blocked first chimeric primer may also be useful for this process.

[0161] The multiplicity of copies of the fragmented cDNA as described above is useful for further analysis and characterization by any of the methods described below. The single stranded amplification products can be labeled by incorporation of labeled nucleotides, tailing of the single stranded amplification products with labeled nucleotides using terminal transferase, and incorporation of nucleotides which can be labeled post-DNA synthesis, such as, for example, aminoallyl-dUTP or various non-enzymatic methods for labeling of nucleic acids (for example, ULS labeling, Kreatech).

Applications using the labeling and/or fragmentation and/or immobilization methods of the invention

[0162] The methods and compositions of the invention can be used for a variety of purposes. For purposes of illustration, methods of producing a hybridization probe, characterizing and/or quantitating nucleic acid, detecting a mutation, preparing a subtractive hybridization probe, detection (using the hybridization probe), and determining a gene expression profile, using the labeled and/or fragmented nucleic acids generated by the methods of the invention, are described.

[0163] Immobilized polynucleotides, for example on a microarray, prepared according to any of the methods of the invention, are also useful for methods of analyzing and characterizing nucleic acids, including methods of hybridizing nucleic acids, methods of characterizing and/or quantitating nucleic acids, methods of

IDT Exhibit 1020.2842

detecting a mutation in a nucleic acids, and methods of determining a gene expression profile, as described below, and these applications likewise apply to immobilized polynucleotides.

Method of producing a hybridization probe

[0164] Labeled polynucleotide fragments obtained by the methods of the invention are useful as hybridization probes. Accordingly, in one aspect, the invention provides methods for nucleic acid hybridization, comprising using a labeled polynucleotide fragment as a hybridization probe, wherein the labeled polynucleotide fragment is produced using a method as described herein. In one embodiment, the invention provides a method for generating hybridization probes, comprising generating labeled polynucleotides using any of the methods described herein, and using the labeled polynucleotides as a hybridization probe. In another embodiment, the invention provides methods for generating a hybridization probe, comprising generating labeled polynucleotide fragments using any of the methods described herein, and using the labeled polynucleotide fragments as a hybridization probe. The labeled polynucleotide fragments can be produced from any template known in the art, including RNA, DNA, genomic DNA (including global genomic DNA amplification), and libraries (including cDNA, genomic or subtractive hybridization library). The invention also provides methods of hybridizing using the hybridization probes described herein.

Characterization of nucleic acids

[0165] The labeled and/or fragmented nucleic acids obtained by the methods of the invention are amenable to further characterization.

[0166] The fragmented nucleic acids, or labeled fragments thereof (*i.e.*, products of any of the methods described herein), can be analyzed using, for example, probe hybridization techniques known in the art, such as Southern and Northern blotting, and hybridizing to probe arrays. They can also be analyzed by electrophoresis-based methods, such as differential display and size characterization, which are known in the art.

IDT Exhibit 1020.2843

[0167] In one embodiment, the methods of the invention are utilized to analyze polynucleotide fragments, wherein the polynucleotide fragments are generated using a method as described herein. In one embodiment, the invention provides a method for analyzing polynucleotides, comprising generating polynucleotide fragments, *e.g.*, labeled polynucleotide fragments, according to a method as described herein, and contacting the polynucleotide fragments with a probe. The polynucleotide fragments can be produced from any template known in the art, including RNA, DNA, genomic DNA (including global genomic DNA amplification), or amplified products thereof, and libraries (including cDNA, genomic or subtractive hybridization library).

[0168] In one embodiment, the methods of the invention are utilized to generate fragmented polynucleotides which are analyzed (for example, detection and/or quantification) by contacting them with, for example, microarrays (of any suitable substrate, 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 and/or identify) a fragmented and labeled polynucleotide by analyzing the labeled products, for example, by hybridization of the labeled products to, for example, probes immobilized at, for example, specific locations on a solid or semi-solid substrate, probes immobilized on defined particles (including beads, such as Bead Array, Illumina), or probes immobilized on blots (such as a membrane), for example arrays, or arrays of arrays. Immobilized probes include immobilized probes generated by the methods described herein, and also include at least the following: cDNA and synthetic oligonucleotides, which can be synthesized directly on the substrate.

[0169] 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 products and probes from solution. The identity of the probes provides characterization of the sequence identity of the products, and thus by extrapolation can also provide characterization of the identity of a template from which the products were prepared (for example, the identity of an RNA in a solution). For example, hybridization of the labeled products is detectable,

IDT Exhibit 1020.2844

and the amount of specific labels that are detected is proportional to the amount of the labeled products prepared from a specific 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. 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 template RNA species in the sample.

[0170] Methods of characterization include sequencing by hybridization (see, e.g., Dramanac, U.S. Patent No. 6,270,961) and global genomic hybridization (also termed comparative genome hybridization) (*see, e.g.*, Pinkel, U.S. Patent No. 6,159,685).

[0171] In another aspect, the invention provides a method of quantitating fragmented polynucleotides comprising use of an oligonucleotide (probe) of defined sequence (which may be immobilized, for example, on a microarray).

Mutation detection utilizing the methods of the invention

[0172] The fragmented polynucleotides generated according to the methods of the invention are also suitable for analysis for the detection of any alteration in the template nucleic acid sequence (from which the fragmented polynucleotides are synthesized), as compared to a reference nucleic acid sequence which is identical to the template nucleic acid sequence other than the sequence alteration. The sequence alterations may be sequence alterations present in the genomic sequence or may be sequence alterations which are not reflected in the genomic DNA sequences, for example, alterations due to post transcriptional alterations, and/or mRNA processing, including splice variants. Sequence alterations (interchangeably called "mutations") include deletion, substitution, insertion and/or transversion of one or more nucleotide.

[0173] Accordingly, the invention provides methods of detecting presence or absence of a mutation in a template, comprising analyzing a polynucleotide fragment generated using a method as described herein, whereby presence or absence of a mutation is detected. In one embodiment, the method comprises: (a) generating a

IDT Exhibit 1020.2845

polynucleotide fragment, *e.g.*, a labeled polynucleotide fragment, by any of the methods described herein; and (b) analyzing the polynucleotide fragment whereby presence or absence of a mutation is detected. In some embodiments, the polynucleotide fragment is compared to a labeled reference template, or fragments thereof. Analyzing the polynucleotide fragment, whereby presence or absence of a mutation is detected, can be performed by any method known in the art. In some embodiments, probes for detecting mutations are provided as a microarray.

[0174] Any alteration in the test nucleic acid sequence, such as base substitution, insertions or deletion, could be detected using this method. The method is expected to be useful for detection of specific single base polymorphism, SNP, and the discovery of new SNPs.

[0175] Other art recognized methods of analysis for the detection of any alteration in the template nucleic acid sequence, as compared to a reference nucleic acid sequence, are suitable for use in the methods of the present invention. For example, essentially any hybridization-based method of detection of mutations is suitable for use with the fragmented polynucleotides produced by the methods of the invention.

Determination of gene expression profile

[0176] The fragmented polynucleotides, *e.g.*, labeled polynucleotide fragments, produced by the methods of the invention are particularly suitable for use in determining the levels of expression of one or more genes in a sample. As described above, fragmented polynucleotides 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 mRNAs, 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 products (for example amplification products) of the sequences, provides for determination of the gene expression profile of the sample source.

IDT Exhibit 1020.2846

[0177] Accordingly, the invention provides methods of determining gene expression profile in a sample, said method comprising: amplifying single stranded (or double stranded) product from at least one RNA sequence of interest in the sample; generating an abasic site in the amplified product; fragmenting the polynucleotide comprising the abasic site according to the methods described herein; and determining amount of fragmented polynucleotide produced from 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.

[0178] Accordingly, the invention provides of determining gene expression profile in a sample, comprising determining the amount of a polynucleotide fragment produced from a polynucleotide template as described herein, wherein the amount of a fragment is indicative of the amount of the polynucleotide template in the sample from which the template was derived, whereby a gene expression profile in the sample is determined. In one embodiment, the method comprises: (a) generating a polynucleotide fragment, *e.g.*, a labeled polynucleotide fragment, from at least one polynucleotide template in the sample using any of the methods described herein; and (b) determining amount of polynucleotide fragment produced from of each polynucleotide template, wherein each said amount is indicative of amount of each polynucleotide template in the sample, whereby the gene expression profile in the sample is determined.

[0179] It is understood that amount of fragmented polynucleotide produced (and thus the amount of product) may be determined using quantitative and/or qualitative methods. Determining amount of fragmented polynucleotides includes determining whether fragmented polynucleotides are present or absent. Thus, an expression profile can include 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.

[0180] The methods of expression profiling are useful in a wide variety of molecular diagnostics, and especially in the study of gene expression in essentially any cell (including a single cell) or cell population. A cell or cell population (e.g. a

IDT Exhibit 1020.2847

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.

Methods of preparing a subtractive hybridization probe

[0181] Fragmented polynucleotides, *e.g.*, labeled fragmented polynucleotides produced by methods of the invention are particularly suitable for use in preparation of subtractive hybridization probes. For example, two nucleic acid populations, one sense and one antisense, can be allowed to mix together with one population present in molar excess ("driver"). Sequence present in both populations will form hybrids, while sequences present in only one population remain single-stranded. Thereafter, various well-known techniques are used to separate the unhybridized molecules representing differentially expressed sequences. See, e.g., Hamson et al., U.S. Patent No. 5,589,339; Van Gelder, U.S. Patent No. 6,291,170.

Comparative hybridization

[0182] In another aspect, the invention provides methods for comparative hybridization (such as comparative genomic hybridization), said method comprising: (a) preparing a first population of polynucleotide fragments from a first template polynucleotide sample using any of the methods described herein; (b) comparing hybridization of the first population to at least one probe with hybridization of a second population of labeled polynucleotides or fragments thereof. In some embodiments, the at least one probe is a chromosomal spread. In still other embodiments, the at least one probe is provided as a microarray. In some embodiments, the first and second population comprise detectably different labels. In other embodiments, a second population of polynucleotide fragments is prepared from a second polynucleotide sample using any of the methods described herein. In some embodiments, comparing comprises determining amount of the products, whereby the amount of the first and second polynucleotide templates is quantified.

IDT Exhibit 1020.2848

[0183] In some embodiments, comparative hybridization comprises preparing a first population of labeled polynucleotide fragments according to any of the methods described herein, wherein the template from which the first population is synthesized is genomic DNA. A second population of labeled polynucleotides (to which the first population is desired to be compared) is prepared from a second genomic DNA template. The first and second populations are labeled with different labels. The hybridized first and second populations are mixed, and hybridized to an array or chromosomal spread. The different labels are detected and compared.

[0184] The following Examples are provided to illustrate, but not limit, the invention.

EXAMPLES

Example 1

Removal of Blocked 3' Termini and Progressive Degradation of Single Stranded DNA with Processive 3' to 5' Exonuclease

[0185] Amplified single stranded cDNA comprising dUTP was prepared by isothermal RNA amplification using the Ovation Biotin System (NuGEN Technologies) according to the manufacturer's instruction. Amplified cDNA products generated by amplification of a few total RNA samples (Universal Human RNA, Stratagene, 20ng each) were purified and pooled. The pooled purified cDNA was used throughout the examples below.

[0186] Pooled cDNA (5ug) was mixed with UNG (USB, 4 units) in reaction buffer containing 32mM DMED , and incubated at 37^{0}C for 30 minutes. As shown previously (U.S. Application No. 2004/0005614), treatment with UNG results in the removal of the base portion of dU residues and formation of abasic sites. DMED cleaves the backbone to generate fragmented DNA with 3'-modified termini. The blocked 3'-end can be removed by a 3' to 5'-exonucleases to generate a 3' hydroxyl group. Insofar as the aim of this procedure is to generate fragmented DNA with 3'-

IDT Exhibit 1020.2849

OH termini, it is desirable to use a non processive exonuclease so as to assure limited hydrolysis of the fragmented DNA. The nuclease activities of exonuclease 1 (Exo1), exonuclease 3 (Exo3), exonuclease T (ExoT), endonuclease D (Endo4), and APE1 (NEB) were tested. The UNG and DMED treated amplified cDNA was purified (DyeEx, Qiagen) and the purified product was incubated with the various enzymes in the respective reaction buffers (as per the manufacturer instructions). Following 30 min. incubation at 37^0C, the products were purified and the size of the treated products was analyzed electrophoretically (BioAnalyzer, Agilent). The more processive the 3' to 5' exonuclease activity of a given enzyme, the greater the expected result in reduced size and reduction in quantity of the fragmented single stranded cDNA population. As shown in Figure 3A, processive exonuclease activity of Exo1 and Exo3 resulted in reduced product recovery and reduced size of the recovered fragmented cDNA, as compared to that of the UNG and DMED fragmented cDNA substrate. The non processive 3'-exonuclease activity of endonuclease 4, exonuclease T, and APE1, on the other hand, resulted in product distribution size similar to the input fragmented cDNA (UNG and DMED treated amplified cDNA). Further evidence for the non processive 3'-deblocking activity of these enzymes was obtained by the demonstration of the ability to end label the fragmented cDNA treated with these enzymes by template-independent extension of the 3-OH termini of the cDNA by terminal transferase, as described in Example 2 below.

[0187] APE1 is a multifunctional repair enzyme comprising an endonuclease activity for the excision of abasic sites and subsequent fragmentation of the DNA template to generate fragments with 3-OH termini, on the one hand, and a 3' to 5' exonuclease activity for the removal of the blocked 3'-end of damaged DNA template. As shown in Figure 3B, treatment of amplified cDNA comprising abasic sites (amplified cDNA treated with the specific glycosylase, UNG), with APE1 (40units) led to partial fragmentation of the cDNA template, as compared to the size distribution of DMED fragmented amplified cDNA comprising abasic sites. The relative inefficiency of APE1 endonuclease activity for full fragmentation of the amplified cDNA template comprising abasic site resulted in the requirement of a large amount of the enzyme. Full fragmentation of the template DNA was achieved in

IDT Exhibit 1020.2850

reactions carried out with 500 to 1000 units of APE1 (data not shown). The 3' to 5' exonuclease activity of the enzyme is very efficient, and the non processive exonuclease activity enables the de-blocking of the 3'-end of template cDNA fragmented by the combined treatment with UNG (to generate abasic sites) and DMED (to fragment the DNA backbone while leaving an aldehyde group at the 3'termini). Use of the exonuclease activity of the enzyme for the generation of fragments with 3'-OH termini was validated by the ability to end label the fragmented and de-blocked template by terminal transferase template-independent extension with labeled nucleotide (biotin end labeling) as described in Example 2 below.

Example 2

Generation of Fragmented cDNA with 3'-OH Termini which are Suitable for Labeling by Template Independent Extension using Terminal Deoxynucleotidyl Transferase (TdT) and Labeled Nucleotide

[0188] Pooled amplified cDNA comprising dU residues was prepared as described in Example 1. Treatment of the pooled cDNA with UNG and DMED, as described in Example 1, was used to generate fragmented cDNA with blocked 3'-OH termini. A cDNA fragment with a blocked 3' terminus can be labeled with an aldehyde reactive conjugate of a desired label to yield fragmented and labeled cDNA target suitable for microarray based analysis (U.S. Application No. 2004/0005614; Dafforn et al. (2004) "*BioTechniques* 37:854-857; Kurn et al. (2005) *Clinical Chemistry* 51:1973-1981). However, extension of the fragmentation process to polymerase extension based end labeling (*e.g.*, terminal transferase template independent labeling) requires deblocking of the 3'-termini. As discussed in Example 1, various non processive 3'-to-5'exonucleases were tested for the ability to deblock the blocked 3' termini of such fragments. The validation of the ability to generate suitable substrates for end labeling was obtained by end labeling of the pooled fragmented and nuclease treated targets for end labeling with TdT (terminal deoxynucleotidyl transferase) and biotin labeled ddUTP. End labeling of the targets

IDT Exhibit 1020.2851

was assessed by hybridization of the targets to high density Human Focus GeneChip arrays (Affymetrix). Various array analysis parameters obtained with the various targets are shown in Table 1.

[0189] End labeling of the various fragmented cDNA products with TdT was carried out under the following conditions: The reactions were carried out in 1x NEB buffer #4 (50mM potassium acetate, 20mM Tris-Acetate, 10mM magnesium acetate, 1mM DTT, pH 7.9), 0.25mM CoCl2, in the presence of 0.5nmol Biotin 2',3'-dideoxy-UTP (Roche), and 40U TdT, in a total volume of 50µl. The TdT reactions were carried out at 37^{0}C for 60 min. followed by TdT inactivation (70^{0}C for 15min). The reaction products were added to a GeneChip hybridization mixture. Hybridization, wash, signal generation and array scanning were carried out as per the manufacturer instructions.

Table 1 - Array Analysis Parameters

Array	Targets	Raw Q	Scaling Factor	Back-ground	% Present	(3'/5') GAPDH
1	UNG, DMED, TdT, ddUTP	0.88	276.1	30.4	4.3	1.40
2	UNG, APE1(30U), TdT, ddUTP	0.93	44.1	32.3	32.3	1.84
3	UNG, DMED, DyeEx ExoT, TdT, ddUTP	0.99	60.6	31.6	25.5	2.20
4	UNG, DMED, DyeEx Endo4, TdT, ddUTP	1.19	24.6	34.8	38.3	2.13
5	UNG, DMED, DyeEx APE10U, TdT, ddUTP	1.04	16.7	28.2	47.2	2.25
6	UNG, DMED, DyeEx APE50U. TdT, ddUTP	1.16	17.1	33.9	45.4	1.90
7	UNG, DMED, APE10U, TdT, ddUTP	0.92	30.9	26.3	41	2.25
8	UNG, DMED100mM, pH 7.4, Mg 0.4mM, APE10U, TdT, ddUTP	1.06	8.3	31.9	56.8	1.54
9	UNG, DMED100mM, pH 7.4, Mg 2mM, APE10U, TdT, ddUTP	1.01	6.5	31.1	60.4	1.47
10	UNG, DMED34mM, pH 7.4, Mg 4mM, APE10U, TdT, ddUTP	1.07	4.0	31.7	64.0	1.50

Results

[0190] Array 1: The target was generated by Biotin 3'-end labeling of fragmented targets which were not further treated with exonuclease to unblock the 3'termini. The high Scaling Factor and low percent of genes called Present (% Present), represent poor labeling and are consistent with the inability to extend the 3'-blocked target for end labeling.

[0191] Removal of the abasic site generated by UNG by APE1, which cleaves the phosphodiester bond 5' to the abasic site sugar, generating a nick with 5' sugar phosphate (dRP) and 3' hydroxyl group at the 3'-end, enabling the biotin labeling of the fragmented cDNA targets by TdT. The endonuclease activity results in lower Scaling Factor and higher % Present results as compared to Array 1. However, this array performance is not optimal as the cDNA fragment sizes are lager than DMED cuts (as seen in Figure 3B). The results indicate that the endonuclease activity of APE1 is not sufficient to efficiently fragment the cDNA comprising abasic sites to the proper size required for efficient hybridization to the high density GeneChip arrays (Affymetrix)

[0192] Arrays 3 and 4: Purified cDNA comprising dU residues was fragmented by the action of UNG and DMED and further treated with exo T or the exonuclease activity of Endo 4, respectively. The results indicate efficient exonuclease activity of these nonprocessive exonucleases so as to unblock the 3'-termini and enable end-labeling by TdT and ddUTP.

[0193] Arrays 5 and 6: Similar labeling efficiency was enabled by unblocking of the DMED fragmented cDNA by the exonuclease activity of APE1 when present at either 10 or 50 units per reaction.

[0194] Array 7: The reaction condition was the same as for the target hybridized to array 5, except that UNG generation of abasic sites, DMED fragmentation of the cDNA at the abasic sites and unblocking of the 3'-end of the fragmented cDNA by APE1 were carried out in a single reaction mixture (30 mi. at 37^0C). The array results obtained with these reaction conditions provided a comparison for arrays 8, 9, and 10 as generation of abasic sites, fragmentation, and unblocking of 3'end occurred in a single reaction.

IDT Exhibit 1020.2853

[0195] Arrays 8, 9 and 10: Reaction conditions for the generation of fragmented cDNA with 3'-OH termini were assessed. The best performance (as per the ability to efficiently biotin-end-label the products by TdT for improved array results) was observed with conditions for target hybridized to array 10. As described for targets hybridized to arrays 8 and 9, UNG, DMED and APE1 were reacted in a single reaction mixture. The general superior performance for this group of arrays is attributed to the reaction buffer condition that favored 3'-to 5'-exonuclease activity of APE1. The buffer condition changes included lower Mg^{2+} concentration, no Na^{+}, and the reaction pH at 7.4. (Chou et al. (2003) *J. Biol. Chem.* 278(20):18289-96.

[0196] All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entireties for all purposes and to the same extent as if each individual publication, patent, or patent application were specifically and individually indicated to be so incorporated by reference.

[0197] Although the foregoing invention has been described in some detail by way of illustration and examples for purposes of clarity of understanding, it will be apparent to those skilled in the art that certain changes and modifications may be practiced without departing from the spirit and scope of the invention. Therefore, the description should not be construed as limiting the scope of the invention, which is delineated by the appended claims.

IDT Exhibit 1020.2854

CLAIMS

We claim:

1. A method for fragmenting a polynucleotide to generate fragments within a desired size range, said method comprising:

(a) chemically cleaving a phosphodiester backbone of a polynucleotide comprising an abasic site at the abasic site, whereby a polynucleotide fragment within the desired size range and comprising a blocked 3' end is generated; and

(b) contacting the polynucleotide fragment with an enzyme capable of unblocking the blocked 3' end of said fragment, whereby a polynucleotide fragment within the desired size range and comprising a 3' end hydroxyl group is generated;

wherein (a) and (b) are performed simultaneously in the same reaction mixture.

2. The method according to claim 1, wherein the phosphodiester backbone is cleaved with a polyamine to generate a polynucleotide fragment within the desired size range and with a blocked 3' end.

3. The method according to claim 2, wherein the polyamine is N, N'-dimethylethylenediamine (DMED).

4. The method according to claim 1, wherein the enzyme capable of unblocking the blocked 3' end comprises a 3' to 5' exonuclease activity.

5. The method according to claim 4, wherein the enzyme comprises non-processive exonuclease activity and does not comprise an endonuclease activity.

6. The method according to claim 4, wherein the exonuclease activity is a non-processive exonuclease activity, wherein the enzyme that comprises an

IDT Exhibit 1020.2855

exonuclease activity also comprises an endonuclease activity, and wherein contacting the polynucleotide fragment with the enzyme is under conditions in which the endonuclease activity is minimized or absent.

7. The method according to claim 4, wherein the enzyme comprising a 3' exonuclease activity is selected from the group consisting of endonuclease 4, exonuclease T, and apurinic/apyrimidinic endonuclease (APE 1).

8. The method according to claim 1, wherein the method further comprises extending the polynucleotide fragment from the 3' hydroxyl group with a template independent polymerase and a labeled nucleotide, whereby a polynucleotide fragment within the desired size range and labeled at the 3' end is generated.

9. The method according to claim 8, wherein the template independent polymerase is terminal deoxynucleotidyl transferase (TdT).

10. The method according to claim 8, wherein the labeled nucleotide is selected from the group consisting of a labeled nucleotide triphosphate (NTP), a labeled deoxynucleotide triphosphate (dNTP), and a labeled dideoxynucleotide triphosphate (ddNTP).

11. The method according to claim 8 or 10, wherein the labeled nucleotide is a biotinylated nucleotide.

12. The method according to claim 8 or 10, wherein the labeled nucleotide comprises a fluorophore.

13. The method according to claim 8, wherein a mixture of labeled and unlabeled nucleotides is used for labeling the polynucleotide fragment.

IDT Exhibit 1020.2856

14. The method according to claim 1, wherein the polynucleotide comprising an abasic site is generated by: (i) cleaving a base portion of a non-canonical nucleotide in a polynucleotide with an enzyme capable of cleaving the base portion of the non-canonical nucleotide, whereby an abasic site is generated or (ii) non-enzymatically converting a canonical or non-canonical nucleotide in a polynucleotide into an abasic site.

15. The method according to claim 14, wherein the non-canonical nucleotide is selected from the group consisting of dUTP, dITP, and 5-OH-Me-dCTP.

16. The method according to claim 14, wherein the enzyme capable of cleaving a base portion of the non-canonical nucleotide is an N-glycosylase.

17. The method according to claim 16, wherein the N-glycosylase is selected from the group consisting of Uracil N-Glycosylase (UNG), hypoxanthine-N-Glycosylase, and hydroxy-methyl cytosine-N-glycosylase.

18. The method according to claim 14, wherein the non-canonical nucleotide is dUTP and the enzyme capable of cleaving a base portion of the non-canonical nucleotide is UNG.

19. The method according to claim 14, wherein the non-canonical nucleotide is dUTP, the enzyme capable of cleaving a base portion of the non-canonical nucleotide is UNG, and the phosphodiester backbone is cleaved with DMED.

20. The method according to claim 14, wherein the polynucleotide comprising a non-canonical nucleotide is synthesized in the presence of two or more different non-canonical nucleotides, whereby a polynucleotide comprising two or more different non-canonical is synthesized

IDT Exhibit 1020.2857

21. The method according to claim 1, wherein the polynucleotide comprising an abasic site is generated by cleaving a base portion of a methylated nucleotide with an agent capable of cleaving a base portion of the methylated nucleotide to create an abasic site, whereby an abasic site is generated.

22. The method according to claim 1, further comprising cleaving a base portion of a methylated nucleotide in a polynucleotide with an agent capable of cleaving a base portion of the methylated nucleotide to create an abasic site, whereby the polynucleotide comprising an abasic site is generated.

23. The method according to claim 1, wherein the polynucleotide comprising an abasic site is generated by cleaving a base portion of a canonical nucleotide with an agent capable of cleaving a base portion of the canonical nucleotide to create an abasic site, whereby an abasic site is generated.

24. The method according to claim 1, further comprising cleaving a base portion of a canonical nucleotide in a polynucleotide with an agent capable of cleaving a base portion of the canonical nucleotide to create an abasic site, whereby the polynucleotide comprising an abasic site is generated.

25. The method according to claim 23 or 24, wherein the canonical nucleotide is cytosine and the agent capable of cleaving a base portion of the canonical nucleotide comprises cytosine deaminase in conjunction with UNG.

26. The method according to claim 1, wherein the polynucleotide comprising an abasic site is synthesized from a polynucleotide template comprising DNA or RNA.

27. The method according to claim 26, wherein the polynucleotide template is selected from the group consisting of mRNA, cDNA, and genomic DNA.

IDT Exhibit 1020.2858

28. The method according to claim 1, wherein the polynucleotide comprising an abasic site is single stranded or double stranded.

29. The method according to claim 1, wherein the polynucleotide comprising an abasic site is synthesized by an amplification method selected from the group consisting of polymerase chain reaction (PCR), strand displacement amplification (SDA), multiple displacement amplification (MDA), rolling circle amplification (RCA), single primer isothermal amplification (SPIA), and Ribo-SPIA.

30. The method according to claim 1, wherein the polynucleotide comprising an abasic site is synthesized by a method selected from the group consisting of reverse transcription, primer extension, limited primer extension, replication, and nick translation.

31. A method for fragmenting and labeling a polynucleotide to generate labeled fragments within a desired size range, said method comprising:

(a) synthesizing a polynucleotide from a polynucleotide template in the presence of a non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated;

(b) cleaving a base portion of the non-canonical nucleotide from the synthesized polynucleotide with an enzyme capable of cleaving the base portion of the non-canonical nucleotide, whereby an abasic site is generated; cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site at the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; contacting the polynucleotide fragment with an enzyme capable of unblocking the blocked 3' end of said fragment, whereby a polynucleotide fragment comprising a 3' hydroxyl group is generated; and contacting the polynucleotide fragment comprising a 3' hydroxyl group with an enzyme capable of extending the polynucleotide fragment from the 3' end and a labeled nucleotide, whereby a labeled polynucleotide fragment within the desired size range is generated.

IDT Exhibit 1020.2859

32. The method according to claim 31, wherein (i) cleaving the base portion of the non-canonical nucleotide from the synthesized polynucleotide and (ii) cleaving the phosphodiester backbone of the polynucleotide are performed simultaneously in the same reaction mixture.

33. The method according to claim 31, wherein (i) cleaving the phosphodiester backbone of the polynucleotide and (ii) contacting the polynucleotide fragment with an enzyme capable of unblocking the blocked 3' end of the fragment are performed simultaneously in the same reaction mixture.

34. The method according to claim 31, wherein the method comprises synthesizing the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range.

35. The method according to claim 31, wherein the polynucleotide comprising an abasic site is synthesized by an amplification method selected from the group consisting of polymerase chain reaction (PCR), strand displacement amplification (SDA), multiple displacement amplification (MDA), rolling circle amplification (RCA), single primer isothermal amplification (SPIA), and Ribo-SPIA.

36. The method according to claim 31, wherein the polynucleotide comprising an abasic site is synthesized by a method selected from the group consisting of reverse transcription, primer extension, limited primer extension, replication, and nick translation.

37. A method for fragmenting and labeling a polynucleotide to generate labeled fragments within a desired size range, said method comprising:

IDT Exhibit 1020.2860

(a) synthesizing a polynucleotide from a polynucleotide template in the presence of a non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated;

(b) cleaving a base portion of the non-canonical nucleotide from the synthesized polynucleotide with an enzyme capable of cleaving the base portion of the non-canonical nucleotide, whereby an abasic site is generated;

(c) cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site at the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated;

(d) contacting the polynucleotide fragment with an enzyme capable of unblocking the blocked 3' end of said fragment, whereby a polynucleotide fragment comprising a 3' hydroxyl group is generated; and

(e) contacting the polynucleotide fragment comprising a 3' hydroxyl group with an enzyme capable of extending the polynucleotide fragment from the 3' end and a labeled nucleotide, whereby a labeled polynucleotide fragment within the desired size range is generated;

wherein (b), (c) and (d) are performed simultaneously in the same reaction mixture.

38. The method according to claim 37, wherein the method comprises synthesizing the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range.

39. The method according to claim 37, wherein the polynucleotide comprising a non-canonical nucleotide is synthesized by an amplification method selected from the group consisting of polymerase chain reaction (PCR), strand displacement amplification (SDA), multiple displacement amplification (MDA), rolling circle amplification (RCA), single primer isothermal amplification (SPIA), and Ribo-SPIA.

IDT Exhibit 1020.2861

40. The method according to claim 37, wherein the polynucleotide comprising a non-canonical nucleotide is synthesized by a method selected from the group consisting of reverse transcription, primer extension, limited primer extension, replication, and nick translation.

41. A method for fragmenting a polynucleotide to generate fragments within a desired size range, said method comprising:

(a) incubating a reaction mixture, said reaction mixture comprising: (i) a polynucleotide template; and (ii) a non-canonical nucleotide; wherein the incubation is under conditions that permit synthesis of a polynucleotide comprising the non-canonical nucleotide, whereby a polynucleotide comprising the non-canonical nucleotide is generated;

(b) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide comprising the non-canonical nucleotide; and (ii) an enzyme capable of cleaving a base portion of the non-canonical nucleotide, wherein the incubation is under conditions that permit cleavage of the base portion of the non-canonical nucleotide, whereby a polynucleotide comprising an abasic site is generated;

(c) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide comprising the abasic site; and (ii) an agent capable of chemically cleaving the phosphodiester backbone of the polynucleotide comprising the abasic site at the abasic site, wherein the incubation is under conditions that permit cleavage of the phosphodiester backbone of the polynucleotide at the abasic site, whereby a polynucleotide fragment comprising a blocked 3' end is generated; and

(d) incubating a reaction mixture, said reaction mixture comprising: (i) the fragment of the polynucleotide comprising a blocked 3' end; and (ii) an enzyme capable of unblocking the blocked 3' end, whereby a polynucleotide fragment within the desired size range and comprising a 3' hydroxyl group is generated;

IDT Exhibit 1020.2862

wherein (b), (c) and (d) are performed simultaneously in the same reaction mixture.

42. The method according to claim 41, further comprising:

(e) incubating a reaction mixture, said reaction mixture comprising: (i) the polynucleotide fragments within the desired size range comprising a 3' hydroxyl group; and (ii) an agent capable of extending the fragment from the 3' hydroxyl group; and (iii) a labeled nucleotide, wherein the incubation is under conditions that permit extension of the polynucleotide fragment from the 3' hydroxyl group, whereby a labeled polynucleotide fragment within the desired size range is generated.

43. The method according to claim 42, wherein the agent capable of extending the polynucleotide fragments within the desired size range from the 3' hydroxyl group is TdT, wherein the polynucleotide fragments within the desired size range are labeled at the 3' hydroxyl group with the labeled nucleotide.

44. The method according to claim 41, wherein the incubation is under conditions that permit synthesis of the polynucleotide from the polynucleotide template in the presence of all four canonical nucleotides and a non-canonical nucleotide, wherein the non-canonical nucleotide is provided at a ratio suitable for generating fragments within the desired size range.

45. A method for characterizing a polynucleotide template of interest, comprising analyzing a labeled polynucleotide fragment within a desired size range produced by the method of claim 1, 31, or 37.

46. The method according to claim 45, wherein analyzing the labeled polynucleotide fragment within the desired size range comprises determining amount of said products, whereby the amount of the polynucleotide template present in a sample is quantified.

IDT Exhibit 1020.2863

47. The method according to claim 45, wherein analyzing the labeled polynucleotide fragment within the desired size range comprises contacting the labeled polynucleotide fragment with at least one probe.

48. The method according to claim 47, wherein the at least one probe is provided as a microarray.

49. A method for determining gene expression profile in a sample, said method comprising determining the amount of a labeled polynucleotide fragment or fragments within a desired size range produced by the method of claim 1, 31, or 37, wherein the amount of said labeled polynucleotide fragment or fragments are indicative of the amount of a polynucleotide template in said sample from which the labeled polynucleotide fragment was generated, whereby a gene expression profile is determined.

50. The method according to claim 49, wherein the polynucleotide template is RNA or mRNA.

51. The method according to claim 49, wherein the amounts of a plurality of polynucleotide fragments within the desired size range derived from a plurality of polynucleotide templates in a sample are determined.

52. A method for nucleic acid hybridization, comprising hybridizing labeled polynucleotide fragments within a desired size range with at least one probe, wherein the labeled polynucleotide fragments within the desired size range are generated according to the method of claim 1, 31, or 37.

53. A method for comparative hybridization, comprising comparing hybridization of a first population of labeled polynucleotide fragments within a desired size range prepared from a first polynucleotide sample by the method of claim

IDT Exhibit 1020.2864

1, 31, or 37 to at least one probe with hybridization of a second population of labeled polynucleotide to the at least one probe.

54. The method according to claim 53, wherein the first population and second population comprise detectably different labels.

55. The method according to claim 53, wherein the second population of labeled polynucleotides is a population of labeled polynucleotide fragments within the desired size range prepared from a second polynucleotide sample by the method of 1, 31, or 37, and wherein comparing comprises determining amount of said labeled polynucleotide fragments, whereby the amount of the first and second polynucleotide templates is quantified.

56. The method according to claim 53, wherein the first polynucleotide sample comprises genomic DNA, the second polynucleotide sample comprises genomic DNA, or both the first and the send polynucleotide samples comprise genomic DNA.

57. A method for detecting presence or absence of a mutation in a template, comprising analyzing a labeled polynucleotide fragments within a desired size range prepared by the method of claim 1, 31, or 37, whereby presence of absence of a mutation is detected, wherein analyzing comprises comparison of the labeled polynucleotide fragment to a polynucleotide prepared from a reference polynucleotide.

58. The method according to claim 57, wherein the mutation is selected from the group consisting of a base substitution, a base insertion, a base deletion, and a single nucleotide polymorphism.

59. A composition comprising: (a) an agent capable of cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide; (b) an agent

IDT Exhibit 1020.2865

capable of cleaving a phosphodiester backbone at an abasic site to produce a polynucleotide fragment with a blocked 3' end within a desired size range; and (c) an enzyme capable of unblocking a blocked 3' end to generate a polynucleotide comprising a 3' hydroxyl group.

60. The composition according to claim 59, wherein (a) is an N-glycosylase, (b) is a polyamine, and (c) is an enzyme comprising a 3' exonuclease activity.

61. The composition according to claim 60, wherein (a) is UNG, (b) is DMED, and (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1.

62. A kit comprising: (a) an agent capable of cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide; (b) an agent capable of cleaving a phosphodiester backbone at an abasic site to produce a polynucleotide fragment with a blocked 3' end within a desired size range; and (c) an enzyme capable of unblocking a blocked 3' end to generate a polynucleotide comprising a 3' hydroxyl group.

63. The kit according to claim 62, further comprising: (d) an agent capable of labeling a 3' hydroxyl group of a polynucleotide.

64. The kit according to claim 63, wherein (a) is an N-glycosylase, (b) is a polyamine, (c) is an enzyme comprising a 3' exonuclease activity; and (d) is a template independent polymerase.

65. The kit according to claim 64, wherein (a) is UNG, (b) is DMED, (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1; and (d) is TdT.

IDT Exhibit 1020.2866

66. The kit according to claim 62, further comprising a non-canonical nucleotide and an enzyme capable of synthesizing a polynucleotide comprising the non-canonical nucleotide.

67. The kit according to claim 66, wherein the non-canonical nucleotide is dUTP and the agent capable of cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide is UNG.

68. The kit according to claim 63, further comprising: (e) a labeled nucleotide.

69. The kit according to claim 68, wherein (a) is an N-glycosylase, (b) is a polyamine, (c) is an enzyme comprising a 3' exonuclease activity; (d) is a template independent polymerase; and (e) is a biotinylated nucleotide.

70. The kit according to claim 68, wherein (e) is selected from the group consisting of a labeled nucleotide triphosphate (NTP), a labeled deoxynucleotide triphosphate (dNTP), and a labeled dideoxynucleotide triphosphate (ddNTP).

71. The kit according to claim 70, wherein the labeled nucleotide is a biotinylated nucleotide.

72. The kit according to claim 68, wherein (a) is UNG, (b) is DMED, (c) is selected from the group consisting of endonuclease 4, exonuclease T, and APE 1; (d) is TdT, and (e) is selected from the group consisting of biotin 2',3'-dideoxy-UTP and biotin 2',3'-dideoxy-CTP.

73. The kit according to claim 62, further comprising a template dependent DNA polymerase; a composite primer, wherein the composite primer comprises an RNA portion and a 3' DNA portion; and an agent capable of cleaving RNA from an RNA-DNA hybrid.

IDT Exhibit 1020.2867

74. The kit according to claim 62, further comprising instructions for use in a method for fragmenting a polynucleotide within the desired size range, said method comprising:

(a) cleaving a base portion of a nucleotide to generate an abasic site in a polynucleotide

(b) chemically cleaving a phosphodiester backbone of a polynucleotide comprising an abasic site at the abasic site, whereby polynucleotide fragments within the desired size range comprising a blocked 3' end are generated; and

(c) contacting the polynucleotide fragments with an enzyme capable of unblocking the blocked 3' end of said fragments, whereby a polynucleotide fragments within the desired size range comprising a 3' end hydroxyl group are generated.

wherein (a) and (b) and (c) are carried out simultaneously in the same reaction mixture

IDT Exhibit 1020.2868

FIGURE 1

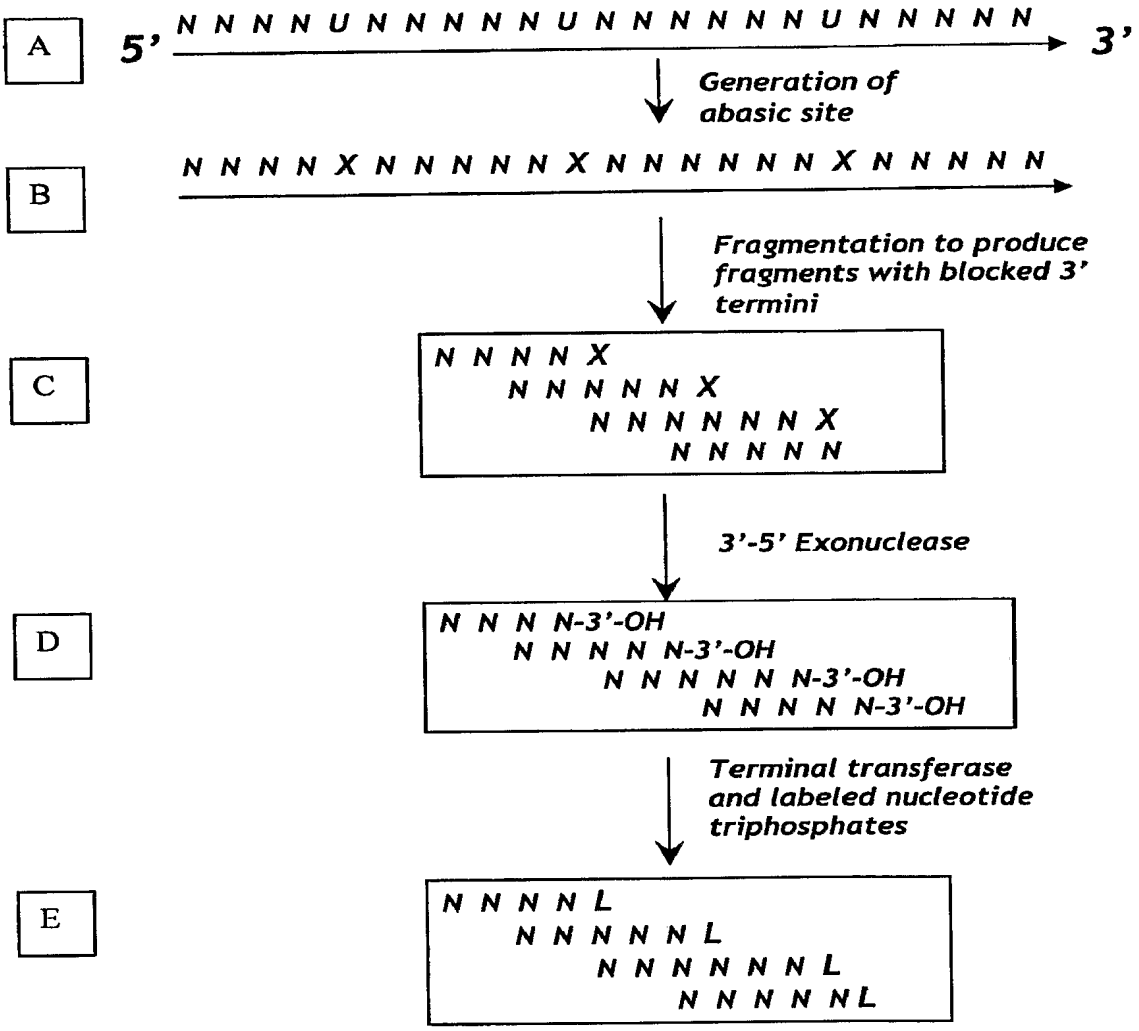

A
5' N N N N U N N N N N U N N N N N N U N N N N N 3'
Generation of abasic site
B
N N N N X N N N N N X N N N N N N X N N N N N
Fragmentation to produce fragments with blocked 3' termini
C
N N N N X
N N N N N X
N N N N N N X
N N N N N
3'-5' Exonuclease
D
N N N N-3'-OH
N N N N N-3'-OH
N N N N N N-3'-OH
N N N N N-3'-OH
Terminal transferase and labeled nucleotide triphosphates
E
N N N N L
N N N N N L
N N N N N N L
N N N N N L

FIGURE 2

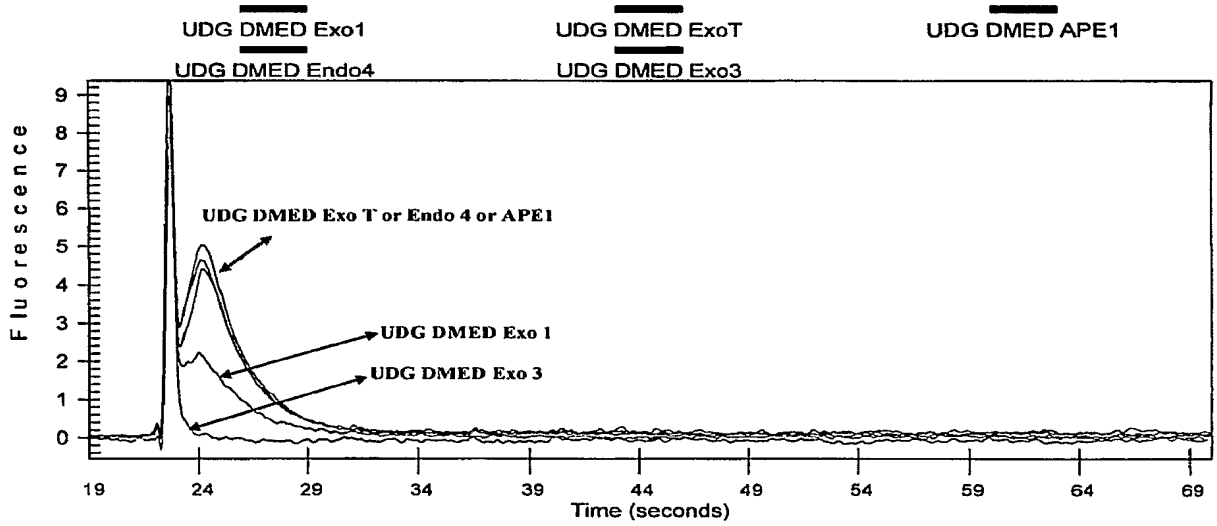
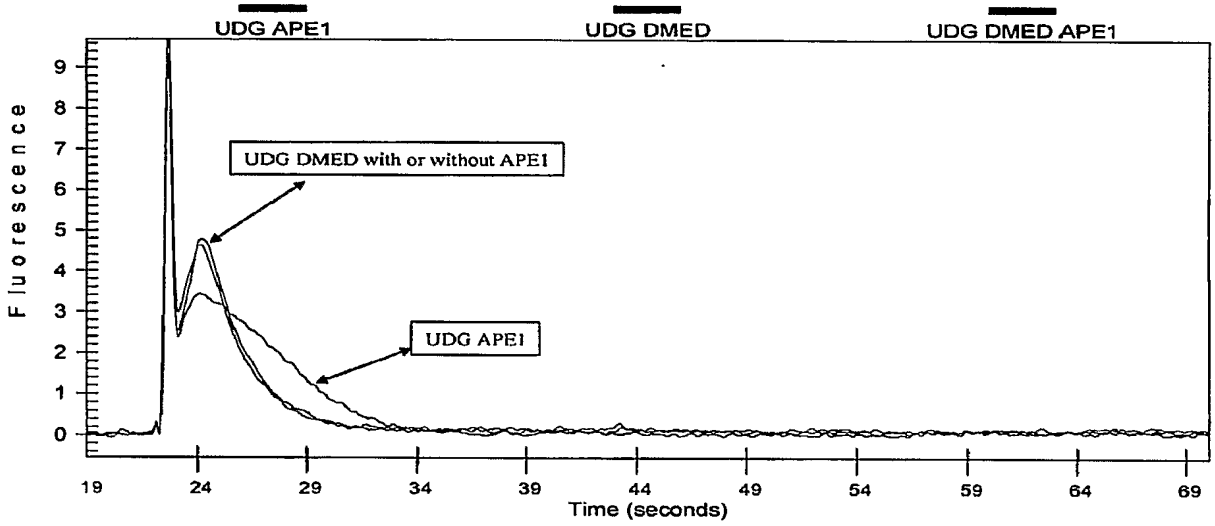

UDG DMED Exo1
UDG DMED Endo4
UDG DMED ExoT
UDG DMED Exo3
UDG DMED APE1
Fluorescence
0
1
2
3
4
5
6
7
8
9
UDG DMED Exo T or Endo 4 or APE1
UDG DMED Exo 1
UDG DMED Exo 3
19
24
29
34
39
44
49
54
59
64
69
Time (seconds)

Figure 3


UDG APE1
UDG DMED
UDG DMED APE1
UDG DMED with or without APE1
UDG APE1
Fluorescence
0
1
2
3
4
5
6
7
8
9
19
24
29
34
39
44
49
54
59
64
69
Time (seconds)

Espacenet

Bibliographic data: WO9416090 (A1) — 1994-07-21

METHOD FOR GENERATING SINGLE-STRANDED DNA MOLECULES

Inventor(s): NIKIFOROV THEO [US]; KNAPP MICHAEL R [US] ± (NIKIFOROV, THEO, ; KNAPP, MICHAEL, R)

Applicant(s): MOLECULAR TOOL INC [US]; NIKIFOROV THEO [US]; KNAPP MICHAEL R [US] ± (MOLECULAR TOOL, INC, ; NIKIFOROV, THEO, ; KNAPP, MICHAEL, R)

Classification: - international: ***C07H21/04; C12N15/09; C12N15/10; C12Q1/68; G01N33/53; G01N33/566; G01N33/58;*** (IPC1-7): C12P19/34
- cooperative: **C12N15/10**; **C12Q1/6848**; **C12Q1/6853**; **C12Q1/6858**; **C12Q1/686**; Y10S435/81 more

Application number: WO1994US00771 19940118 Global Dossier

Priority number (s): US19930005061 19930115 ; US19930155746 19931123

Also published as: AT239090 (T) AU6126294 (A) AU674211 (B2) CA2153898 (A1) CA2153898 (C) more

Abstract of WO9416090 (A1)

A method for generating single-stranded nucleic acid molecules. The molecules contain nuclease resistant modified nucleotides, such that they are resistant to 5' -> 3' exonucleases.

IDT Exhibit 1020.2872

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12P 19/34 — A1

(11) International Publication Number: WO 94/16090

(43) International Publication Date: 21 July 1994 (21.07.94)

(21) International Application Number: PCT/US94/00771

(22) International Filing Date: 18 January 1994 (18.01.94)

(30) Priority Data:

08/005,061	15 January 1993 (15.01.93)	US
08/155,746	23 November 1993 (23.11.93)	US

(71) Applicant *(for all designated States except US)*: MOLECULAR TOOL, INC. [US/US]; Alpha Center, Hopkins Bayview Research Campus, 5210 Eastern Avenue, Baltimore, MD 21224 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NIKIFOROV, Theo [BG/US]; 8 Charles Plaza, Apartment 605 NT, Baltimore, MD 21201 (US). KNAPP, Michael, R. [US/US]; 2630 North Calvert, Baltimore, MD 21218 (US).

(74) Agents: AUERBACH, Jeffrey, I. et al.; Howrey & Simon, 1299 Pennsylvania Avenue, NW, Washington, DC 20004 (US).

(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, US, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD FOR GENERATING SINGLE-STRANDED DNA MOLECULES

(57) Abstract

A method for generating single-stranded nucleic acid molecules. The molecules contain nuclease resistant modified nucleotides, such that they are resistant to 5' ← 3' exonucleases.

IDT Exhibit 1020.2873

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	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

5

TITLE OF THE INVENTION:

METHOD FOR GENERATING SINGLE-STRANDED DNA MOLECULES

10

FIELD OF THE INVENTION:

The invention relates to a method for generating single-
15 stranded DNA molecules. More specifically, it relates to the use of modified nucleotides and a 5'→3' exonuclease to produce single-stranded nucleic acid molecules following primer-mediated extension.

20 ## CROSS-REFERENCE TO RELATED APPLICATIONS:

This application is a continuation-in-part of U.S. Patent Application Serial No. 08/ 005,061 (filed January 15, 1993, and now pending), herein incorporated by reference.

25

BACKGROUND OF THE INVENTION:

The analysis of the structure, organization and sequence of nucleic acid molecules is of profound importance to the
30 prediction, diagnosis and treatment of human and animal disease, in forensics, in epidemiology and public health, and in the elucidation of the factors that control gene expression and development.

Three areas of particular importance involve the
35 development of nucleic acid molecules that are capable of hybridizing to a desired sequence, the generation of nucleic acid molecules that are single-stranded, and the determination of the nucleotide sequence of a nucleic acid molecule.

I. Nucleic Acid Hybridization

The capacity of a nucleic acid "probe" molecule to hybridize (i.e. base pair) to a complementary nucleic acid "target" molecule forms the cornerstone for a wide array of diagnostic and therapeutic procedures.

Hybridization is used to detect and identify causal agents of infectious disease, to provide information on paternity and lineage, to predict the likelihood that an individual will suffer from a genetic disease, or to identify tissue samples. The diagnostic value of such procedures hinges on their sensitivity. Sensitivity can be increased through the use of probes that are detectably labelled. The most common label involves the use of radioisotopic labels (Falkow et al. (U.S. Patent 4,358,535); Berninger (U.S. Patent 4,446,237). Methods of labeling and performing such hybridization reactions are disclosed by, for example, Sambrook, J. et al. (In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)), and Haymes, B.D., et al. (In: Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, DC (1985)) which references are included herein by reference.

The sensitivity of nucleic acid hybridization detection assays may also be increased by altering the manner in which detection is reported or signaled to the observer. Thus, for example, assay sensitivity can be increased through the use of detectably labeled reagents. A wide variety of such labels have been used for this purpose. Kourilsky et al. (U.S. Patent 4,581,333) describe the use of enzyme labels to increase sensitivity in a detection assay. Fluorescent labels (Albarella et al., EP 144914), chemical labels (Sheldon III et al., U.S. Patent 4,582,789; Albarella et al., U.S. Patent 4,563,417), modified bases (Miyoshi et al., EP 119448), etc. have also been used in an effort to improve the efficiency with which hybridization can be observed.

Hybridization assays employing synthetically or enzymatically made single-stranded nucleic acid probes can be performed in solution (Berk, A.J. et al., Cell 12:721-732 (1977); Hood, L.E. et al., In: Molecular Biology of Eukaryotic Cells: A Problems Approach, Menlow Park, CA: Benjamin-Cummings (1975); Wetmer, J.G., Ann. Rev. Biophys. Bioeng. 5:337-361 (1976); Itakura, K. et al., Ann. Rev. Biochem. 53:323-356 (1984)) or in conjunction with gel electrophoresis or nucleic acid-binding membrane blotting methods. Such methods also allow the detection of nucleic acid molecules with sequences that are complementary to all or part of the probe (Alwine, J.C. et al., Proc. Natl. Acad. Sci. (U.S.A.) 74:5350-5354 (1977); Southern, E.M., J. Molec. Biol. 98:503-517 (1975)); Berk, A.J. et al., Cell 12:721-732 (1977); Itakura, K. et al., Ann. Rev. Biochem. 53:323-356 (1984); Ruddle, F.H., Nature 294:115-119 (1981); White, R. et al., Sci. Amer. 258:40-48 (1988); McGinnis, W. et al., Cell 37:403-408 (1984)). Single-stranded nucleic acid probes can also be used in situ to locate specific nucleic acid sequences in a method termed "in situ hybridization" (Abelson, J. et al., Science 209:1317-1438 (1980); Gilbert, W. et al., Sci. Amer. 242:74-94 (1980)).

Hybridization assays may also be performed using affinity chromatography methods. In this method, one single-stranded nucleic acid molecule, usually an oligonucleotide, is immobilized to a solid support matrix and used as a probe to hybridize a second complementary single-stranded nucleic acid molecule. Efficient detection or recovery of a single desired nucleic molecule is enhanced when the two complementary single-stranded sequences are present in non-limiting concentrations and each are in substantially pure form. For example, single-stranded oligonucleotides of high purity have been isolated from solution by affinity chromatography using immobilized (i.e bound to solid support matrix) oligonucleotides complementary to the oligomers in solution, as discussed for example, in Gilham et al. (J. Amer. Chem. Soc. 86:4982 (1964)) and Kremsky et al. (Nucl. Acids Res. 15:3131-3139 (1987)).

The capacity of DNA molecules to hybridize to complementary mRNA molecules, and thereby attenuate the translation of specific proteins forms one basis for the therapeutic application of hybridization technology. Such "antisense" technology has significant potential in anti-viral and anti-cancer therapy. Antisense technology is discussed in European Patent Application Publication Nos. 263,740; 335,451; and 329,882, and in PCT Publication No. WO90/00624, all of which references are incorporated herein by reference.

Hybridization technology is also exploited to aid in the recovery of RNA. In the case of eukaryotic mRNA, this has been accomplished using affinity matrix chromatography columns having polydeoxythymidine oligonucleotides bound to a solid support matrix comprised of cellulose (i.e. Oligo (dT)-cellulose columns). Such oligonucleotides are capable of hybridizing to the polyadenine mRNA "tails" normally found on the 3' end of all eukaryotic mRNA molecules (Gilham, P.T., J. Amer. Chem. Soc. 86:4982 (1971)). Such methods of isolating single-stranded nucleic acid molecules require large quantities of starting material.

II. The Amplification of Nucleic Acid Molecules

The ability to detect the presence of a desired target nucleic acid molecule in a sample is often limited by the concentration of the molecule in either its double-stranded and single-stranded forms. In many such situations, the concentration of the target can be amplified through the use of either in vivo or in vitro based amplification systems.

In vivo based amplification systems include amplification of a target nucleotide molecule through its propagation (i.e. replication and amplification) in cloning or expression vectors. Cloning and expression vectors are disclosed, for example, in Sambrook, J. et al. (In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)).

Commonly employed in vitro based amplification systems include enzymatic methods using DNA dependent or RNA dependent DNA or RNA polymerases. The most widely used method of nucleic acid amplification, the "polymerase chain reaction" ("PCR"), involves template-dependent extension using thermally stable DNA polymerase (Mullis, K. et al., Cold Spring Harbor Symp. Quant. Biol. 51:263-273 (1986); Erlich H. et al., EP 50,424; EP 84,796, EP 258,017, EP 237,362; Mullis, K., EP 201,184; Mullis K. et al., US 4,683,202; Erlich, H., US 4,582,788; and Saiki, R. et al., US 4,683,194), which references are incorporated herein by reference. PCR achieves the amplification of a specific nucleic acid sequence using two oligonucleotide primers complementary to regions of the sequence to be amplified. Extension products incorporating the primers then become templates for subsequent replication steps. Reviews of the polymerase chain reaction are provided by Mullis, K.B. (Cold Spring Harbor Symp. Quant. Biol. 51:263-273 (1986)); Saiki, R.K. et al. (Bio/Technology 3:1008-1012 (1985)); and Mullis, K.B. et al. (Meth. Enzymol. 155:335-350 (1987), which references are incorporated herein by reference).

Other nucleic acid amplification procedures include transcription-based amplification systems (Kwoh D et al., Proc. Natl. Acad. Sci. (U.S.A.) 86:1173 (1989); Gingeras T.R. et al., PCT appl. WO 88/10315 (priority: US Patent applications serial nos. 064,141 and 202,978); Miller, H.I. et al., PCT appl. WO 89/06700 (priority: US Patent application serial no. 146,462); Davey, C. et al. (European Patent Application Publication no. 329,822)) and ligation-based amplification systems (Wu, D.Y. et al., Genomics 4:560 (1989)).

Although amplification technologies can be used to achieve the rapid and extensive amplification of a polynucleotide molecule, such methods generally result in the production of double-stranded DNA. Thus, the methods, in general fail to provide a selective means for amplifying and isolating a single-strand of a double-stranded target molecule.

Single-stranded DNA molecules may be produced using the single-stranded DNA bacteriophage M13 (Messing, J. et al., Meth. Enzymol. 101:20 (1983); see also, Sambrook, J. et al. (In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). Use of M13 to produce single-stranded DNA, however, requires cloning and other time consuming manipulations, and thus M13 is primarily used for DNA sequencing. In general, the method entails cloning a target DNA molecule into the single-stranded DNA chain of M13. Once introduced into a host bacterium, the recombinant M13 vector directs the formation and extrusion of bacteriophage particles that contain single-stranded DNA.

There are several major disadvantages in using M13 to produce single-stranded DNA molecules. First, the method requires the cloning of the target DNA, and the ultimate isolation and purification of the mature bacteriophage particles. Hence, the method is quite time consuming. More significantly, the isolated target DNA is unavoidably attached to the M13 viral DNA sequences. A further disadvantage to the M13 system is due to the instability of DNA target molecules larger than 1000 nucleotides which often results in loss of the desired recombinant M13 phage. For the forgoing reasons, M13 is not used to generate single-stranded DNA for most applications other than DNA sequencing.

Several methods are currently used for generating single-stranded DNA molecules. Gyllensten, U. et al. (Proc. Natl. Acad. Sci. (U.S.A.) 85:7652-7656 (1988) and Mihovilovic, M. et al. (BioTechniques 7:14 (1989)) describe a method involving a modification of the standard "PCR" method which is normally used for amplifying double-stranded DNA molecules. This modified PCR procedure, termed "asymmetric PCR," employs amplification primers present in different molar concentrations. When asymmetric primer concentrations are used in the "asymmetric PCR" technique, the primer in limiting concentration becomes exhausted after the first 10 to 15

amplification cycles. Continued cycling generates single-stranded DNA originating from the non-limiting primer.

There are, unfortunately, several disadvantages to employing "asymmetric PCR" to obtain single-stranded DNA. Single-stranded DNA amplification only occurs linearly with cycle number in contrast to exponential DNA amplification when using the standard PCR method. Additionally, to optimize the yield of single-stranded product, it is often necessary to carry out several separate amplification reactions that contain varying concentrations and ratios of the priming oligonucleotides. The advantages to generating single-stranded DNA in combination with an exponentially producing amplification reaction, such as the standard non-asymmetric PCR, are thus apparent.

Higuchi, R.G. et al. (Nucl. Acids Res. 17:5865 (1985)) exemplifies an additional method currently used for generating single-stranded amplification products. The method entails phosphorylating the 5' terminus of one strand of a double-stranded amplification product, and then permitting a 5'→3' exonuclease (such as λ exonuclease) to preferentially degrade the phosphorylated strand. The method thus has several drawbacks. The efficiency of the method depends both on the extent and specificity of the phosphorylation reaction, and on the degree of preference exhibited by the exonuclease.

5'→3' exonucleases have been used to prepare single-stranded DNA fragments from full-length double-stranded DNA molecules. When such a full-length molecule is incubated in the presence of a 5'→3' exonuclease, degradation occurs from the 5' terminus of each strand. The degradation of the first strand continues until the exonuclease acting on that strand reaches the region in which the second strand has been degraded by the exonuclease acting on that strand. Thus, the method produces two "half-length," non-complementary molecules from a full-length duplex DNA molecule.

Additionally, other methods have exploited the nuclease resistant properties of phosphorothioate derivatives for

generating single-stranded DNA molecules. Benkovic et al. (U.S. Patent No. 4,521,509; June 4, 1985)) used the restriction endonuclease and the 3'→5' exonuclease resistant properties of phosphorothioate-containing nucleic acid sequences to generate single-stranded DNA molecules. This method employs the use of a restriction endonuclease to form a double-stranded molecule having a single recessed 3' hydroxyl terminus. Phosphorothioate nucleotides are used to modify this terminus, thereby producing a strand that is resistant to exonuclease attack, which permits the generation of a single-stranded product. This method is limiting since the target DNA sequence must contain two desired restriction endonuclease sites: the first must create a recessed 3'-OH terminus, and be at one end of the target molecule and the second must create a recessed 5' terminus, and be at the second end. A further limitation to this method is that production of a desired single-stranded DNA product requires a high concentration of target double-stranded DNA molecules.

Sayers, J.R. et al. (Nucl. Acids Res. 16:791-802 (1988)) exemplifies a method that uses the restriction endonuclease-resistant properties of phosphorothioate-containing DNA to produce single-stranded DNA. In the method, a primer is permitted to hybridize to a circular target molecule. Primer extension then occurs in the presence of phosphorothioate nucleotides, such that the nucleotide derivatives are incorporated into the extension product. The termini of the extension product are then ligated, to form a double-stranded circular molecule. The presence of the phosphorothioate residues in the circularized extension product renders that strand resistant to restriction endonucleases. Thus, upon incubation with such endonucleases, the target strand is cleaved. Such cleavage generates termini that can then be attacked by exonucleases. Significantly, the exonuclease resistance of the phosphorothioate-containing strand cannot be evaluated, since that strand, being circular, is not a substate for an exonuclease.

Phosphorothioate-containing oligonucleotides have been found to protect oligonucleotide primers from degradation by the 5'→3' "mismatch" exonuclease activity of polymerase I (Ott, J. et al., Biochem 26:8237-8241 (1987). The method of Ott et al., since it employs a polymerase, is incapable of producing single-stranded DNA.

Although the method is suitable for site-directed mutagenesis, it is limited by its dependency on using the cumbersome and limited bacteriophage M13 system described above. Additionally, the Sayers et al. method requires the presence of a restriction endonuclease cleavage site in the target molecule.

In sum, the ability to manipulate and exploit nucleic acid molecules often requires the isolation of a single-stranded molecular species. Present methods of nucleic acid amplification typically lead to the formation of double-stranded species, and thus require additional processing steps in order to obtain purified preparations of single-stranded molecules.

III. The Sequencing of Nucleic Acid Molecules

Initial attempts to determine the sequence of a DNA molecule employed extensions of techniques that had been developed to permit the sequencing of RNA molecules (Sanger, F., J. Molec. Biol. 13:373 (1965); Brownlee, G.G. et al., J. Molec. Biol. 34:379 (1968)). Such early methods involved the specific cleavage of DNA into smaller fragments by (1) enzymatic digestion (Robertson, H.D. et al., Nature New Biol. 241:38 (1973); Ziff, E.B. et al., Nature New Biol. 241:34 (1973)); (2) nearest neighbor analysis (Wu, R. et al., J. Molec. Biol. 57:491 (1971)), and (3) the "Wanderings Spot" method (Sanger, F., Proc. Natl. Acad. Sci. (U.S.A.) 70:1209 (1973)).

More recent advances have led to the development of two highly utilized methods for elucidating the sequence of a DNA molecule: the "Dideoxy-Mediated Chain Termination Method," also known as the "Sanger Method" (Sanger, F. et al., J. Molec. Biol.

94:441 (1975)) and the "Maxam-Gilbert Chemical Degradation Method" (Maxam, A.M. et al., Proc. Natl. Acad. Sci. (U.S.A.) 74:560 (1977), both references herein incorporated by reference). Methods for sequencing DNA using either the dideoxy-mediated method or the Maxam-Gilbert method are widely known to those of ordinary skill in the art. Such methods are, for example, disclosed in Maniatis, T., et al., Molecular Cloning, a Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), and in Zyskind, J.W. et al., Recombinant DNA Laboratory Manual, Academic Press, Inc., New York (1988), both herein incorporated by reference.

In the dideoxy-mediated method, the sequence of a target molecule is determined using four separate primer extension reactions, each of which are conducted using a polymerase, an oligonucleotide primer, and the four nucleotide triphosphates needed to polymerize DNA. Each of the reactions is carried out in the additional presence of a 2',3' dideoxy derivative of either an A, T, C, or G nucleoside triphosphate. Such derivatives differ from conventional nucleotide triphosphates in that they lack a hydroxyl residue at the 3' position of deoxyribose. Thus, although they can be incorporated into the newly synthesized primer extension, such incorporation results in the termination of the extension reaction. The net result of each of the four reactions is the production of a set of nested oligonucleotides each of which is terminated by the particular dideoxy derivative used in the reaction. Such reaction products can be readily analyzed to yield the sequence of the target molecule.

The Maxam-Gilbert method of DNA sequencing is a degradative method. In this procedure, a fragment of DNA is labeled at one end and partially cleaved in four separate chemical reactions, each of which is specific for cleaving the DNA molecule at a particular base (G or C) at a particular type of base (A/G, C/T, or A>C). As in the above-described dideoxy method, the effect of such reactions is to create a set of nested molecules whose lengths are determined by the locations of a particular base along the length of the DNA molecule being

sequenced. The nested reaction products can be analyzed to yield the sequence of the target molecule.

In general, multiple sets of nested oligonucleotides must be evaluated in order to determine the sequence of the target molecule, however, various modifications, such as the use of multiple, distinguishable labels has led to the development of "multiplexing" methods that are capable of yielding increased sequence data (Church, G.M. et al., Science 240:185-188 (1988); Church, G.M. et al., U.S. Patent No. 4,942,124; Tabor, et al., U.S. Patent No. 4,962,020; Prober, J.M. et al., Science 238:336-340 (1987)).

Other "multiplexing" sequencing method such as described by Macevicz, S.C., (U.S. Patent 5,002,867) are directed to methods for determining the nucleotide sequence of a DNA or an RNA molecule using multiple mixed oligonucleotide probes. Sequence information is obtained by carrying out a series of hybridizations whose results provide for each probe the number of times the complement of the probe's sequence occurs in the RNA or DNA whose sequence is to be determined. The nucleotide sequence of the RNA or DNA is reconstructed from this information and from a knowledge of the probes' sequences. The nucleic acid whose sequence is to be determined is referred to herein as the target sequence.

The double-stranded structure of DNA complicates the sequence analysis process. Because the two strands of DNA are symmetrical and chemically identical, a sequence analysis that is conducted using both strands of a DNA molecule will yield two indistinguishable sets of sequence data. Thus, it is highly desirable to perform sequence analysis using preparations of DNA that contain only one of the two strands. Unfortunately, because the DNA strands are chemically indistinguishable, it is in general quite difficult to obtain DNA preparations that contain only one strand. The dideoxy sequencing method attempts to avoid this problem by employing either a DNA source that is single-stranded (such as a bacteriophage M13 or phagemid vector (Sambrook, J. et al., Molecular Cloning, a

Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), herein incorporated by reference)) or a primer that is capable of binding to only one strand of a target molecule. As will be appreciated, since the sequence of the target molecule is unknown, there can be no a priori assurance that a particular primer will not hybridize to both DNA strands. In the case of the Maxam-Gilbert method, it is in general necessary to label both strands of the target molecule, and then to selectively remove the label from one of the strands. These manipulations complicate the determinations of nucleic acid sequence.

In view of the deficiencies of the above-described methods for preparing single-stranded DNA, and in view of the importance of such methods to a variety of molecular biology and medical procedures, a method that would preferentially produce a single strand of a desired target molecule, and which could be used in conjunction with nucleic acid amplification procedures would be extremely desirable. The present invention provides such a method.

SUMMARY OF THE INVENTION:

The present invention provides a method for generating single-stranded DNA molecules following a primer-mediated extension or amplification reaction. Such molecules are useful as hybridization probes and in nucleic acid sequencing.

In detail, the invention provides a method for generating a desired single-stranded nucleic acid molecule, substantially free of any nucleic acid molecule of complementary sequence, said method comprising the steps:

A) incubating a preselected nucleic acid molecule in the presence of a primer molecule; wherein said primer molecule is capable of hybridizing to said preselected molecule, and wherein said primer molecule contains a region that is resistant to a 5'→3' exonuclease;

B) permitting template-dependent extension of said primer to thereby form said desired nucleic acid molecule; and

C) adding to said incubation a 5'→3' exonuclease, under conditions sufficient to eliminate said preselected molecule, and to thereby generate said desired single-stranded molecule substantially free of any nucleic acid molecule of complementary sequence.

The invention additionally includes the embodiment of the above method wherein in step B, after forming the desired nucleic acid molecule, the molecule is incubated in the presence of a second primer molecule capable of hybridizing thereto, and of being extended in a template-dependent manner to thereby form a nucleic acid molecule having a sequence substantially complementary to that of the desired molecule.

The invention also provides a composition of matter comprising a target nucleic acid molecule, hybridized to a primer molecule having a length of from about 10 to about 30 nucleotides and containing a nucleotide that confers 5'→3' exonuclease resistance to the primer molecule.

The invention particularly concerns the embodiment of the above methods wherein the exonuclease resistance of the region that is resistant to a 5'→3' exonuclease is caused by a plurality of phosphorothioate nucleotide derivatives.

The invention also provides a method of determining the identity of a nucleotide base at a specific position in a nucleic acid of interest, which comprises:

A) treating a sample containing the nucleic acid of interest, if the nucleic acid is double-stranded, so as to obtain unpaired nucleotide bases spanning the specific position, or directly employing step (B) if the nucleic acid of interest is single-stranded, wherein the nucleic acid of interest contains a sufficient number of 5'→3' exonuclease resistant nucleotide derivatives to confer 5'→3' exonuclease resistance to a region of said nucleic acid of interest;

B) contacting the sample from step (A), under hybridizing conditions, with an oligonucleotide primer which is capable of hybridizing with a stretch of nucleotide bases present in the nucleic acid of interest immediately adjacent to the nucleotide base to be identified, so as to form a duplex between the primer and the nucleic acid of interest such that the nucleotide base to be identified is the first unpaired base in the template immediately downstream of the 3' terminus of the primer in the duplex; and

C) contacting the duplex from step (B), in the substantial absence of dATP, dCTP, dGTP or dTTP, with at least two different nucleotide triphosphate derivatives, the derivatives including a derivative complimentary to the first unpaired base, and being terminators of a nucleic acid template-dependent, primer extension reaction; wherein at least one of the terminators is labeled with a detectable marker; and wherein the contacting is under conditions sufficient to permit base pairing of the complementary terminator derivative with the first unpaired base;

D) permitting occurrence of a template-dependent primer extension reaction sufficient to incorporate the complementary terminator derivative onto the 3' terminus of the primer;

E) determining the identity of the incorporated derivative, and thereby determining the identity of the nucleotide base at the specific position in the nucleic acid of interest.

The invention additionally includes the embodiments of the above method wherein in step (C), the duplex from step (B) is contacted with four terminators, wherein only one of the terminators has a detectable marker, and wherein the step (C) is performed four times, each time with a different one of the terminators being labelled; or wherein in step (C), the duplex

from step (B) is contacted with four labeled terminators, each with a different detectable label.

The invention also includes a method for detecting a desired exonuclease resistant amplification product of a polymerase chain reaction which comprises:

A) conducting a polymerase chain reaction with two primer molecules, wherein one of the primer molecules contains a sufficient number of phosphorothioate nucleotide derivatives (most preferably, about 4) at that primer's 5' terminus to render the terminus resistant to a 5'→3' exonuclease; the reaction being sufficient to form double-stranded amplification products;

B) subsequently treating the amplification products with a 5'→3' exonuclease under conditions to degrade oligonucleotides that lack a sufficient number of phosphorothioate bonds to render the oligonucleotides resistant to the exonuclease

C) detecting the desired amplification product of the polymerase chain reaction by permitting the product to hybridize to a complementary oligonucleotide bound to a solid support.

The invention also includes a method for minimizing cross contamination between polymerase chain reactions which comprises conducting a polymerase chain reaction wherein at least one of the primer molecules of the reaction contains a sufficient number of phosphorothioate nucleotide derivatives (i.e. about 4) at that primer's 3' terminus to render the terminus resistant to a 5'→3' exonuclease; and wherein, subsequent to conducting the polymerase chain reaction, amplification products of the reaction are incubated in the presence of said 5'→3' exonuclease under conditions sufficient to permit the degradation of oligonucleotide regions of unused primers and of amplification products that lack the sufficient number of phosphorothioate bonds; the degradation rendering the unused primer and the amplification products substantially incapable of serving as substrates in an additional polymerase chain reaction, and thereby minimizing cross contamination between polymerase chain reactions.

The invention also includes a kit, being specially adapted to contain in close compartmentalization a first container which contains a first primer, the first primer containing a phosphorothioate nucleotide derivative; and a second container which contains a second primer lacking any phosphorothioate nucleotide derivatives, such that the two primers can be used to amplify a predetermined gene sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS:

I. Introduction

The present invention provides a method for generating single-stranded DNA molecules, especially following the preparation of double-stranded nucleic acid molecules by an in vitro amplification procedure, such as PCR. The method employs nuclease resistant nucleotides derivatives, and incorporates, by chemical synthesis or enzymatic means, these derivatives into primer molecules in place of naturally occurring nucleotides.

The molecules that can be generated through the use of the present method can have a length ranging from a few nucleotides to several kilobases. The "desired" molecules of the invention are said to have a sequence that is "complementary," or substantially complementary to the sequence of a "target" strand of a nucleic acid molecule.

As used herein, two molecules are said to be complementary if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions (see, Sambrook, J. et al., Molecular Cloning, a Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), herein incorporated by reference).

The target molecule may be either DNA, cDNA or RNA; it may be either single-stranded or double-stranded. Where the "target" molecule is double-stranded, the invention distinguishes these strands as being either a "target" strand or a

"complement" strand (whose sequence is complementary to that of the target sequence). If the target molecule is double-stranded, the method of the present invention can be used to generate either of such strands.

Significantly, the method of the present invention permits one to generate single-stranded molecules that have the same length as the target molecule. The capacity to form full-length molecules (rather than molecules comprising only fragments of full-length molecules) greatly simplifies sequencing analysis, and facilitates the preparation of hybridization probes.

The present invention is capable of generating single-stranded molecules regardless of the nature, origin or sequence of the target molecule. Thus, the present invention can be used to generate single-stranded molecules that have a naturally occurring sequences, such as a sequence present in a virus (e.g. rhinovirus, hepatitis virus, herpes virus, HIV, etc.), a bacterium (e.g. Escherichia, Clostridium, Mycobacterium, Neisseria, Mycoplasma, Vibrio, Chlamydia, Rickettsia, etc.), a yeast, a fungus, or other lower eukaryote. In particular, the present invention can be used to generate single-stranded molecules that have sequence present in a plant cell, or an animal cell (especially a mammalian cell, such as from a horse, cow, dog, cat or human). The present invention can also be used to generate single-stranded molecules that are purely or partially synthetic (i.e. non-naturally occurring).

Significantly, the method of the invention permits the generated single-stranded molecules to be obtained "substantially free" of other sequences with which it is naturally associated. As used herein, the term "substantially free" denotes the reduction or elimination of at least one other sequence which is naturally associated with the obtained sequence or its complement.

The present invention accomplishes the generation of such single-stranded molecules through the use and extension of "primer" molecules that contain exonuclease resistant nucleotide derivatives. Examples of such modified nucleotide

derivatives are disclosed by Zon, G. et al. (Anti-Cancer Drug Design 6:539-568 (1991)) and Goodchild, J. et al. (Bioconjugate Chem. 1:613-629 (1990)), both of which references are incorporated herein by reference. In general, suitable nucleotide derivatives include derivatives in which one or two of the non-bridging oxygens of the phosphate moiety of a nucleotide has been replaced with a sulfur-containing group (especially a phosphothioate), an alkyl group (especially a methyl or ethyl alkyl group), a nitrogen-containing group (especially an amine), and/or a selenium-containing group, etc. For the purpose of the present invention, phosphorothioate nucleotide derivatives are the most preferable derivatives. A phosphorothioate nucleotide derivative (e.g. a nucleoside 5'-O-1-thiotriphosphate) contains a non-bridging (i.e. monocoordinate) sulfur in place of the oxygen atom of the orthophosphate moiety. As will be appreciated, the introduction of the sulfur permits two stereoisomers to form. Such a racemic mixture is suitable for the purposes of the present invention.

Importantly, the selected nucleotide derivative must be suitable for in vitro primer-mediated extension and provide nuclease resistance to the region of the nucleic acid molecule in which it is incorporated. In the most preferred embodiment, it must confer resistance to exonucleases that attack double-stranded DNA from the 5' end ("5'→3' exonucleases"). Examples of such exonucleases include bacteriophage T7 gene 6 exonuclease ("T7 exonuclease) and the bacteriophage lambda exonuclease ("λ exonuclease"). Both T7 exonuclease and λ exonuclease are inhibited to a significant degree by the presence of phosphorothioate bonds so as to allow the selective degradation of one of the strands. However, any double-strand specific, 5'→3' exonuclease can be used for this process, provided that its activity is affected by the presence of the bonds of the nuclease resistant nucleotide derivatives. The preferred enzyme when using phosphorothioate derivatives is the T7 gene 6 exonuclease, which shows maximal enzymatic activity in the same buffer used for many DNA dependent

polymerase buffers including Taq polymerase. The 5'→3' exonuclease resistant properties of phosphorothioate derivative-containing DNA molecules are discussed, for example, in Kunkel, T.A. (In: Nucleic Acids and Molecular Biology, Vol. 2, 124-135 (Eckstein, F. et al., eds.), Springer-Verlag, Berlin, (1988)). The 3'→5' exonuclease resistant properties of phosphorothioate nucleotide containing nucleic acid molecules are disclosed in Putney, S.D. et al. (Proc. Natl. Acad. Sci. (U.S.A.) 78:7350-7354 (1981)) and Gupta, A.P. et al. (Nucl. Acids. Res., 12:5897-5911 (1984)).

In addition to being resistant to such exonucleases, nucleic acid molecules that contain phosphorothioate derivatives at restriction endonuclease cleavage recognition sites are resistant to such cleavage. Taylor, J.W. et al. (Nucl. Acids Res. 13:8749-8764 (1985)) discusses the endonuclease resistant properties of phosphorothioate nucleotide containing nucleic acid molecules.

The nuclease resistance of phosphorothioate bonds has been utilized in a DNA amplification protocol (Walker, T.G. et al. (Proc. Natl. Acad. Sci. (U.S.A.) 89:392-396 (1992)). In the Walker et al. method, phosphorothioate nucleotide derivatives are installed within a restriction endonuclease recognition site in one strand of a double-stranded DNA molecule. The presence of the phosphorothioate nucleotide derivatives protects that strand from cleavage, and thus results in the nicking of the unprotected strand by the restriction endonuclease. Amplification is accomplished by cycling the nicking and polymerization of the strands.

Similarly, this resistance to nuclease attack has been used as the basis for a modified "Sanger" sequencing method (Labeit, S. et al. (DNA 5:173-177 (1986)). In the Labeit et al. method, ^{35}S-labelled phosphorothioate nucleotide derivatives were employed in lieu of the dideoxy nucleotides of the "Sanger" method.

As indicated, other methods (such as asymmetric PCR, etc.) have been used in attempts to generate single-stranded

molecules. The methods of the present invention offer the advantage that the double-stranded PCR product is quantitatively converted to a single-stranded product of exactly the same length. Second, the exonuclease used shows optimal enzymatic activity in PCR salts, thus no purification or buffer exchange is required prior to the exonuclease treatment. Finally, the resulting single-stranded molecule is completely resistant to any further degradation by the T7 gene 6 exonuclease.

The term "primer," as used herein, refers to a single-stranded oligonucleotide or a single-stranded polynucleotide that is capable of being extended by the covalent addition of a nucleotide in a "template-dependent" extension reaction. In order to possess such a capability, the primer must have a 3' hydroxyl terminus, and be hybridized to a second nucleic acid molecule (i.e. the "template"). A primer is typically 11 bases or longer; most preferably, a primer is 25 bases, however, primers of shorter or greater length may suffice. "Template-dependent" extension refers to the capacity of a polymerase to mediate the extension of a primer such that the extended sequence is complementary to the sequence of a nucleic acid template. A "polymerase" is an enzyme that is capable of incorporating nucleoside triphosphates to extend a 3' hydroxyl group of a nucleic acid molecule, if that molecule has hybridized to a suitable template nucleic acid molecule. Polymerase enzymes are discussed in Watson, J.D., In: Molecular Biology of the Gene, 3rd Ed., W.A. Benjamin, Inc., Menlo Park, CA (1977), which reference is incorporated herein by reference, and similar texts. For amplification purposes, a preferred DNA polymerase is Taq polymerase (Cetus). Other polymerases such as the large proteolytic fragment of the DNA polymerase I of the bacterium E. coli, commonly known as "Klenow" polymerase, E. coli DNA polymerase I, and bacteriophage T7 DNA polymerase, may also be used to perform the method described herein.

Conditions or agents which increase the rate or the extent of priming, primer elongation, or strand displacement, may

increase the extent of the amplification obtained with the methods of the present invention. For instance, the addition of helicases or single-stranded nucleic acid binding proteins may increase the strand displacement rate of a DNA polymerase, or may allow the use of a DNA polymerase that might not ordinarily give substantial amplification.

All of the enzymes used in an amplification reaction may be active under the same reaction conditions. Indeed, buffers exist in which all enzymes are near their optimal reaction conditions. It is desirable to provide to the reaction mixture an amount of required co-factors such as Mg^{++}, and dATP, dCTP, dGTP, dTTP, ATP, CTP, GTP, UTP or other nucleoside triphosphates in sufficient quantity to support the degree of amplification desired. Equivalent nucleoside triphosphate analogues, etc. (Piccirilli, J.A. et al., Nature 343:33-37 (1990) can be substituted or added to those specified above, provided that the base pairing, polymerase and strand displacing functions are not adversely affected to the point that the amplification does not proceed to the desired extent.

When defining conditions to be used in any specific embodiment of the present invention, primer mediated, target-independent reactions may occur which may reduce amplification efficiency, and should be examined during assay optimization. For this reason, primers should be chosen which cannot prime on themselves. Primers can also act as DNA templates in unusual promoter-independent transcription reactions (Krupp, G., Nucl. Acids Res. 17:3023-3036 (1989)). To minimize the likelihood of potential interfering reactions, candidate primers should preferably be tested in reactions which address these issues prior to their use in the amplification process.

In a preferred embodiment of the invention, the single-stranded molecules of the present invention, or amplification products thereof are detectably labelled. Any suitable means of detectable labelling may be employed; thus, the label may be an enzyme label, a fluorescent label, a radioisotopic label, a

chemiluminescent label, etc. Examples of suitable enzyme labels include alkaline phosphatase, acetylcholine esterase, alpha-glycerol phosphate dehydrogenase, alkaline phosphatase, asparaginase, ß-galactosidase, catalase, delta-5-steroid isomerase, glucose oxidase, glucose-6-phosphate dehydrogenase, glucoamylase, glycoamylase, luciferase, malate dehydrogenase, peroxidase, ribonuclease, staphylococcal nuclease, triose phosphate isomerase, urease, and yeast-alcohol dehydrogenase. Examples of suitable fluorescent labels include a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, a fluorescamine label, etc. Examples of suitable chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, an aequorin label, etc.

II. The Preferred Method of the Invention

As indicated, the invention accomplishes the generation of single-stranded molecules through the use of primer molecules that contain exonuclease resistant nucleotide derivatives, most preferably phosphorothioate deoxyribonucleotide or ribonucleotide derivatives. Any of a variety of chemical methods may be used to produce such phosphorothioate derivatives (see, for example, Zon, G. et al., Anti-Canc. Drug Des. 6:539-568 (1991); Kim, S.G. et al., Biochem. Biophys. Res. Commun. 179:1614-1619 (1991); Vu, H. et al., Tetrahedron Lett. 32:3005-3008 (1991); Taylor, J.W. et al., Nucl. Acids Res. 13:8749-8764 (1985); Eckstein, F. et al., Biochemistry 15:1685-1691 (1976); Ludwig, J. et al., J. Org. Chem. 54:631-635 (1989)). Phosphorothioate nucleotide derivatives can also be obtained commercially from Amersham or Pharmacia.

In the most preferred embodiment, the phosphorothioate derivative is included in the primer. Preferably, the primer molecules will be approximately 25 nucleotides in length, and

contain from about 4% - to about 100%, and more preferably from about 4% - to about 40%, and most preferably about 16%, phosphorothioate residues (as compared to total residues). The nucleotides may be incorporated into any position of the primer, and may be adjacent to one another, or interspersed across all or part of the primer. Most preferably, however, the phosphorothioate residues will be adjacent to one another, and will be incorporated at the 5' terminus of the primer.

In one embodiment, the present invention can be used in concert with an amplification protocol, for example, PCR. In this embodiment, it is preferred to limit the number of phosphorothioate bonds of the primers to about 10 (or approximately half of the length of the primers), so that the primers can be used in a PCR reaction without any changes to the PCR protocol that has been established for non-modified primers. When the primers contain more phosphorothioate bonds, the PCR conditions may require adjustment, especially of the annealing temperature, in order to optimize the reaction. Incorporation of less than 4 phosphorothioates leads to incomplete exonuclease protection. The use of primers that contain 4 phosphorothioate bonds is, therefore, preferred.

The incorporation of such nucleotide derivatives into DNA or RNA can be accomplished enzymatically, using a DNA polymerase (Vosberg, H.P. et al., Biochemistry 16: 3633-3640 (1977); Burgers, P.M.J. et al., J. Biol. Chem. 254:6889-6893 (1979); Kunkel, T.A., In: Nucleic Acids and Molecular Biology, Vol. 2, 124-135 (Eckstein, F. et al., eds.), Springer-Verlag, Berlin, (1988); Olsen, D.B. et al., Proc. Natl. Acad. Sci. (U.S.A.) 87:1451-1455 (1990); Griep, M.A. et al., Biochemistry 29:9006-9014 (1990); Sayers, J.R. et al., Nucl. Acids Res. 16:791-802 (1988)). Alternatively, phosphorothioate nucleotide derivatives can be incorporated synthetically into an oligonucleotide (Zon, G. et al., Anti-Canc. Drug Des. 6:539-568 (1991)).

The primer molecules are permitted to hybridize to a complementary target nucleic acid molecule, and are then extended, preferably via a polymerase, to form an extension

product. The presence of the phosphorothioate nucleotides in the primers renders the extension product resistant to nuclease attack. As indicated, the amplification products containing phosphorothioate or other suitable nucleotide derivatives are substantially resistant to "elimination" (i.e. degradation) by "5'→3'" exonucleases such as T7 exonuclease or λ exonuclease, and thus a 5'→3' exonuclease will be substantially incapable of further degrading a nucleic acid molecule once it has encountered plurality of phosphorothioate residues (most preferably, about 4 (i.e. 3-5). The use of additional numbers of phosphorothioate residues is equivalnent to the use of four such residues.

Since the target molecule lacks nuclease resistant residues, the incubation of the extension product and its template - the target - in the presence of a 5'→3' exonuclease results in the destruction of the template strand, and thereby achieves the preferential production of the desired single strand.

III. Uses of the Single-Stranded Molecules Generated by the Present Invention

A. Hybridization Substrates

As indicated, the target molecule can be either single-stranded or double-stranded, and can be either DNA or RNA. Although the method of the present invention is capable of generating a single molecular species upon the amplification of a double-stranded molecule, there is no constraint on which of the strands is to be amplified. Because methods, such as PCR, result in the amplification of double-stranded molecules regardless of whether the source target molecule was initially single-stranded or double-stranded, the present invention permits either strand of an initial double-stranded molecule to be generated. Similarly, either the initial strand of a single-stranded molecule, or either the complement of that strand can be generated by the method of the present invention.

Thus, for example, the present invention can be used to either form cDNA corresponding in sequence to an mRNA molecule, or it can be used to generate an "antisense" molecule, capable of hybridizing to that mRNA molecule. "Antisense" molecules may be used to detect and identify pathogens (either viral or bacterial) in tissue (including blood, spinal fluid, tumorous tissue, etc.), food, water, milk, etc. They may also be used to evaluate the persistence or significance of latent viral or bacterial infection. In one embodiment of such a use, the single-stranded molecules generated by the invention are preferably detectably labeled, and used as hybridization probes of the target molecule. In another embodiment, the single-stranded molecules of the present invention (either labelled or unlabelled) can be amplified, using PCR, or other means, to produce amplification products that have been detectably labelled. Since such labelling can, if desired, be incorporated throughout the amplification product, this embodiment permits a higher specific activity of labelling than is obtainable through end-labelling.

The therapeutic use of antisense molecules derives from the capacity of such molecules, if incorporated into a cell, to hybridize to an mRNA molecule of complementary sequence, and thereby impair (i.e. attenuate or prevent) the translation of that mRNA molecule into a gene product. To act as an antisense oligonucleotide, the nucleic acid molecule must be capable of binding to or hybridizing with that portion of target mRNA molecule (or gene) which mediates the translation of the target mRNA.

The single-stranded nucleic acid molecules generated by the presently disclosed method can also be used to obtain oligonucleotides, such as are employed in oligonucleotide-based diagnostic assays of nucleic acid sequence variation, and in particular, the "Genetic Bit Analysis ("GBA™") method disclosed by Goelet, P. et al. (WO 92/15712, herein incorporated by reference). GBA™ is a method for detecting single nucleotide genetic polymorphisms in nucleic acid samples that relies on a

rapid, non-radioactive, solid-phase assay procedure. In essence, locus-specific DNA primers are coupled to a solid phase and hybridized to genomic templates, then extended, preferably by Klenow or T7 DNA polymerases in a sequence-directed fashion; the substrates for this chain-extension reaction are preferably novel chain terminating dideoxynucleotides having a covalently attached biotin moiety. The particular base(s) incorporated in a given reaction can then be read via a colorimetric reaction using commercially available enzyme conjugates. The reactions have been adapted to an ELISA-like 96-well format and automated using standard robotic liquid handling systems.

Modern gene mapping strategies rely on the accumulation of informative genetic markers at closely-spaced intervals along a genome. One of the advantages of GBA™ is that its use of standard reaction conditions enables tests for newly-defined single nucleotide polymorphisms to be readily developed. GBA™ also allows preliminary allelic frequencies in a population to be rapidly determined so that the informativeness of a new marker can be conveniently assessed.

Thus, in GBA™, purified oligonucleotides having a defined sequence (complementary to a target molecule) are bound to a solid support. A sample, suspected to contain the target molecule is placed in contact with the support, and any target molecules present are permitted to hybridize to the bound oligonucleotide. In one embodiment, the 5' terminus of the oligonucleotide is attached to the solid support, as described, for example by Nickerson et al. (Proc. Natl. Acad. Sci. (U.S.A.) 87:8923-8927 (1990)), such that the 3' end can serve as a substrate for primer extension. The presence of the desired molecule is determined by the incorporation of a labelled nucleotide to the 3' terminus of the bound oligonucleotide by a primer-dependent polymerase.

The methods of the present invention can be used to prepare modified single-stranded oligonucleotides including oligonucleotides modified for the attachment of detectable

reporter groups or oligonucleotides modified for attachment to a solid support matrix (Ruth, J.L., U.S. Patent 4,948,882)).

The method of the present invention provides several salient advantages. The present invention provides a highly convenient and reliable method for preparing full-length, or partial length, single-stranded DNA molecules following the synthesis of double-stranded DNA molecules by a primer-directed nucleic acid amplification reaction, e.g., PCR. Significantly, the degradation of the nuclease sensitive strand can be carried out without prior isolation or purification of the double-stranded PCR amplification product.

In contrast to the previously discussed method of Higuchi, R.G. et al., which typically gives only 50-70% conversion, even when excess λ exonuclease is used, the method of the present invention typically yields fully quantitative degradation of the nuclease sensitive strand.

B. Amplification

As suggested above, the method of the present invention is advantageously coupled with an in vitro amplification procedure, in order to specifically amplify a single strand of a double-stranded molecule. This aspect of the invention is illustrated below by reference to PCR, however any of the previously described amplification procedures may alternatively be used.

For this purpose, PCR is performed using two primers, only one of which has been modified to contain nuclease resistant nucleotide derivatives, such as phosphorothioate nucleotides. The resulting nuclease resistant bonds become an integral part of the "target strand" of the double-stranded PCR amplification product. In contrast, the "complement strand" of the PCR amplification product, which is formed from the primer that lacked nuclease resistant nucleotide derivatives, is sensitive of nuclease degradation. Following the PCR amplification, the double-stranded DNA product obtained will contain phosphorothioate bonds at the 5' terminus of only one strand.

The use of a suitable double-strand-specific, 5'→3' exonuclease therefore converts this product to a single-stranded molecule by the selective degradation of the non-protected complement strand. The phosphorothioate bonds present in the desired strand protect it from enzymatic hydrolysis. Preferably then, after the PCR reaction the exonuclease (preferably, T7 gene 6 exonuclease) can simply be added directly to the reaction mixture and the hydrolysis of the non-protected strand can be carried out, either at room temperature or, more preferably, at 37°C for 15-30 minutes. When the λ exonuclease is used, the reaction mixture is most preferably adjusted to a pH of 9.4 (the optimal pH of this enzyme); significantly more enzyme should be used if the complete degradation of the nuclease-sensitive strand is desired. Since λ exonuclease shows a significant preference for 5'-phosphorylated substrates over non-phosphorylated ones, the nuclease sensitive PCR primer is most preferably 5'-phosphorylated in order to obtain optimal results with this enzyme.

Thus, since the 5'→3' exonuclease causes the "complement strand" to be degraded, and a preparation of "target strand" that is substantially free of natural contaminants is obtained. The single-stranded target molecules can be used as hybridization probes, as sequencing templates, or in other applications that require single-stranded DNA.

C. Sequence Analysis

As indicated, the single-stranded molecules generated by the present invention can be used to sequence a target molecule. In one embodiment of the invention, the primer that contains the phosphorothioate nucleotide derivatives is preferably labelled, such that the extension product that is formed from the primer can be readily detected or visualized. Any suitable label, such as a radioisotope, enzyme, fluorescent moiety, chemiluminescent moiety, etc., may be used for this purpose. In an alternate embodiment, the label will be incorporated into the

phosphorothioate nucleotide derivative, as through the use of a radioactive sulfur isotope (i.e. ^{35}S). In yet another embodiment, the single-stranded molecules of the present invention (either labelled or unlabelled) can be amplified, using PCR, or other means, to produce amplification products that have been detectably labelled. As indicated above, such labelling can, if desired, be incorporated into an amplification product, obtained from PCR, or by other means, in order to obtain higher specific activity of labelling than would be obtainable through end-labelling.

Thus, the method of the invention permits the preparation of single-stranded molecules that are labelled either at its 5' terminus, or optionally, throughout the molecule. As such, the molecules can be rapidly and efficiently sequenced using the previously described Maxam-Gilbert sequencing method.

The present invention includes articles of manufacture, such as "kits." Such kits will, typically, be specially adapted to contain in close compartmentalization a first container which contains a first primer containing a phosphorothioate nucleotide derivative, and a second container which contains a second primer, not containing any phosphorothioate nucleotide derivatives, such that the two primers can be used to amplify a predetermined gene sequence. The kit may additionally contain buffers, enzymes, instructional brochures, and the like.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

EXAMPLE 1

PRODUCTION OF SINGLE-STRANDED DNA

A single-stranded molecular species corresponding to a 257 bp region of horse genomic DNA was generated through the use of PCR and two 25 residue long primers having four

phosphorothioate bonds at their 5' ends ("ps" denotes a phosphorothioate bond):

SEQ ID NO:1

C(ps)T(ps)C(ps)C(ps)ACCCTTATGAACTCATTGAAT

SEQ ID NO:2

T(ps)G(ps)T(ps)G(ps)TGTGCTATAAATGCCACTAAC

The phosphorothioate bonds were introduced during the automated synthesis of the oligonucleotides using the commercially available reagent tetraethyl thiuram disulfide (TETD). The phosphorothioate-modified PCR primers were purified using commercially available Oligonucleotide Purification Cartridges (OPC).

The same primers (SEQ ID NO:1 and SEQ ID NO:2) were also prepared in their non-modified form. In all PCR reactions, where generation of single-stranded product was required, one of the PCR primers was phosphorothioate-modified and the other was non-modified. PCR amplifications were carried out for 30 or 35 cycles where each cycle of amplification includes one minute denaturation, two minutes annealing at 60°C, and three minutes extension at 72°C. Following the PCR amplification, a dilution of T7 gene 6 exonuclease (approximately 16 units of enzyme for 100 µl of PCR reaction) was added and the mixture incubated for 15 minutes at 37°C. The reaction was stopped by the addition of EDTA to 10 mM and analyzed by polyacrylamide gel electrophoresis.

The PCR amplification resulted in the exponential amplification of the 257 base pair sequence separating the respective primer binding sites. Treatment with exonuclease resulted in the complete degradation of the nuclease sensitive strand. Electrophoretic analysis of the amplified material after exonuclease treatment revealed that the material had been converted into a 257 base long single-stranded form.

EXAMPLE 2

STABILITY OF PHOSPHOROTHIOATE BONDS TO HYDROLYSIS BY T7 GENE 6 EXONUCLEASE

The bacteriophage T7 gene 6 exonuclease hydrolyses double-stranded DNA in the 5' to 3' direction. In order to study the effect on the enzyme activity of the substitution of regular phosphodiester bonds with phosphorothioates, the following 3' biotinylated, self-complementary oligonucleotides were synthesized (45 mers; "X" indicates the presence of a phosphorothioate linkage between the adjacent nucleotides; B denotes a biotin residue):

Oligo #1 (SEQ ID NO. 3): CCGCGTGGTGCCTGGTGCCCTTTTT
GGGCACCAGGCACCACGCGG-B

Oligo # 2 (SEQ ID NO: 4): $C_XCGCGTGGTGCCTGGTGCCCTTTTT$
GGGCACCAGGCACCACGCGG-B

Oligo # 3 (SEQ ID NO: 5): $C_XC_XG_XC_XGTGGTGCCTGGTGCCCTTTTT$
GGGCACCAGGCACCACGCGG-B

Oligonucleotides #1-3 were synthesize trityl-on, purified by reversed-phase HPLC, detritylated by treatment with 80% acetic acid, and desalted. Oligonucleotide #1 does not contain phosphorothioate bonds at its 5' end. Oligonucleotide #2 contains one phosphorothioate bond at the 5" end. It is therefore a mixture of two diastereoisomers, Rp and Sp, depending upon the orientation of the phosphorothioate residue. These two diastereoisomers were well separated by reverse-phase HPLC at the trityl-on level and obtained in pure from after detritylation. The two individual diastereoisomers of oligonucleotide #2 thus obtained are referred to below as peak A (eluting earlier) and peak B (eluting later).

Oligonucleotides #1-3 were designed to form stable hairpin-type self-complementary secondary structures, with a single-stranded loop of five thymidine residues. Upon treatment with T7 gene 6 exonuclease, these oligonucleotides should be

hydrolyzed from the 5' end up to the thymidine loop, and would thereby be converted to single-stranded molecules. To capture these resulting 3' biotinylated single-stranded oligonculeotides onto a solid phrase by hybridization, the oligonucleotide #4 was immobilized in 96 well plates. This oligonucleotide has the sequence:

Oligo #4 (SEQ ID NO:6): AGCCTCAGACCGCGTGGTGCCTGGT

The sixteen 3' terminal bases of this oligonucleotide are complementary to the 3' ends of the biotinylated oligonucleotides #1-3.

Approximately 60 pmoles of the purified oligonucleotides #1-3 were treated with either 0 or 4 units/µl of T7 gene 6 exonuclease, at 37°C in a total volume of 100 µl. Following this treatment, aliquots were removed at intervals and mixed with an equal volume of 3 M NaCl, 20 mM EDTA. After an additional dilution step in 1.5 M NaCl, 10 mM EDTA, aliquots containing approx. 1 pmol of oligonucleotide were added to the wells of a 96 well plate containing the immobilized oligonucleotide #4.

The presence of absence of biotin was then detected in a colorimetric assay. This assay was performed in the following manner. After the hybridization, the plate was incubated with a 1:1200 dilution of anti-biotin horseradish peroxidase conjugate (Vector Laboratories, Burlingame, CA) in 1% BSA in TNTw, for 30 min at room temperature. The plate was then washed six times with TNTw and a solution of 1 mg/ml of o-phenylenediamine (OPD) in 0.1 M citrate buffer, PH 4.5, containing 0.012% H_2O_2 was added. The plate was immediately placed in a plate reader (v_{max}, Molecular Devices), and the development of color was followed at 450 nm for 2 min. The results were expressed as mOD^{450}/min.

The results of this assay are summarized in Table 1. The signals given in this Table are those obtained after 15 min treatment with exonuclease. No increase in signal was seen upon longer incubation. Table 1 thus shows the effect of

phosphorothioate residues on the activity of T7 gene 6 exonuclease.

TABLE 1

Oligo #	Number of 5' phosphorotioate residues	Signal without exonuclease treatment	Signal after treatment with 4 u/µl of exonuclease
# 1	0	4	196
#2 peak A	1	4	220
#2 peak B	1	4	180
# 3	4	5	6

Several important results emerged from these experiments. As expected, none of the self-complementary, double-stranded oligonucleotides was able to hybridize to the solid-phase immobilized oligonucleotide. Hybridization only took place if a single-stranded, biotinylated oligonucleotide was obtained by treatment with T7 gene 6 exonuclease. In this assay, oligonucleotide #1, as well as both diasteroisomers of oligonucleotide #2 were found to be equally good substrates for the exonuclease. Thus, the presence of only one phosphorothioate residue does not provide sufficient protection. In contrast, four phosphorothioate residues at the 5' end of oligonucleotide #3 provided complete protection from the hydrolytic activity of T7 gene 6 exonuclease. Most likely, the enzyme is capable of bypassing the 5' terminal phosphorothioate bond and starting the hydrolysis from the next phosphodiester.

EXAMPLE 3

COLORIMETRIC DETECTION OF PCR PRODUCTS IN 96 WELL PLATES

Having established that phosphorothioate bonds can provide protection from the hydrolytic action of T7 gene 6 exonuclease, PCR primers were prepared containing four internucleotidic phosphorothioate bonds at their 5' ends. A fifth phosphorothioate bond links the 5' terminal nucleotide of these

primers to a biotin residue, which allows the nonradioactive detection of the PCR products. These labeled primers were used together with unmodified opposite strand primers to amplify fragments from horse genomic DNA. The sequences of the PCR primers and capture probes are as follows:

Primer pair A
SEQ ID NO: 7: B_x-$C_xC_xA_xA_x$AGGAGCTGGGTCTGAAACAAA
SEQ ID NO: 8: ATGGCTTCCCACCCTACCCATCCCG

The amplification product of primer pair A was 93 base pairs long, and was captured using a capture probe having the following sequence:

SEQ ID NO:9: TGTTCTGGGAAAGACCACATTATTT

Primer pair B:
SEQ ID NO: 10: B_x-$A_xT_xG_xC_x$TCCCAGGTGATTCCAGTGTGC
SEQ ID NO: 11: GGTGCTGTGCGAGGTACACTTGACTG

The amplification product of primer pair B was 201 base pairs long, and was captured using a capture probe having the following sequence:

SEQ ID NO: 12: AGAAACACAAGGCCCAAGAACAGGA

Primer pair C:
SEQ ID NO: 13: B_x-$G_xG_xA_xT_x$CCAGATGAACAACCAGATGAA
SEQ ID NO: 14: CTGCAGCCCACTGGGCCTTCTTTGT

The amplification product of primer pair C was 547 base pairs long, and was captured using a capture probe having the following sequence:

SEQ ID NO: 15: CCTTTGTGTAGAGTAGTTCAAGGAC

For all PCR reactions, negative controls were carried out that contained all reactions components with the exception of

the horse genomic DNA. A positive result of such a control reaction would indicate contamination of one of the reaction components by a previously obtained PCR product.

Following the PCR amplification, aliquots of the reaction mixtures were withdrawn and saved as double-stranded PCR controls, while the rest of the mixtures was treated with T7 gene 6 exonuclease. Analysis was then carried out using polyacrylamide gel electrophoresis and also by hybridization of the single-stranded products of the exonuclease reaction to oligonucleotide probes immobilized in 96 well plates. The capture oligonucleotides were designed to hybridize to internal regions of the PCR products, thereby eliminating the possible capture of primer-dimers. Following the hybridization step, the presence or absence of biotin was determined with a colorimetric reaction using an anti-biotin horseradish peroxidase conjugate.

The results of polyacrylamide gel electrophoretic analysis of the PCR products illustrated that the exonuclease used hydrolysed the unmodified DNA strand, and left the phosphorothioated strand intact.

To demonstrate the specificity of hybridization, each of the same three PCR products after the exonuclease treatment was hybridized to wells that contained each of the three capture oligonucleotides. Thus, the products of PCR reactions, A, B, and C were rendered single-stranded by treatment with 2 units/μl of T7 gene 6 exonuclease and aliquots corresponding to 5 μl of the initial PCR action were added to the wells of a microliter plate containing the appropriate capture oligonucleotides for hybridization. The results of the colorimetric assay are presented in mOD_{450}/min. All experiments were carried out in duplicate; the results shown are averages (NT=not tested, "-control"=negative control). The results of the microliter plate hybridization assay are summarized in Table 2. It should be noted that no hybridization signals were obtained using the double-stranded PCR products directly, without the exonuclease step. This again illustrates that the exonuclease used

hydrolyses the unmodified DNA strand, and leaves the phosphorothioated strand intact. The results of this cross-hybridization experiment are also included in Table 2. Each of the three PCR products hybridized only to its specific capture oligonucleotide.

TABLE 2

PCR reaction	hybridization to capture oligo for reaction	Signal without exonuclease treatment	Signal after exonuclease treatment (2 u/µl)
A	A	2	450
	B	NT	3
	C	NT	1
A (- control)	A	NT	1
B	A	NT	4
	B	1	630
	C	NT	1
B (- control)	B	NT	4
C	A	NT	3
	B	NT	1
	C	2	450
C (- control)	C	NT	4

EXAMPLE 4

USE OF PHOSPHOROTHIOATE PCR PRIMERS FOR PCR PRODUCTS STERILIZATION

One embodiment of the present invention concerns the placement of phosphorothioate bonds at the 3' rather than the 5' end of the PCR primers. Upon treatment with T7 gene 6 exonuclease, the 5' unmodified parts of the double-stranded PCR products will be degraded up to the phosphorothioate bonds. The resulting product can be either single-stranded or double-stranded, depending on whether only one PCR primer contained phosphorothioates, or both. In both cases, assuming a very high efficiency of the exonuclease reaction, the resulting products

should not be reamplifiable in a subsequent polymerase chain reaction that uses the same primers, since the parts of the molecule where the primers should hybridize will have been destroyed. This could constitute an alternative method to preventing PCR crosscontamination.

EXAMPLE 5

TYPING OF DNA SINGLE-BASE POLYMORPHISMS BY GBA™.

As indicated above, Genetic Bit™ Analysis (GBA™) is a solid-phase method for the typing of single-nucleotide polymorphisms. In this method, oligonucleotide primers (called GBA™ primers) are immobilized on solid phases like polystyrene or glass, hybridized to single-stranded PCR templates obtained by the method of the present invention, and subjected to enzymatic extension at their 3' ends by a single, labeled ddNTP. The nature of the incorporated ddNTP is determined by the nucleotide that is located in the opposite strand (the polymorphic nucleotide). This assay can be conveniently carried out in polystyrene ELISA plates, on polystyrene pins, or on glass slides. A typical example of GBA™ carried out in a polystyrene plate is given below. In this example, GBA™ is used to type a diallelic polymorphism in equine genomic DNA.

The use of phosphorothioate-containing oligonucleotides in GBA™ is illustrated by the use of PCR primers to amplify a 112 bp region from equine genomic DNA that contains a single-base polymorphism. The PCR primers had the following sequences:

(SEQ ID NO: 16) ATAATACAGAAGTTCTGAGAGGCTA
(SEQ ID NO: 17) $_xG_xG_xA_x$TCCAGGTCTGCTTCTGCTTCCC

The PCR primer of SEQ ID NO:17 contains at its 5' end four phosphorothioate bonds. These protect this end of the double-stranded PCR product from the exonucleolytic action of the T7 gene 6 exonuclease and allow the preparation of single-stranded PCR product.

Genomic DNA isolated from four different horses was used. The amplification by PCR was carried out by standard techniques, using the oligonucleotides SEQ ID NO:16 and SEQ ID NO:17 as primers. The double-stranded PCR product was converted to the single-stranded form as described, and it was hybridized to a GBA™ primer having the following sequence:

(SEQ ID NO: 18) 5' AAGAGAAAGAGTTTTGCCTCAATCC

This GBA™ primer was immobilized on a polystyrene 96 well plate. Following the hybridization of the PCR-derived, single-stranded DNA fragment to the immobilized GBA™ primer, the 3' end of the latter was enzymatically extended by one labelled ddNTP, in the presence of the large fragment of DNA polymerase I from E. coil (Klenow polymerase). The extension mixture used contained the following components: 20mM Tris-HCl, pH 7.5; 10 mM $MgCl_2$; 25 mM NaCl; 10 mM $MnCl_2$; 15 mm sodium isocitrate; 1.5 μM of each of three unlabeled 2', 3'-dideoxynucleoside 5'-triphosphates and either 1.5 μM of biotin-labelled 2',3'-dideoxyadenosine 5'-triphosphate or 1.5 μM of biotin-labelled 2',3'-dideoxyguanosine 5'-triphosphate; and 0.15 units of the Klenow polymerase per well. The extension was carried out in separate wells, each containing a different labeled ddNTP. The presence of biotin was then revealed by a colorimetric detection as described above. The results of this experiment are shown in Table 3.

Table 3

Horse #	mOD_{450}/min Base A	mOD_{450}/min Base G	Genotype
1	115	80	AG
2	2	150	GG
3	75	90	AG
4	85	1	AA

These results show that, for this polymorphism, horses 1 and 3 are heterozygotes, horse 2 is a G homozygote, and horse 4 is an A homozygote.

The GBA™ (genetic bit analysis) method is thus a simple, convenient, and automatable genotyping method. In this method, sequence-specific annealing to a solid phase-bound primer is used to select a unique polymorphic site in a nucleic acid sample, and interrogation of this site is via a highly accurate DNA polymerase reaction using a set of novel non-radioactive dideoxynucleotide analogs. One of the most attractive features of the GBA™ approach is that, because the actual allelic discrimination is carried out by the DNA polymerase, one set of reaction conditions can be used to interrogate many different polymorphic loci. This feature permits cost reductions in complex DNA tests by exploitation of parallel formats and provides for rapid development of new tests.

The intrinsic error rate of the GBA™ procedure in its present format is believed to be low; the signal-to-noise ratio in terms of correct vs. incorrect nucleotide incorporation for homozygotes appears to be approximately 20:1. GBA™ is thus sufficiently quantitative to allow the reliable detection of heterozygotes in genotyping studies. The presence in the DNA polymerase-mediated extension reaction of all four dideoxynucleoside triphosphates as the sole nucleotide substrates heightens the fidelity of genotype determinations by suppressing misincorporation. GBA™ can be used in any application where point mutation analyses are presently employed -- including genetic mapping and linkage studies, genetic diagnoses, and identity/paternity testing --assuming that the surrounding DNA sequence is known.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come

within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Nikiforov, Theo T.
Knapp, Michael R.

(ii) TITLE OF INVENTION: METHOD FOR GENERATING SINGLE-STRANDED DNA MOLECULES

(iii) NUMBER OF SEQUENCES: 18

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Howrey & Simon
(B) STREET: 1299 Pennsylvania Avenue, NW
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: US
(F) ZIP: 20004

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Auerbach, Jeffrey I.
(B) REGISTRATION NUMBER: 32,680

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 383-7451
(B) TELEFAX: (202) 383-6610

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACCCTT ATGAACTCAT TGAAT 25

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGTGTGTGCT ATAAATGCCA CTAAC 25

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGCGTGGTG CCTGGTGCCC TTTTTGGGCA CCAGGCACCA CGCGG 45

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCGCGTGGTG CCTGGTGCCC TTTTTGGGCA CCAGGCACCA CGCGG 45

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGCGTGGTG CCTGGTGCCC TTTTTGGGCA CCAGGCACCA CGCGG 45

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCCTCAGAC CGCGTGGTGC CTGGT 25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCAAAGGAGC TGGGTCTGAA ACAAA 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGCTTCCC ACCCTACCCA TCCCG 25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGTTCTGGGA AAGACCACAT TATTT 25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGCTCCCAG GTGATTCCAG TGTGC 25

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGTGCTGTGC GAGGTACACT TGACTG 26

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGAAACACAA GGCCCAAGAA CAGGA 25

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGATCCAGAT GAACAACCAG ATGAA 25

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTGCAGCCCA CTGGGCCTTC TTTGT 25

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTTTGTGTA GAGTAGTTCA AGGAC 25

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATAATACAGA AGTTCTGAGA GGCTA 25

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGATCCAGGT CTGCTTCTGC TTCCC 25

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Equus caballus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAGAGAAAGA GTTTTGCCTC AATCC 25

WHAT IS CLAIMED IS:

1. A method for generating a desired single-stranded nucleic acid molecule, substantially free of any nucleic acid molecule of complementary sequence, said method comprising the steps:

A) incubating a preselected nucleic acid molecule in the presence of a primer molecule; wherein said primer molecule is capable of hybridizing to said preselected molecule, and wherein said primer molecule contains a region that is resistant to a 5'→3' exonuclease;

B) permitting template-dependent extension of said primer to thereby form said desired nucleic acid molecule; and

C) adding to said incubation a 5'→3' exonuclease, under conditions sufficient to eliminate said preselected molecule, and to thereby generate said desired single-stranded molecule substantially free of any nucleic acid molecule of complementary sequence.

2. The method of claim 1, wherein in step B, after forming said desired nucleic acid molecule, said molecule is incubated in the presence of a second primer molecule capable of hybridizing thereto, and of being extended in a template-dependent manner to thereby form a nucleic acid molecule having a sequence substantially complementary to that of said desired molecule.

3. The method of claim 1, wherein in step B, after forming said desired nucleic acid molecule, said molecule is incubated in the presence of a second primer molecule capable of hybridizing thereto, and with at least one dideoxynucleotide derivative, but in the absence of any non-terminating deoxynucleotide derivatives; wherein said incubation is under conditions sufficient to permit the extension of said primer in a template-dependent manner.

4. The method of claim 1, wherein said primer has a length of from about 10 to about 30 nucleotides.

5. The method of claim 4, wherein said primer has a length of from about 20 to about 25 nucleotides.

6. The method of claim 1, wherein the exonuclease resistance of said region that is resistant to a 5'→3' exonuclease is caused by a plurality of phosphorothioate nucleotide derivatives.

7. The method of claim 6, wherein said region contains about 4 phosphorothioate nucleotide derivatives.

8. The method of claim 1, wherein said primer is detectably labelled.

9. The method of claim 8, wherein said nucleotide that confers 5' → 3' exonuclease resistance to said primer molecule is detectably labelled.

10. The method of claim 7, wherein said detectable label is selected from the group consisting of an enzyme label, a fluorescent label, a radioisotopic label, and a chemiluminescent label.

11. The method of claim 1, wherein said 5'→3' exonuclease is T7 gene 6 exonuclease.

12. The method of claim 1, wherein said 5'→3' exonuclease is λ exonuclease.

13. The method of claim 1, wherein said desired single-stranded nucleic acid molecule is detectably labelled by the incorporation of labelled nucleotides during the template-dependent extension of the primer.

14. The method of claim 13, wherein said detectable label is selected from the group consisting of an enzyme label, a fluorescent label, a radioisotopic label, and a chemiluminescent label.

15. A composition of matter comprising a target nucleic acid molecule, hybridized to a primer molecule having a length of from about 10 to about 30 nucleotides and containing a region that confers 5' → 3' exonuclease resistance to said primer molecule.

16. The composition of matter of claim 15, wherein said region that confers 5'→3' exonuclease resistance to said primer molecule contains about four phosphorothioate bonds.

17. The composition of matter of claim 15, wherein said primer molecule is bound to a solid support.

18. A method of determining the identity of a nucleotide base at a specific position in a nucleic acid of interest, which comprises:

A) treating a sample containing the nucleic acid of interest, if the nucleic acid is double-stranded, so as to obtain unpaired nucleotide bases spanning the specific position, or directly employing step (B) if the nucleic acid of interest is single-stranded, wherein said nucleic acid of interest contains at least one 5' → 3' exonuclease resistant nucleotide derivative;

B) contacting the sample from step (A), under hybridizing conditions, with an oligonucleotide primer which is capable of hybridizing with a region of nucleotides present in the nucleic acid of interest immediately adjacent to the nucleotide base to be identified, so as to form a duplex between the primer

and the nucleic acid of interest such that the nucleotide base to be identified is the first unpaired base in the template immediately downstream of the 3' terminus of the primer in said duplex; and

C) contacting the duplex from step (B), in the substantial absence of dATP, dCTP, dGTP or dTTP, with at least two different nucleotide triphosphate derivatives, said derivatives including a derivative complimentary to said first unpaired base, and being terminators of a nucleic acid template-dependent, primer extension reaction; wherein at least one of said terminators is labeled with a detectable marker; and wherein said contacting is under conditions sufficient to permit base pairing of said complementary terminator derivative with said first unpaired base;

D) permitting occurrence of a template-dependent primer extension reaction sufficient to incorporate said complementary terminator derivative onto the 3' terminus of the primer;

E) determining the identity of said incorporated derivative, and thereby determining the identity of the nucleotide base at the specific position in the nucleic acid of interest.

19. The method of claim 18, wherein in step (C), the duplex from step (B) is contacted with four terminators, wherein only one of the terminators has a detectable marker, and wherein said step (C) is performed four times, each time with a different one of the terminators being labelled.

20. The method of claim 18, wherein in step (C), the duplex from step (B) is contacted with four labeled terminators, each with a different detectable label.

21. A method for detecting a desired exonuclease resistant amplification product of a polymerase chain reaction which comprises:

A) conducting a polymerase chain reaction with two primer molecules, wherein one of said primer molecules contains a sufficient number of phosphorothioate nucleotide derivatives at said primer's 5' terminus to render said terminus resistant to a 5'→3' exonuclease; said reaction being sufficient to form double-stranded amplification products;

B) subsequently treating said amplification products with a 5'→3' exonuclease under conditions to degrade oligonucleotides that lack a sufficient number of phosphorothioate bonds to render said oligonucleotides resistant to said exonuclease

C) detecting said desired amplification product of said polymerase chain reaction by permitting said product to hybridize to a complementary oligonucleotide bound to a solid support.

22. The method of claim 21, wherein said sufficent number of phosphorothioate nucleotide derivatives is about four.

23. A method for minimizing cross contamination between polymerase chain reactions which comprises conducting a polymerase chain reaction wherein at least one of the primer molecules of said reaction contains a sufficient number of phosphorothioate nucleotide derivatives at said primer's 3' terminus to render said terminus resistant to a 5'→3' exonuclease; and wherein, subsequent to conducting said polymerase chain reaction, amplification products of said reaction are incubated in the presence of said 5'→3' exonuclease under conditions sufficient to permit the degradation of oligonucleotide regions of unused primers and of amplification products that lack said sufficient number of phosphorothioate bonds; said degradation rendering said unused primer and said amplification products substantially incapable of serving as

substrates in an additional polymerase chain reaction, and thereby minimizing cross contamination between polymerase chain reactions.

24. The method of claim 23, wherein said sufficent number of phosphorothioate nucleotide derivatives is about four.

25. A kit, being specially adapted to contain in close compartmentalization a first container which contains a first primer, said first primer containing a phosphorothioate nucleotide derivative; and a second container which contains a second primer lacking any phosphorothioate nucleotide derivatives, such that the two primers can be used to amplify a predetermined gene sequence.

26. The kit of claim 25 which additionally includes a 5'→3' exonuclease.

INTERNATIONAL SEARCH REPORT

Int. ational application No.
PCT/US94/00771

A. CLASSIFICATION OF SUBJECT MATTER
IPC(5) :C12P 19/34
US CL : 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : C12Q 1/68; C07H 19/00, 17/00; A01N 43/04; C12N 15/00; 435/6, 91.1, 91.2; 536/24.3, 22.1; 514/44; 935/1, 17

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CHEMICAL ABSTRACTS; BIOSIS; MEDLINE; APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NUCLEIC ACIDS RESEARCH, Symposium Series No. 18, issued 1987, Gish and Eckstein, "DNA and RNA sequencing utilizing phosphorothioate chemistry", pages 253-256, see entire document.	1-26
Y	DNA, Volume 5, No. 2, issued 1986, Labeit *et alii*, "Laboratory Methods: A New Method of DNA Sequencing Using Deoxynucleoside α-Thiotriphosphates", pages 173-177, see entire document.	1-26
Y	US, A, 4,683,194, (SAIKI ET AL.) 28 JULY 1987, see entire document.	1-26
Y	US, A, 4,683,195 (MULLIS ET AL.) 28 JULY 1987, see entire document.	1-26

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"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
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"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
"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)	"&"	document member of the same patent family
"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		

Date of the actual completion of the international search
27 MARCH 1994

Date of mailing of the international search report
APR 12 1994

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
BRADLEY SISSON
Jill Warden for
Telephone No. (703) 305-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

IDT Exhibit 1020.2928

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BIOCONJUGATE CHEMISTRY, Volume 1, Number 3, issued May/June 1990, Goodchild, "Conjugates of Oligonucleotides and Modified Oligonucleotides: A Review of Thier Synthesis and Properties:, pages 165-187, see entire document.	1-26
A	BIOCHEMISTRY, Volume 15, No. 8, issued 1976, Eckstein and Goody, "Synthesis and Propertie of Diateriosomers of Adenosine 5'(*O*-1-Thiotriphosphate) and Adenosine 5'-(*O*-2-Thiotriphosophate)", paes 1685-1691, see entire document.	1-24
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, UNITED STATES OF AMERICA, Volume 78, Number 12, issued 1981, Putney et al., "A DNA fragment with an α-phosphorothioate nucleotide at one end is asymmetircally blocked from digesion by exonuclease III and can be replicated *in vivo*", pages 7350-7354, see entire document.	1, 4-20, 23-26
A	BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, Volume 179, Number 3, issued 30 September 1991, Kim et al., "Phosphorotioate Analogues of Oligodeoyribonucleotide: Synthesis and Activity as Inhibitors of Replication of Human Immunodeficiency Virus", pages 1614-1619, see entire document.	1-24
A	JOURNAL OF ORGANIC CHEMISTRY, Volume 54, issued 1989, Ludwig et al., "Rapid and Efficient Synthesis of Nucleoside 5'-*O*-(1-thiotriphosphates), 5'-Triphosphates and 2'3'-Cyclophosphorotioates Using 2-Chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one", pages 631-635, see entire document.	1-24
Y	NUCLEIC ACIDS RESEARCH, Volume 13, Number 24, issued 1985, Taylor et al., "The Rapid Generation of Oligonucleotide-directed Mutations at High Frequency Using Phosphorotioate-modified DNA", pages 8765-8785, see entire document.	1-26
A	NUCLEIC ACIDS RESEARCH, Volume 16, Number 3, issued 1988, Sayers et al., "5'-3' Exonucleases iin Phosphorotioatebased Oligonucleotide-Directed Mutagenesis", apges 791-802, see entire document.	1-24
Y	BIOCHEMISTRY, Volume 26, issued 1987, Ott et al., "Protection of Oligonucleotide Primers against Degradation by DNA Polymerase I", pages 8237-8241, see pages 8237, 8239-8341.	1-26

IDT Exhibit 1020.2929

INTERNATIONAL SEARCH REPORT

Int. ational application No.
PCT/US94/00771

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NUCLEIC ACIDS RESEARCH, Volume 14, Number 24, issued 1986, Nakamaye et al., "Inhibition of Restriction Endonuclease Nci I cleavage by Phosphorothioate groups and its application of Oligonucleotide-directed Mutagenesis", pages 9679-9698, see entire document.	1-24

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

Espacenet

Bibliographic data: WO2013059740 (A1) — 2013-04-25

NOVEL ALK AND NTRK1 FUSION MOLECULES AND USES THEREOF

Inventor(s): YELENSKY ROMAN [US]; LIPSON DORON [US]; CRONIN MAUREEN [US] ± (YELENSKY, ROMAN, ; LIPSON, DORON, ; CRONIN, MAUREEN)

Applicant(s): FOUNDATION MEDICINE INC [US] ± (FOUNDATION MEDICINE, INC)

Classification: **- international:** ***C07H21/04; C12P21/02***
- cooperative: **A61K31/4545**; **C07K14/4738**; **C07K14/4748**; **C07K16/40**; **C12N9/12**; **C12N9/1205**; **C12Q1/6886**; **G01N33/5011**; **G01N33/57496**; A61K38/00; C12Y207/10001

Application number: WO2012US61211 20121019 Global Dossier

Priority number (s): US201161550327P 20111021

Also published as: US2014336236 (A1)

Abstract of WO2013059740 (A1)

Novel ALK and NTRKI fusion molecules and uses are disclosed.

IDT Exhibit 1020.2931

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 April 2013 (25.04.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/059740 A1

(51) International Patent Classification:
C12P 21/02 (2006.01) *C07H 21/04* (2006.01)

(21) International Application Number:
PCT/US2012/061211

(22) International Filing Date:
19 October 2012 (19.10.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/550,327 21 October 2011 (21.10.2011) US

(71) Applicant: **FOUNDATION MEDICINE, INC.** [US/US]; One Kendall Square, Suite B6501, Cambridge, MA 02139 (US).

(72) Inventors: **YELENSKY, Roman**; 514 Lowell Avenue, Newton, MA 02460 (US). **LIPSON, Doron**; 142 Middlesex Road, #1, Chestnut Hill, MA 02467 (US). **CRONIN, Maureen**; 232 Beacon St., #6, Boston, MA 02116 (US).

(74) Agent: **COLLAZO, Diana M.**; Lando & Anastasi LLP, Riverfront Office Park, One Main Street, Suite 1100, Cambridge, MA 02142 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

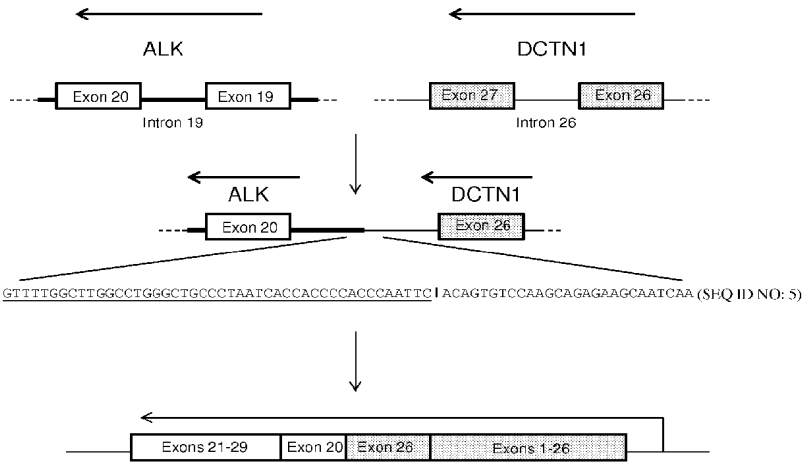
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

(54) Title: NOVEL ALK AND NTRK1 FUSION MOLECULES AND USES THEREOF

FIG. 1



(57) Abstract: Novel ALK and NTRK1 fusion molecules and uses are disclosed.

WO 2013/059740 A1

IDT Exhibit 1020.2932

NOVEL ALK AND NTRK1 FUSION MOLECULES AND USES THEREOF

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 61/550,327, filed October 21, 2011, the contents of which are incorporated herein by reference in their entirety.

BACKGROUND

Cancer represents the phenotypic end-point of multiple genetic lesions that endow cells with a full range of biological properties required for tumorigenesis. Indeed, a hallmark genomic feature of many cancers, including, for example, B cell cancer, lung cancer, breast cancer, ovarian cancer, pancreatic cancer, and colon cancer, is the presence of numerous complex chromosome structural aberrations, including translocations, intra-chromosomal inversions, point mutations, deletions, gene copy number changes, gene expression level changes, and germline mutations, among others.

The need still exists for identifying novel genetic lesions associated with cancer. Such genetic lesions can be an effective approach to develop compositions, methods and assays for evaluating and treating cancer patients.

SUMMARY

The invention is based, at least in part, on the discovery of novel translocation events that include a fragment of an ALK gene ("Anaplastic lymphoma kinase") and a fragment of a DCTN1 gene ("Dynactin-1") referred to herein as "DCTN1-ALK." The term "DCTN1-ALK" or "DCTN1-ALK fusion" is used generically herein, and includes any fusion molecule (*e.g.*, gene, gene product (*e.g.*, cDNA, mRNA, or polypeptide), and variant thereof) that includes a fragment of DCTN1 and a fragment of ALK, including, *e.g.*, a DCTN1-ALK. Expression of DCTN1-ALK was detected in a Spitz tumor, which is a family of neoplasms that can range from benign Spitz nevi to Spitzoid melanomas. Expression of DCTN1-ALK was also detected in a non-Langerhans cell histiocytosis.

The invention is also based, in part, on the discovery of a novel deletion resulting in the formation of an LMNA-NTRK1 fusion. The term "LMNA-NTRK1" or "LMNA-NTRK1 fusion" is used generically herein, and includes any fusion molecule (*e.g.*, gene,

IDT Exhibit 1020.2933

gene product (*e.g.*, cDNA, mRNA, or polypeptide), and variant thereof) that includes a fragment of NTRK1 ("neurotrophic tyrosine kinase receptor, type 1") and a fragment of LMNA ("lamin A/C"), including, *e.g.*, a LMNA-NTRK1. Expression of LMNA-NTRK1 was detected in Spitz tumors.

ALK and NTRK1 receptor tyrosine kinase domains have been associated with cancerous phenotypes, including lung, thyroid, non-Hodgkin's lymphoma, neuroblastoma, among others. Expression of the fusion molecules described herein in Spitz tumors and histiocytosis suggests a further association with neoplastic growth (including benign, pre-malignant, or malignant and/ or metastatic growth (*e.g.*, a cancer), including, for example, melanocytic neoplasms, melanoma and metastatic disease. Accordingly, the invention provides methods of: identifying, assessing or detecting a DCTN1-ALK and/or LMNA-NTRK1 fusion; methods of identifying, assessing, evaluating, and/or treating a subject having a cancer, *e.g.*, a cancer having a DCTN1-ALK and/or LMNA-NTRK1 fusion; isolated DCTN1-ALK and/or LMNA-NTRK1 nucleic acid molecules, nucleic acid constructs, host cells containing the nucleic acid molecules; purified DCTN1-ALK and/or LMNA-NTRK1 polypeptides and binding agents; detection reagents (*e.g.*, probes, primers, antibodies, kits, capable, *e.g.*, of specific detection of a DCTN1-ALK and/or LMNA-NTRK1 nucleic acid or protein); screening assays for identifying molecules that interact with, *e.g.*, inhibit, 5'DCTN1-3'ALK and/or 5'LMNA-3'NTRK1 fusions, *e.g.*, novel kinase inhibitors; as well as assays and kits for evaluating, identifying, assessing and/or treating a subject having a cancer, *e.g.*, a cancer having a DCTN1-ALK and/or LMNA-NTRK1 fusion. The compositions and methods identified herein can be used, for example, to identify new DCTN1-ALK and/or LMNA-NTRK1 inhibitors; to evaluate, identify or select a subject, *e.g.*, a patient, having a cancer; and to treat or prevent a cancer.

Each of these fusion molecules is described herein in more detail.

DCTN1-ALK fusions: In one embodiment, a DCTN1-ALK fusion includes an in-frame fusion of an exon of DCTN1 (*e.g.*, one more exons of DCTN1, such as exons 1-26 or a fragment thereof) and an exon of ALK (*e.g.*, one or more exons encoding the ALK receptor tyrosine kinase or a fragment thereof). For example, the DCTN1-ALK fusion can occur in chromosome 2 and include an in-frame fusion within an intron of DCTN (*e.g.*, intron 26) or a fragment thereof, with an intron of

IDT Exhibit 1020.2934

ALK (*e.g.,* intron 19) or a fragment thereof. In one embodiment, the fusion junction of the DCTN1-ALK fusion comprises a nucleotide sequence of: GTTTTGGCTTGGCCTGGGCTGCCCTAATCACCACCCCACCCAATTCACAGT GTCCAAGCAGAGAAGCAATCAA (SEQ ID NO: 5), or a fragment thereof. The sequence of SEQ ID NO:5 is the reference genome orientation.

In certain embodiments, the DCTN1-ALK fusion is in a 5'- DCTN1 to 3'-ALK configuration referred to herein as "5' DCTN1-ALK." A DCTN1-ALK fusion polypeptide encoded by a 5'DCTN1-3'ALK nucleic acid is sometimes referred to herein as a 5'DCTN1-3'ALK polypeptide. In an embodiment, the 5'DCTN1-3'ALK fusion comprises sufficient DCTN1 and sufficient ALK sequence such that the 5'DCTN1-3'ALK fusion has kinase activity, *e.g.,* has elevated activity, *e.g.*, ALK tyrosine kinase activity, as compared with wild type ALK, *e.g.*, in a cell of a cancer referred to herein (*e.g.,* melanoma). In one embodiment, the 5'DCTN1-3'ALK fusion comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or more exons from DCTN1 and at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or more, ALK exons. In one embodiment, the 5'DCTN1-3'ALK fusion polypeptide includes an ALK receptor tyrosine kinase domain or a functional fragment thereof. In one embodiment, the DCTN1-ALK fusion comprises a nucleotide sequence shown in FIG. 2 (SEQ ID NO: 6), or a fragment thereof. In another embodiment, the DCTN1-ALK fusion comprises an amino acid sequence shown in FIG. 2 (SEQ ID NO:7).

The ALK receptor tyrosine kinase is known to be associated with cancerous phenotypes including inflammatory myofibroblastic tumors, neuroblastoma, lung cancer, non-Hodgkin's lymphoma, and anaplastic large cell lymphoma, among others. For example, a chromosomal rearrangement that generates a fusion gene resulting in the juxtaposition of the N-terminal region of nucleophosmin (NPM) with the kinase domain of ALK are known to be associated with non-Hodgkin's lymphoma (Morris, SW (1994) *Science* 263:1281-1284).

In other embodiments, the DCTN1-ALK fusion includes an in-frame fusion having a breakpoint in an intron of DCTN1 (*e.g.,* intron 26) or a fragment thereof and a breakpoint in an intron of ALK (*e.g.,* intron 19) or a fragment thereof. In certain embodiments, the DCTN1-ALK fusion is in a 5'- ALK to 3'-DCTN1 configuration referred to herein as "5'ALK-3'DCTN1"). For example, the ALK-DCTN1 fusion can occur in chromosome 2 and include an in-frame fusion within an intron of DCTN (*e.g.,*

IDT Exhibit 1020.2935

intron 26) or a fragment thereof, with an intron of ALK (*e.g.,* intron 19) or a fragment thereof. In one embodiment, the fusion junction of the ALK-DCTN1 fusion comprises a nucleotide sequence of: TGAAGCCACATGAACTCAGTGAGAAAA| CAGGCACCTGTGGCACAGCCTGAGACACTATTCAGTCCTGCCTTCCTGC (SEQ ID NO: 13), or a fragment thereof. The sequence of SEQ ID NO:5 is the reference genome orientation.

An ALK-DCTN1 fusion polypeptide encoded by a 5'ALK-3'DCTN1 nucleic acid is sometimes referred to herein as a 5'ALK-3'DCTN1 polypeptide. In one embodiment, the 5'ALK-3'DCTN1 fusion comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or more exons from ALK and at least 1, 2, 3, 4, 5, or 6, or more, DCTN1 exons. In one embodiment, the ALK-DCTN1 fusion comprises a nucleotide sequence shown in FIGs. 4A-4D (SEQ ID NO:11), or a fragment thereof. In another embodiment, the ALK-DCTN1 fusion comprises an amino acid sequence shown in FIGs. 4A-4D (SEQ ID NO:12).

LMNA-NTRK1 fusion: Certain aspects featured in the invention include LMNA-NTRK1 fusion nucleic acids and polypeptides. An LMNA-NTRK1 fusion includes an in-frame fusion of at least exon 2 of LMNA (*e.g.,* one or more of exons 1-2 or a fragment thereof) and exon 11 of NTRK1 (*e.g.,* one more exons encoding a tyrosine kinase domain or a fragment thereof) (FIG. 6). For example, the LMNA-NTRK1 fusion can include an in-frame fusion within an intron of NTRK1 (*e.g.,* intron 10) or a fragment thereof, with an intron of LMNA (*e.g.,* intron 2) or a fragment thereof. In certain embodiments, the LMNA-NTRK1 fusion is in a 5'- LMNA to 3'-NTRK1 configuration referred to herein as "5' LMNA-3'NTRK1." In one embodiment, the LMNA-NTRK1 fusion has a breakpoint comprising the nucleotide sequence of: ATAAAAATTTAAAGAAATTAGCTGGGCATAGGGGTCCCCAGGGGAGGATGA GGCAGGTCTGGAGACCT (SEQ ID NO:8), or a fragment thereof (FIG. 6). The sequence of SEQ ID NO:8 is presented in the reference genome orientation.

In one embodiment, the LMNA-NTRK1 fusion has a nucleotide sequence as shown in FIGs. 7A-7B (SEQ ID NO: 9), or a fragment thereof. In another embodiment, the LMNA-NTRK1 fusion has an amino acid sequence as shown in FIGs. 7A-7B (SEQ ID NO:10), or a fragment thereof.

IDT Exhibit 1020.2936

An LMNA-NTRK1 fusion polypeptide encoded by a 5'LMNA-3'NTRK nucleic acid is sometimes referred to herein as a 5'LMNA-3'NTRK1 polypeptide. In an embodiment, the 5'LMNA-3'NTRK1 fusion comprises sufficient LMNA and sufficient NTRK1 sequence such that the 5'LMNA-3'NTRK1 fusion has kinase activity, *e.g.*, has elevated activity, *e.g.*, receptor tyrosine kinase activity, as compared with wild type NTRK1, *e.g.*, in a cell of a cancer referred to herein (*e.g.*, melanoma). In one embodiment, the 5'LMNA-3'NTRK1 fusion comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or more exons from LMNA and at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or more, NTRK1 exons. In one embodiment, the 5'LMNA-3'NTRK1 fusion polypeptide includes *e.g.*, an NTRK1 receptor tyrosine kinase domain or a functional fragment thereof.

The NTRK1 receptor tyrosine kinase (also referred to as the TrkA receptor) is known to be associated with cancerous phenotypes including colon tumors and papillary thyroid carcinomas, among others. For example, a chromosomal rearrangement that generates a fusion gene resulting in the juxtaposition of the TFG gene with the kinase binding domain of NTRK1 are known to be associated with papillary thyroid carcinomas. Other NTRK1 rearrangements resulting in fusions include, *e.g.*, NTRK1-TPM3 and TPR-NTRK1. NTRK1 fusion genes are reviewed in, *e.g.*, Greco, A. *et al.*, *Mol Cell Endocrinol* (2010) 321(1):44-49.

In other embodiments, the LMNA-NTRK1 fusion includes an in-frame fusion having a breakpoint in an intron of LMNA (*e.g.*, intron 2) or a fragment thereof and a breakpoint in an intron of NTRK1 (*e.g.*, intron 10) or a fragment thereof. In another embodiment, the LMNA-NTRK1 is an in-frame fusion comprising at least exon 2 or LMNA and at least exon 11 of NTRK1. In certain embodiments, the LMNA-NTRK1 fusion is in a 5'- LMNA to 3'-NTRK1 configuration referred to herein as "5'LMNA-3'NTRK1").

Accordingly, the invention provides, methods of: identifying, assessing or detecting a LMNA-NTRK1 fusion; isolated LMNA-NTRK1 nucleic acid molecules, nucleic acid constructs, host cells containing the nucleic acid molecules; purified LMNA-NTRK1 polypeptides and binding agents; detection reagents (*e.g.*, probes, primers, antibodies, kits, capable, *e.g.*, of specific detection of a LMNA-NTRK1 nucleic acid or protein); screening assays for identifying molecules that interact with, e.g., inhibit, LMNA-NTRK1 fusions, *e.g.*, novel kinase inhibitors.

IDT Exhibit 1020.2937

Nucleic Acid Molecules

In one aspect, the invention features an isolated nucleic acid molecule, or an isolated preparation of nucleic acid molecules, that includes a genetic alteration disclosed herein. Such nucleic acid molecules or preparations thereof can include a genetic alteration described herein or can be used to detect, *e.g.,* sequence, a genetic alteration disclosed herein.

In certain embodiments, the alteration in the nucleic acid molecule, isolated preparation, or tumor member, is chosen from one or more of:

(i) a rearrangement (*e.g.,* translocation) of a fragment of a Dynactin-1 gene (DCTN1) and a fragment of an Anaplastic Lymphoma receptor tyrosine Kinase (ALK) gene (a "DCTN1-ALK fusion"); and

(ii) a rearrangement (*e.g.,* deletion) of a fragment of a lamin A/C gene (LMNA) and a fragment of an NRTK1 (neurotrophic tyrosine kinase receptor, type 1) gene ("a "LMNA-NRTK1 fusion").

DCTN1-ALK nucleic acid molecules

In one aspect, the invention features a nucleic acid molecule (*e.g.*, an isolated or purified) nucleic acid molecule that includes a fragment of a DCTN1 gene and a fragment of an ALK receptor tyrosine kinase. In one embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, between an intron of DCTN1 (*e.g.*, intron 26, or a fragment thereof), and an intron of ALK (*e.g.*, intron 19, or a fragment thereof).

In another embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, of at least exon 26 of DCTN1 or a fragment thereof (*e.g.*, exons 1-26 of DCTN1 or a fragment thereof), and at least exon 20 or a fragment thereof (*e.g.*, exons 20-29 of ALK or a fragment thereof). In one embodiment, the nucleic acid molecule includes the nucleotides sequence of 1-3195 of SEQ ID NO:6 (corresponding to exons 1-26 of a DCTN1 gene) or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of 3196-4907 of SEQ ID NO:6 (corresponding to exons 20-29 of the ALK gene) or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in

IDT Exhibit 1020.2938

FIGs. 2A-2F (*e.g.*, SEQ ID NO:6) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:6 or a fragment thereof. In yet another embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition to a nucleotide sequence complementary to SEQ ID NO:6 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary 5'DCTN1-3'ALK fusion is shown in SEQ ID NO:6, and the predicted amino acid sequence is shown in SEQ ID ID NO:7.

In an embodiment the 5'DCTN1-3'ALK nucleic acid molecule comprises sufficient DCTN1 and sufficient ALK sequence such that the encoded 5'DCTN1-3'ALK fusion has kinase activity, *e.g.*, has elevated activity, *e.g.*, ALK kinase activity, as compared with wild type ALK, e.g., in a cell of a cancer referred to herein. In an embodiment the encoded 5'DCTN-3'ALK fusion comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or 11 exons from DCTN1 and at least 1, 2, 3, 4, 5, 6, 7, 9, or 10, ALK exons. In one embodiment, the encoded 5'DCTN1-3'ALK fusion polypeptide includes an ALK tyrosine kinase domain or a functional fragment thereof.

In one embodiment, the nucleic acid molecule includes a nucleotide sequence that has an in-frame fusion of intron 26 of DCTN1 (*e.g.*, NM_004082) with intron 19 of ALK (*e.g.*, NM_004304). In other embodiments, the nucleic acid molecules includes a nucleotide sequence in the region of nucleotides 29,447,851-29,448,653 of chromosome 2 coupled to (*e.g.*, juxtaposed to) nucleotides in the region of nucleotides 74,591,512-74,592,314 of chromosome 2 (FIG. 5). In another embodiment, the nucleic acid molecule includes a nucleotide sequence that includes a breakpoint, *e.g.*, a breakpoint identified in FIG. 1. For example, the nucleic acid molecule includes a nucleotide sequence that includes the fusion junction between the DCTN1 gene and the ALK gene, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:5 (*e.g.*, the breakpoint between intron 26 of DCTN1 and intron 19 of ALK).

In other embodiments, the nucleic acid molecule includes a DCTN1-ALK fusion having a configuration shown in FIG. 1. For example, the DCTN1-ALK fusion can include an in-frame fusion resulting from a break-point of at least intron 26 of DCTN1 or

IDT Exhibit 1020.2939

a fragment thereof with at least intron 19 of ALK or a fragment thereof. In certain embodiments, the DCTN1-ALK fusion is in a 5'-DCTN1 to 3'-ALK configuration referred to herein as "5'DCTN1-3'ALK"). In one embodiment, the nucleic acid molecule includes the nucleotide sequence of SEQ ID NO:5 (corresponding to the breakpoint of a DCTN1-ALK fusion), or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of nucleotides 1-27 of SEQ ID NO:5 (*e.g.*, corresponding to intron 26 of DCTN1), or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of nucleotides 28-73 of SEQ ID NO:5 (*e.g.*, corresponding to intron 19 of ALK), or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIGs. 2A-2F (*e.g.*, SEQ ID NO:6) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:6 or a fragment thereof. In yet other embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to a nucleotide sequence complementary to SEQ ID NO:6 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary breakpoint for a 5'DCTN1-3'ALK fusion is shown in SEQ ID NO:6.

In other embodiments, the nucleic acid molecule includes a nucleotide sequence encoding a DCTN1-ALK fusion polypeptide that includes a fragment of a DCTN1 gene and a fragment of an ALK receptor tyrosine kinase. In one embodiment, the nucleotide sequence encodes a DCTN1-ALK fusion polypeptide that includes *e.g.,* an ALK tyrosine kinase domain or a functional fragment thereof. In another embodiment, the nucleotide sequence encodes a fragment of the DCTN1 polypeptide including the amino acid sequence of amino acids 1062-1069 of SEQ ID NO:7 or a fragment thereof, or a sequence substantially identical thereto. For example, the nucleic acid molecule can include a nucleotide sequence encoding an ALK kinase domain of a DCTN1-ALK fusion polypeptide that includes amino acids 1124-1370 of SEQ ID NO:7 or a fragment thereof. In yet other embodiments, the nucleic acid molecule includes a nucleotide sequence

IDT Exhibit 1020.2940

encoding the amino acid sequence shown in FIGs. 2A-2F (*e.g.*, SEQ ID NO:7) or a fragment thereof, or a sequence substantially identical thereto.

In another embodiment, the nucleic acid molecule includes a DCTN1-ALK fusion having the configuration shown in FIG. 1. In one embodiment, the nucleic acid molecule includes a nucleotide sequence that includes a fusion junction between the DCTN1 transcript and the ALK transcript, *e.g.*, a nucleotide sequence within SEQ ID NO:6.

In another embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, of at least exon 19 of ALK or a fragment thereof (*e.g.*, exons 1-19 of ALK or a fragment thereof), and at least exon 27 or a fragment thereof (*e.g.*, exons 27-32 of DCTN1 or a fragment thereof). In certain embodiments, the DCTN1-ALK fusion is in a 5'- ALK to 3'- DCTN1 configuration referred to herein as "5'ALK-3'DCTN1"). In one embodiment, the nucleic acid molecule includes the nucleotides sequence of 1-3172 of SEQ ID NO:11 (corresponding to exons 1-19 of an ALK gene) or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotides sequence of 3173-3831 of SEQ ID NO:11 (corresponding to exons 27-32 of the a DCTN1 gene) or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIGs. 4A-4D (*e.g.*, SEQ ID NO:11) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:11 or a fragment thereof. In yet another embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition to a nucleotide sequence complementary to SEQ ID NO:11 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary 5'ALK-3'DCTN1 fusion is shown in SEQ ID NO:11, and the predicted amino acid sequence is shown in SEQ ID ID NO:12.

In a related aspect, the invention features nucleic acid constructs that include the DCTN1-ALK nucleic acid molecules described herein. In certain embodiments, the nucleic acid molecules are operatively linked to a native or a heterologous regulatory sequence. Also included are vectors and host cells that include the DCTN1-ALK nucleic

IDT Exhibit 1020.2941

acid molecules described herein, *e.g.*, vectors and host cells suitable for producing the nucleic acid molecules and polypeptides described herein.

In a related aspect, methods of producing the nucleic acid molecules and polypeptides described herein are also described.

In another aspect, the invention features nucleic acid molecules that reduce or inhibit the expression of a nucleic acid molecule that encodes a DCTN1-ALK fusion described herein. Examples of such nucleic acid molecules include, for example, antisense molecules, ribozymes, RNAi, triple helix molecules that hybridize to a nucleic acid encoding DCTN1-ALK, or a transcription regulatory region of DCTN1-ALK, and blocks or reduces mRNA expression of DCTN1-ALK.

LMNA-NTRK1 nucleic acid molecules

In one aspect, the invention features a nucleic acid molecule (*e.g.*, an isolated or purified) nucleic acid molecule that includes a fragment of a LMNA gene and a fragment of an NTRK1 receptor tyrosine kinase. In one embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, of an intron of LMNA (*e.g.*, expressing one more exons of LMNA, *e.g.,* at least exon 2, or a fragment thereof), and an intron of NTRK1 (*e.g.*, expressing one or more exons encoding an NRTK1 tyrosine kinase domain, *e.g.,* at least exon 11 or a fragment thereof).

In one embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, of at least exon 2 of LMNA or a fragment thereof (*e.g.*, exons 1 and 2 of LMNA or a fragment thereof), and at least exon 11 or a fragment thereof (*e.g.*, exons 11-19 of NTRK1 or a fragment thereof). In one embodiment, the nucleic acid molecule includes the nucleotides sequence of 1-513 of SEQ ID NO:9 (corresponding to exons 1-2 of an LMNA gene) or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of 514-1740 of SEQ ID NO:9 (corresponding to exons 11-19 of the an NTRK1 gene) or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIGs. 7A-7B (*e.g.*, SEQ ID NO:9) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency

IDT Exhibit 1020.2942

condition described herein to SEQ ID NO:9 or a fragment thereof. In yet another embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition to a nucleotide sequence complementary to SEQ ID NO:9 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary 5'LMNA-3'NTRK1 fusion is shown in SEQ ID NO:9, and the predicted amino acid sequence is shown in SEQ ID ID NO:10.

In an embodiment the 5'LMNA-3'NTRK1 nucleic acid molecule comprises sufficient LMNA and sufficient NTRK1 sequence such that the encoded 5'LMNA-3'NTRK1 fusion has kinase activity, *e.g.*, has elevated activity, *e.g.*, NTRK1 kinase activity, as compared with wild type NTRK1, e.g., in a cell of a cancer referred to herein. In an embodiment the encoded 5'LMNA-3'NTRK1 fusion comprises at least 1 or 2 exons or more from LMNA and at least 1, 2, 3, 4, 5, 6, 7, or 9 NTRK1 exons. In one embodiment, the encoded 5'LMNA-3'NTRK1 fusion polypeptide includes an NTRK1 tyrosine kinase domain or a functional fragment thereof.

In one embodiment, the nucleic acid molecule includes a nucleotide sequence that has an in-frame fusion of intron 2 of LMNA (*e.g.*, NM_005572) with intron 10 of NTRK1 (*e.g.*, NM_002529). In other embodiments, the nucleic acid molecules includes a nucleotide sequence in the region of 156,100,970-156,101,459 of chromosome 1 coupled to (*e.g.*, juxtaposed to) nucleotides in the region of nucleotides 156,844,390-156,844,879 of chromosome 1 (FIG. 8). In another embodiment, the nucleic acid molecule includes a nucleotide sequence that includes a breakpoint, *e.g.*, a breakpoint identified in FIG. 6. For example, the nucleic acid molecule includes a nucleotide sequence that includes the fusion junction between the LMNA gene and the NTRK1 gene, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:8 (*e.g.*, the breakpoint between intron 2 of LMNA and intron 10 of NTRK1).

In other embodiments, the nucleic acid molecule includes a LMNA-NTRK1 fusion having a configuration shown in FIG. 6. For example, the LMNA-NTRK1 fusion can include an in-frame fusion of at least intron 2 of LMNA or a fragment thereof with at least intron 10 of NTRK1 or a fragment thereof. In one embodiment, the LMNA-NTRK2 fusion comprises a fusion of at least exon 2 of LMNA and at least exon 11 of NTRK1. In certain embodiments, the LMNA-NTRK1 fusion is in a 5'-LMNA to 3'-NTRK1

IDT Exhibit 1020.2943

configuration referred to herein as "5'LMNA-3'NTRK1"). In one embodiment, the nucleic acid molecule includes the nucleotide sequence of SEQ ID NO:8 (corresponding to the breakpoint of a LMNA-NTRK1 fusion), or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of nucleotides 1-29 of SEQ ID NO:8 (*e.g.*, corresponding to intron 2 of LMNA), or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of nucleotides 30-68 of SEQ ID NO:8 (*e.g.*, corresponding to intron 10 of NTRK1), or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIG 6 (*e.g.*, SEQ ID NO:8) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:8 or a fragment thereof. In yet other embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to a nucleotide sequence complementary to SEQ ID NO:9 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary breakpoint for a 5'LMNA-3'NTRK1 fusion is shown in SEQ ID NO:9.

In other embodiments, the nucleic acid molecule includes a nucleotide sequence encoding a LMNA-NTRK1 fusion polypeptide that includes a fragment of a LMNA gene and a fragment of an NTRK1 receptor tyrosine kinase. In one embodiment, the nucleotide sequence encodes a LMNA-NTRK1 fusion polypeptide that includes *e.g.*, an NTRK1 tyrosine kinase domain or a functional fragment thereof. In another embodiment, the nucleotide sequence encodes a fragment of the LMNA polypeptide including the amino acid sequence of amino acids 168-175 of SEQ ID NO:10 or a fragment thereof, or a sequence substantially identical thereto. For example, the nucleic acid molecule can include a nucleotide sequence encoding a LMNA kinase domain of a LMNA-NTRK1 fusion polypeptide that includes amino acids 289-560 of SEQ ID NO:10 or a fragment thereof. In yet other embodiments, the nucleic acid molecule includes a

IDT Exhibit 1020.2944

nucleotide sequence encoding the amino acid sequence shown in FIG. 4 (*e.g.*, SEQ ID NO:10) or a fragment thereof, or a sequence substantially identical thereto.

In another embodiment, the nucleic acid molecule includes the nucleotides sequence of 1-513 of SEQ ID NO:9 (corresponding to exons 1-2 of an LMNA gene) or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotide sequence of 514-1740 of SEQ ID NO:9 (corresponding to exons 11-19 of an NTRK1 gene) or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIGs. 7A-7B (*e.g.*, SEQ ID NO:9) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:9 or a fragment thereof. In yet other embodiments, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition to a nucleotide sequence complementary to SEQ ID NO:9 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary 5'LMNA-3'NTRK1 fusion is shown in SEQ ID NO:9, and the predicted amino acid sequence is shown in SEQ ID ID NO:10.

In a related aspect, the invention features nucleic acid constructs that include the LMNA-NTRK1 nucleic acid molecules described herein. In certain embodiments, the nucleic acid molecules are operatively linked to a native or a heterologous regulatory sequence. Also included are vectors and host cells that include the LMNA-NTRK1 nucleic acid molecules described herein, *e.g.*, vectors and host cells suitable for producing the nucleic acid molecules and polypeptides described herein.

In a related aspect, methods of producing the nucleic acid molecules and polypeptides described herein are also described.

In another aspect, the invention features nucleic acid molecules that reduce or inhibit the expression of a nucleic acid molecule that encodes a LMNA-NTRK1 fusion described herein. Examples of such nucleic acid molecules include, for example, antisense molecules, ribozymes, RNAi, triple helix molecules that hybridize to a nucleic

IDT Exhibit 1020.2945

acid encoding LMNA-NTRK1, or a transcription regulatory region of LMNA-NTRK1, and blocks or reduces mRNA expression of LMNA-NTRK1.

Nucleic Acid Detection and Capturing Reagents

The invention also features a nucleic acid molecule, *e.g.*, nucleic acid fragment, suitable as probe, primer, bait or library member that includes, flanks, hybridizes to, which are useful for identifying, or are otherwise based on, the DCTN1-ALK or LMNA-NTRK1 fusions described herein. In certain embodiments, the probe, primer or bait molecule is an oligonucleotide that allows capture, detection or isolation of a DCTN1-ALK or an LMNA-NTRK1 fusion nucleic acid molecule described herein. The oligonucleotide can comprise a nucleotide sequence substantially complementary to a fragment of the DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecules described herein. The sequence identity between the nucleic acid fragment, *e.g.*, the oligonucleotide, and the target DCTN1-ALK or LMNA-NTRK1 sequence need not be exact, so long as the sequences are sufficiently complementary to allow the capture, detection or isolation of the target sequence. In one embodiment, the nucleic acid fragment is a probe or primer that includes an oligonucleotide between about 5 and 25, *e.g.*, between 10 and 20, or 10 and 15 nucleotides in length. In other embodiments, the nucleic acid fragment is a bait that includes an oligonucleotide between about 100 to 300 nucleotides, 130 and 230 nucleotides, or 150 and 200 nucleotides, in length.

In one embodiment, the nucleic acid fragment can be used to identify or capture, *e.g.*, by hybridization, a DCTN1-ALK fusion or a LMNA-NTRK1 fusion. For example, the nucleic acid fragment can be a probe, a primer, or a bait, for use in identifying or capturing, *e.g.*, by hybridization, a DCTN1-ALK fusion or an LMNA-NTRK1 fusion described herein. In one embodiment, the nucleic acid fragment can be useful for identifying or capturing a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g,*. SEQ ID NO:5). In one embodiment, the nucleic acid fragment hybridizes to a nucleotide sequence within a chromosomal rearrangement that creates an in-frame fusion of intron 26 of DCTN1 with intron 19 of ALK. In one embodiment, the nucleic acid fragment hybridizes to a nucleotide sequence in the region of 29,447,851-29,448,653 of chromosome 2 coupled to (*e.g.,* juxtaposed to) nucleotides in the region of nucleotides 74,591,512-74,592,314 of chromosome 2. In one embodiment, the nucleic acid fragment

IDT Exhibit 1020.2946

hybridizes to a nucleotide sequence that includes a breakpoint, *e.g.*, a breakpoint as identified in FIG. 1. For example, the nucleic acid fragment can hybridize to a nucleotide sequence that includes the fusion junction between the DCTN1 gene and the ALK gene, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:6 (*e.g.*, a nucleotide sequence within introns 26 of a DCTN1 gene and 19 of an ALK gene).

In another embodiment, the nucleic acid fragment can be useful for identifying or capturing a LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g,*. SEQ ID NO:8). In one embodiment, the nucleic acid fragment hybridizes to a nucleotide sequence within a chromosomal rearrangement that creates an in-frame fusion of intron 2 of LMNA with intron 10 of NTRK1. In other embodiments, the nucleic acid molecules includes a nucleotide sequence in the region of 156,100,970-156,101,459 of chromosome 1 coupled to (*e.g.,* juxtaposed to) nucleotides in the region of nucleotides 156,844,390-156,844,879 of chromosome 1. In one embodiment, the nucleic acid fragment hybridizes to a nucleotide sequence that includes a breakpoint, *e.g.*, a breakpoint as identified in FIG. 6. For example, the nucleic acid fragment can hybridize to a nucleotide sequence that includes the fusion junction between the LMNA transcript and the NRTK1 transcript, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:8 (*e.g.*, a nucleotide sequence within introns 2 of a LMNA gene and 10 of an NRTK1 gene).

The probes or primers described herein can be used, for example, for FISH detection or PCR amplification. In one exemplary embodiment where detection is based on PCR, amplification of the DCTN1-ALK fusion junction or the LMNA-NTRK1 fusion junction can be performed using a primer or a primer pair, *e.g.*, for amplifying a sequence flanking the fusion junctions described herein, *e.g.*, the mutations or the junction of a chromosomal rearrangement described herein, *e.g.,* DCTN1-ALK or LMNA-NTRK1.

In one embodiment, a pair of isolated oligonucleotide primers can amplify a region containing or adjacent to a position in the ALK-DCTN1 fusion or the NTRK1-LMNA fusion. For example, reverse primers can be designed to hybridize to a nucleotide sequence within ALK genomic or mRNA sequence (*e.g.*, a nucleotide sequence within exon 20 of ALK, or to a sequence corresponding to nucleotides 3196-3231 of SEQ ID NO:6), and the forward primers can be designed to hybridize to a nucleotide sequence within DCTN1 (*e.g.*, a nucleotide sequence within exon 26 of DCTN1, or a sequence corresponding to nucleotides 3163-3195 of SEQ ID NO:6). Also,

IDT Exhibit 1020.2947

forward primers can be designed to hybridize to a nucleotide sequence within LMNA genomic or mRNA sequence (*e.g.*, a nucleotide sequence within exon 2 of LMNA or a sequence corresponding to nucleotides 484-514 of SEQ ID NO:9), and the reverse primers can be designed to hybridize to a nucleotide sequence of NTRK1 (*e.g.*, a nucleotide sequence within exon 11 of NTRK1, or nucleotides 514-540 of SEQ ID NO:9).

In another embodiment, the nucleic acid fragments can be used to identify, *e.g.*, by hybridization, a 5'DCTN1-3'ALK fusion or the 5'LMNA-3'NTRK1 fusion. In one embodiment, the nucleic acid fragment hybridizes to a nucleotide sequence that includes a fusion junction between the DCTN1 transcript and the ALK transcript or a fusion junction between the LMNA transcript and the NTRK1 transcript.

In other embodiments, the nucleic acid fragment includes a bait that comprises a nucleotide sequence that hybridizes to a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule described herein, and thereby allows the capture or isolation said nucleic acid molecule. In one embodiment, a bait is suitable for solution phase hybridization. In other embodiments, a bait includes a binding entity, *e.g.*, an affinity tag, that allows capture and separation, *e.g.*, by binding to a binding entity, of a hybrid formed by a bait and a nucleic acid hybridized to the bait.

In other embodiments, the nucleic acid fragment includes a library member comprising a DCTN1-ALK or LMNA-NTRK1 nucleic acid molecule described herein. In one embodiment, the library member includes a rearrangement that results in a DCTN1-ALK or LMNA-NTRK1 fusion described herein.

The nucleic acid fragment can be detectably labeled with, *e.g.,* a radiolabel, a fluorescent label, a bioluminescent label, a chemiluminescent label, an enzyme label, a binding pair label, or can include an affinity tag; a tag, or identifier (*e.g.*, an adaptor, barcode or other sequence identifier).

DCTN1-ALK Fusion Polypeptides

In another aspect, the invention features a DCTN1-ALK fusion polypeptide (*e.g.*, a purified DCTN1-ALK fusion polypeptide), a biologically active or antigenic fragment thereof, as well as reagents (*e.g.*, antibody molecules that bind to a

IDT Exhibit 1020.2948

DCTN1-ALK fusion polypeptide), methods for modulating a DCTN1-ALK polypeptide activity and detection of a DCTN1-ALK polypeptide.

In one embodiment, the DCTN1-ALK fusion polypeptide has at least one biological activity, *e.g.*, an ALK kinase activity, dynein binding activity, kinesin binding activity, and/or a dimerizing or multimerizing activity. In one embodiment, at least one biological activity of the DCTN1-ALK fusion polypeptide is reduced or inhibited by an anti-cancer drug, *e.g.*, a kinase inhibitor (*e.g.*, a multikinase inhibitor or an ALK-specific inhibitor). In one embodiment, at least one biological activity of the DCTN1-ALK fusion polypeptide is reduced or inhibited by an ALK kinase inhibitor chosen from *e.g.,* TAE-684 (also referred to herein as "NVP-TAE694"), PF02341066 (also referred to herein as "crizotinib" or "1066"), AF-802, LDK-378, ASP-3026, CEP-37440, CEP-28122, CEP-18050 and AP26113.

In other embodiments, the nucleic acid molecule includes a nucleotide sequence encoding a DCTN-ALK fusion polypeptide that includes a fragment of a DCNT1 gene and a fragment of an ALK gene. In one embodiment, the nucleotide sequence encodes a DCTN1-ALK fusion polypeptide that includes a dynein associated domain or a functional fragment thereof, and an ALK tyrosine kinase domain or a functional fragment thereof. In another embodiment, the nucleotide sequence encodes a fragment of the DCTN1 polypeptide including the amino acid sequence of amino acids 1-1065 of SEQ ID NO:7 or a fragment thereof, or a sequence substantially identical thereto. For example, the nucleic acid molecule can include a nucleotide sequence encoding a dynein association domain of a DCTN1-ALK fusion polypeptide that includes amino acids 526-805 of SEQ ID NO:7 or a fragment thereof. In other embodiments, the nucleic acid molecule includes a fragment of the ALK gene encoding the amino acid sequence of amino acids 1066-1640 of SEQ ID NO:7 or a fragment thereof, or a sequence substantially identical thereto. For example, the nucleic acid molecule can include a nucleotide sequence encoding an ALK kinase domain of a DCTN1-ALK fusion polypeptide that includes amino acids 1116-1362 of SEQ ID NO:7 or a fragment thereof. In yet other embodiments, the nucleic acid molecule includes a nucleotide sequence encoding the amino acid sequence shown in FIGs. 2A-2F (*e.g.*, SEQ ID NO:7) or a fragment thereof, or a sequence substantially identical thereto.

IDT Exhibit 1020.2949

In another embodiment, the nucleic acid molecule includes a DCTN1-ALK fusion having the configuration shown in FIGs. 1 and 2A-2F. In one embodiment, the nucleic acid molecule includes a nucleotide sequence that includes a fusion junction between the DCTN1 transcript and the ALK transcript, *e.g.*, a nucleotide sequence within SEQ ID NO:6 (*e.g.*, a sequence comprising nucleotides 3181-3220, 3184-3210, or 3178-3213 of SEQ ID NO:6 (see FIG. 2D)). In another embodiment, the nucleic acid molecule includes a fusion, *e.g.*, an in-frame fusion, of at least exon 20 of ALK or a fragment thereof (*e.g.*, exons 20-29 ALK or a fragment thereof), and at least exon 26 or a fragment thereof (*e.g.*, exons 1-26 of DCNT1 or a fragment thereof).

In certain embodiments, the DCNT1-ALK fusion is in a 5'- ALK to 3'-DCTN1 configuration referred to herein as "5'ALK 3'DCTN1"). In one embodiment, the nucleic acid molecule includes the nucleotides sequence of 1-3172 of SEQ ID NO:11 (corresponding to exons 1-19 of ALK gene) or a fragment thereof, or a sequence substantially identical thereto. In another embodiment, the nucleic acid molecule includes the nucleotides sequence of 3173-3813 of SEQ ID NO:11 (corresponding to exons 27-32 of the a DCTN1 gene) or a fragment thereof, or a sequence substantially identical thereto. In yet other embodiments, the nucleic acid molecule includes the nucleotide sequence shown in FIGs. 4A-4D (*e.g.*, SEQ ID NO:11) or a fragment thereof, or a sequence substantially identical thereto. In one embodiment, the nucleic acid molecule is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition described herein to SEQ ID NO:11 or a fragment thereof. In yet other embodiment, the nucleic acid molecule hybridizes to a nucleotide sequence that is complementary to at least a portion of a nucleotide sequence disclosed herein, *e.g.*, is capable of hybridizing under a stringency condition to a nucleotide sequence complementary to SEQ ID NO:11 or a fragment thereof. The nucleotide sequence of a cDNA encoding an exemplary 5'ALK-3'DCTN1 fusion is shown in SEQ ID NO:11, and the predicted amino acid sequence is shown in SEQ ID ID NO:12.

In yet other embodiments, the DCTN1-ALK fusion polypeptide is encoded by a nucleic acid molecule described herein. In one embodiment, the DCTN1-ALK fusion polypeptide is encoded by an in-frame fusion of intron 26 of DCTN1 with intron 19 of ALK (*e.g.*, a sequence on chromosome 2). In other embodiments, the DCTN1-ALK

IDT Exhibit 1020.2950

fusion polypeptide is encoded by a nucleotide sequence in the region of 29,447,851-29,448,653 of chromosome 2 coupled to (*e.g.,* juxtaposed to) nucleotides in the region of nucleotides 74,591,512-74,592,314 of chromosome 2. In another embodiment, the DCTN1-ALK fusion polypeptide includes an amino acid sequence encoded by a nucleotide sequence comprising a fusion junction between the DCTN transcript and the ALK transcript, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:6 (*e.g.*, a nucleotide sequence within exon 26 of a DCTN1 and exon 20 of an ALK gene).

In an embodiment, the 5'DCTN1-3'ALK fusion polypeptide comprises sufficient DCTN1 and sufficient ALK sequence such that it has kinase activity, *e.g.*, has elevated activity, *e.g.*, ALK kinase activity, as compared with wild type ALK, e.g., in a cell of a cancer referred to herein. In an embodiment the 5'DCTN1-3'ALK fusion polypeptide comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or 11 exons from DCTN1 and at least 1, 2, 3, 4, 5, 6, 7, 9, or 10, ALK exons. In one embodiment, the 5'DCTN1-3'ALK fusion polypeptide includes a dynein or kinesin II binding domain or a functional fragment thereof, and an ALK tyrosine kinase domain or a functional fragment thereof. In a related aspect, the invention features DCTN1-ALK fusion polypeptide or fragments operatively linked to heterologous polypeptides to form fusion proteins.

In another embodiment, the DCTN1-ALK fusion polypeptide or fragment is a peptide, *e.g.*, an immunogenic peptide or protein, that contains a fusion junction described herein. Such immunogenic peptides or proteins can be used to raise antibodies specific to the fusion protein. In other embodiments, such immunogenic peptides or proteins can be used for vaccine preparation. The vaccine preparation can include other components, *e.g.*, an adjuvant.

In another aspect, the invention features antibody molecules that bind to a DCTN1-ALK fusion polypeptide or fragment described herein. In embodiments the antibody can distinguish wild type ALK (or DCTN1) from DCTN1-ALK.

LMNA-NTRK1 Fusion Polypeptides

In another aspect, the invention features a LMNA-NTRK1 fusion polypeptide (*e.g.*, a purified LMNA-NTRK1 fusion polypeptide), a biologically active or antigenic fragment thereof, as well as reagents (*e.g.*, antibody molecules that bind to a LMNA-NTRK1 fusion polypeptide), and detection of an LMNA-NTRK1 polypeptide.

IDT Exhibit 1020.2951

In one embodiment, the LMNA-NTRK1 fusion polypeptide has at least one biological activity, *e.g.*, an NTRK1 kinase activity, lamin matrix activity, and/or a dimerizing or multimerizing activity.

In other embodiments, the LMNA-NTRK1 fusion polypeptide includes a fragment of a LMNA polypeptide and a fragment of an NTRK1 polypeptide. In one embodiment, the LMNA-NTRK1 fusion polypeptide includes amino acids 167-175 of SEQ ID NO:10 or a fragment thereof (*e.g.*, amino acids 1-175 of SEQ ID NO:10 or a fragment thereof), and amino acids 168-176 of SEQ ID NO:10 or a fragment thereof (*e.g.*, amino acids 168-580 of SEQ ID NO:10 or a fragment thereof). In yet other embodiments, the LMNA-NTRK1 fusion polypeptide includes an amino acid sequence substantially identical to an in-frame fusion of amino acids 168-176 of SEQ ID NO:10 or a fragment thereof (*e.g.*, amino acids 1-176 of SEQ ID NO:10 or a fragment thereof), and amino acids 168-176 of SEQ ID NO:10 or a fragment thereof (*e.g.*, amino acids 168-580 of SEQ ID NO:10 or a fragment thereof).

In other embodiments, the LMNA-NTRK1 fusion polypeptide includes a LMNA dynein association domain or a fragment thereof, and an NTRK1 tyrosine kinase domain or a fragment thereof. In another embodiment, the LMNA-NTRK1 fusion polypeptide includes the amino acid sequence of amino acids 1-176 of SEQ ID NO:10 or a fragment thereof, or a sequence substantially identical thereto. For example, the LMNA-TRK1 fusion polypeptide can include a laminin binding domain of LMNA or a fragment thereof. In other embodiments, the LMNA-NTRK1 fusion polypeptide includes the amino acid sequence of amino acids 289-560 of SEQ ID NO:10 or a fragment thereof, or a sequence substantially identical thereto. For example, the LMNA-NTRK1 fusion polypeptide can include an NTRK1 kinase domain that includes amino acids 269-560 of SEQ ID NO:10 or a fragment thereof. In yet other embodiments, the LMNA-NTRK1 fusion polypeptide includes the amino acid sequence shown in FIGs. 7A-7B (*e.g.*, SEQ ID NO:10) or a fragment thereof, or a sequence substantially identical thereto.

In yet other embodiments, the LMNA-NTRK1 fusion polypeptide is encoded by a nucleic acid molecule described herein. In one embodiment, the LMNA-NTRK1 fusion polypeptide is encoded by an in-frame fusion of intron 2 of LMNA with intron 10 of NTRK1. In other embodiments, the nucleic acid molecules includes a nucleotide sequence in the region of 156,100,970-156,101,459 of chromosome 1 coupled to (*e.g.*,

IDT Exhibit 1020.2952

juxtaposed to) nucleotides in the region of nucleotides 156,844,390-156,844,879 of chromosome 1. In another embodiment, the LMNA-NTRK1 fusion polypeptide includes an amino acid sequence encoded by a nucleotide sequence comprising a fusion junction between the LMNA transcript and the NTRK1 transcript, *e.g.*, a nucleotide sequence that includes a portion of SEQ ID NO:8 (*e.g.*, a nucleotide sequence within introns 2 of an LMNA gene and 10 of an NTRK1 gene) (*e.g.*, SEQ ID NO:10).

In yet other embodiments, the LMNA-NTRK1 fusion polypeptide is encoded by a 5'- LMNA to 3'- NTRK1 nucleic acid molecule described herein. In one embodiment, the LMNA-NTRK1 fusion polypeptide is encoded by a nucleotide sequence that includes a fusion junction between the LMNA transcript and the NTRK1 transcript, *e.g.*, a sequence comprising nucleotides 502-525 of SEQ ID NO:9. In yet other embodiments, the LMNA-TRK1 fusion polypeptide is encoded by the nucleotide sequence shown in FIGs. 7A-7B (*e.g.*, SEQ ID NO:9) or a fragment thereof, or a sequence substantially identical thereto.

In an embodiment, the 5'LMNA-3'NTRK1 fusion polypeptide comprises sufficient LMNA and sufficient NTRK1 sequence such that it has kinase activity, *e.g.*, has elevated activity, *e.g.*, NTRK1 kinase activity, as compared with wild type NTRK1, e.g., in a cell of a cancer referred to herein. In an embodiment the 5'LMNA-3'NTRK1 fusion polypeptide comprises at least 1, 2, 3, 4, 5, 6, 7, 9, 10, or 11 exons from LMNA and at least 1, 2, 3, 4, 5, 6, 7, 9, or 10, NTRK1 exons. In one embodiment, the 5'LMNA-3'NTRK1 fusion polypeptide includes a CAAX farnesylation motif, an autophosphorylation site, or a functional fragment thereof, and an NRTK1 tyrosine kinase domain or a functional fragment thereof. In a related aspect, the invention features LMNA-NTRK1 fusion polypeptide or fragments operatively linked to heterologous polypeptides to form fusion proteins.

In another embodiment, the LMNA-NTRK1 fusion polypeptide or fragment is a peptide, *e.g.*, an immunogenic peptide or protein that contains a fusion junction described herein. Such immunogenic peptides or proteins can be used to raise antibodies specific to the fusion protein. In other embodiments, such immunogenic peptides or proteins can be used for vaccine preparation. The vaccine preparation can include other components, *e.g.*, an adjuvant.

IDT Exhibit 1020.2953

In another aspect, the invention features antibody molecules that bind to a LMNA-NTRK1 fusion polypeptide or fragment described herein. In embodiments the antibody can distinguish wild type NRTK1 (or LMNA) from LMNA-NTRK1.

Detection Reagents and Detection of Mutations

In another aspect, the invention features a detection reagent, *e.g.*, a purified or an isolated preparation thereof. Detection reagents can distinguish a nucleic acid, or protein sequence, having a breakpoint, *e.g.*, a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g.*, SEQ ID NO:5); or an LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g.*, SEQ ID NO:8), from a reference sequence. In one embodiment, the detection reagent detects (*e.g.*, specifically detects) a DCTN1-ALK fusion nucleic acid or a polypeptide (*e.g.*, distinguishes a wild type ALK or another ALK fusion (or DCTN1) from a DCTN1-ALK nucleic acid (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:6); or a DCTN1-ALK polypeptide (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:7). In another embodiment, the detection reagent detects (*e.g.*, specifically detects) an LMNA-NTRK1 fusion nucleic acid or polypeptide (e.g., distinguishes a wild type NTRK1 or another NTRK1 fusion (or LMNA) from LMNA-NTRK1 nucleic acid (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:9); or LMNA-NTRK1 polypeptide (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:10).

Detection reagents, *e.g.*, nucleic acid-based detection reagents, can be used to identify mutations in a target nucleic acid, *e.g.*, DNA, *e.g.*, genomic DNA or cDNA, or RNA, *e.g.*, in a sample, *e.g.*, a sample of nucleic acid derived from a neoplastic or tumor cell, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cell. Detection reagents, *e.g.*, antibody-based detection reagents, can be used to identify mutations in a target protein, *e.g.*, in a sample, *e.g.*, a sample of protein derived from, or produced by, a neoplastic or tumor cell, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cell.

Nucleic Acid-based Detection Reagents

In an embodiment, the detection reagent comprises a nucleic acid molecule, *e.g.*, a DNA, RNA or mixed DNA/RNA molecule, comprising sequence which is complementary with a nucleic acid sequence on a target nucleic acid (the sequence on

IDT Exhibit 1020.2954

the target nucleic acid that is bound by the detection reagent is referred to herein as the "detection reagent binding site" and the portion of the detection reagent that corresponds to the detection reagent binding site is referred to as the "target binding site"). In an embodiment, the detection reagent binding site is disposed in relationship to the interrogation position such that binding (or in embodiments, lack of binding) of the detection reagent to the detection reagent binding site allows differentiation of mutant and reference sequences for a mutant described herein (*e.g.*, a translocation having a breakpoint, *e.g.*, a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g.*, SEQ ID NO:5); or an LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g.*, SEQ ID NO:8)), from a reference sequence. The detection reagent can be modified, *e.g.*, with a label or other moiety, *e.g.*, a moiety that allows capture.

In an embodiment, the detection reagent comprises a nucleic acid molecule, *e.g.*, a DNA, RNA or mixed DNA/RNA molecule, which, *e.g.*, in its target binding site, includes the interrogation position and which can distinguish (*e.g.*, by affinity of binding of the detection reagent to a target nucleic acid or the ability for a reaction, *e.g.*, a ligation or extension reaction with the detection reagent) between a mutation, *e.g.*, a translocation described herein, and a reference sequence. In embodiments, the interrogation position can correspond to a terminal, *e.g.*, to a 3' or 5' terminal nucleotide, a nucleotide immediately adjacent to a 3' or 5' terminal nucleotide, or to another internal nucleotide, of the detection reagent or target binding site.

In embodiments, the difference in the affinity of the detection reagent for a target nucleic acid comprising the mutant and that for a target nucleic acid comprising the reference sequence allows determination of the presence or absence of the mutation (or reference) sequence. Typically, such detection reagents, under assay conditions, will exhibit substantially higher levels of binding only to the mutant or only to the reference sequence, *e.g.*, will exhibit substantial levels of binding only to the mutant or only to the reference sequence.

In embodiments, binding allows (or inhibits) a subsequent reaction, *e.g.*, a subsequent reaction involving the detection reagent or the target nucleic acid. *E.g.*, binding can allow ligation, or the addition of one or more nucleotides to a nucleic acid, *e.g.*, the detection reagent, *e.g.*, by DNA polymerase, which can be detected and

IDT Exhibit 1020.2955

used to distinguish mutant from reference. In embodiments, the interrogation position is located at the terminus, or sufficiently close to the terminus, of the detection reagent or its target binding site, such that hybridization, or a chemical reaction, e.g., the addition of one or more nucleotides to the detection reagent, *e.g.*, by DNA polymerase, only occurs, or occurs at a substantially higher rate, when there is a perfect match between the detection reagent and the target nucleic acid at the interrogation position or at a nucleotide position within 1, 2, or 3 nucleotides of the interrogation position.

In an embodiment, the detection reagent comprises a nucleic acid, *e.g.*, a DNA, RNA or mixed DNA/RNA molecule wherein the molecule, or its target binding site, is adjacent (or flanks), *e.g.*, directly adjacent, to the interrogation position, and which can distinguish between a mutation, *e.g.*, a translocation described herein, and a reference sequence, in a target nucleic acid.

In embodiments, the detection reagent binding site is adjacent to the interrogation position, *e.g.*, the 5' or 3'terminal nucleotide of the detection reagent, or its target binding site, is adjacent, *e.g.*, between 0 (directly adjacent) and 1,000, 500, 400, 200, 100, 50, 10, 5, 4, 3, 2, or 1 nucleotides from the interrogation position. In embodiments, the outcome of a reaction will vary with the identity of the nucleotide at the interrogation position allowing one to distinguish between mutant and reference sequences. *E.g.*, in the presence of a first nucleotide at the interrogation position a first reaction will be favored over a second reaction. *E.g.*, in a ligation or primer extension reaction, the product will differ, *e.g.*, in charge, sequence, size, or susceptibility to a further reaction (*e.g.*, restriction cleavage) depending on the identity of the nucleotide at the interrogation position. In embodiments the detection reagent comprises paired molecules (e.g., forward and reverse primers), allowing for amplification, *e.g.*, by PCR amplification, of a duplex containing the interrogation position. In such embodiments, the presence of the mutation can be determined by a difference in the property of the amplification product, *e.g.*, size, sequence, charge, or susceptibility to a reaction, resulting from a sequence comprising the interrogation position and a corresponding sequence having a reference nucleotide at the interrogation positions. In embodiments, the presence or absence of a characteristic

IDT Exhibit 1020.2956

amplification product is indicative of the identity of the nucleotide at the interrogation site and thus allows detection of the mutation.

In embodiments, the detection reagent, or its target binding site, is directly adjacent to the interrogation position, *e.g.*, the 5' or 3'terminal nucleotide of the detection reagent is directly adjacent to the interrogation position. In embodiments, the identity of the nucleotide at the interrogation position will determine the nature of a reaction, *e.g.*, a reaction involving the detection reagent, *e.g.*, the modification of one end of the detection reagent. *E.g.*, in the presence of a first nucleotide at the interrogation position a first reaction will be favored over a second reaction. By way of example, the presence of a first nucleotide at the interrogation position, *e.g.*, a nucleotide associated with a mutation, can promote a first reaction, *e.g.*, the addition of a complementary nucleotide to the detection reagent. By way of example, the presence of an A at the interrogation position will cause the incorporation of a T, having, *e.g.*, a first colorimetric label, while the presence of a G and the interrogation position will cause the incorporation for a C, having, *e.g.*, a second colorimetric label. In an embodiment, the presence of a first nucleotide at the nucleotide will result in ligation of the detection reagent to a second nucleic acid. *E.g.*, a third nucleic acid can be hybridized to the target nucleic acid sufficiently close to the interrogation site that if the third nucleic acid has an exact match at the interrogation site it will be ligated to the detection reagent. Detection of the ligation product, or its absence, is indicative of the identity of the nucleotide at the interrogation site and thus allows detection of the mutation.

A variety of readouts can be employed. *E.g.*, binding of the detection reagent to the mutant or reference sequence can be followed by a moiety, *e.g.*, a label, associated with the detection reagent, *e.g.*, a radioactive or enzymatic label. In embodiments the label comprises a quenching agent and a signaling agent and hybridization results in altering the distance between those two elements, *e.g.*, increasing the distance and un-quenching the signaling agent. In embodiments, the detection reagent can include a moiety that allows separation from other components of a reaction mixture. In embodiments, binding allows cleavage of the bound detection reagent, *e.g.*, by an enzyme, *e.g.*, by the nuclease activity of the DNA polymerase or by a restriction enzyme. The cleavage can be detected by the

IDT Exhibit 1020.2957

appearance or disappearance of a nucleic acid or by the separation of a quenching agent and a signaling agent associated with the detection reagent. In embodiments, binding protects, or renders the target susceptible, to further chemical reaction, *e.g.*, labeling or degradation, *e.g.*, by restriction enzymes. In embodiments binding with the detection reagent allows capture separation or physical manipulation of the target nucleic acid to thereby allow for identification. In embodiments binding can result in a detectable localization of the detection reagent or target, *e.g.*, binding could capture the target nucleic acid or displace a third nucleic acid. Binding can allow for determination of the presence of mutant or reference sequences with FISH, particularly in the case of rearrangements. Binding can allow for the extension or other size change in a component, *e.g.*, the detection reagent, allowing distinction between mutant and reference sequences. Binding can allow for the production, *e.g.*, by PCR, of an amplicon that distinguishes mutant from reference sequence.

In an embodiment the detection reagent, or the target binding site, is between 5 and 500, 5 and 300, 5 and 250, 5 and 200, 5 and 150, 5 and 100, 5 and 50, 5 and 25, 5 and 20, 5 and 15, or 5 and 10 nucleotides in length. In an embodiment the detection reagent, or the target binding site, is between 10 and 500, 10 and 300, 10 and 250, 10 and 200, 10 and 150, 10 and 100, 10 and 50, 10 and 25,10 and 20, or 10 and 15, nucleotides in length. In an embodiment the detection reagent, or the target binding site, is between 20 and 500, 20 and 300, 20 and 250, 20 and 200, 20 and 150, 20 and 100, 20 and 50, or 20 and 25 nucleotides in length. In an embodiment the detection reagent, or the target binding site, is sufficiently long to distinguish between mutant and reference sequences and is less than 100, 200, 300, 400, or 500 nucleotides in length.

Preparations of Mutant Nucleic Acid and Uses Thereof

In another aspect, the invention features purified or isolated preparations of a neoplastic or tumor cell nucleic acid, *e.g.*, DNA, *e.g.*, genomic DNA or cDNA, or RNA, containing an interrogation position described herein, useful for determining if a mutation disclosed herein is present. The nucleic acid includes the interrogation position, and typically additional fusion sequence on one or both sides of the interrogation position. In addition the nucleic acid can contain heterologous

IDT Exhibit 1020.2958

sequences, *e.g.*, adaptor or priming sequences, typically attached to one or both terminus of the nucleic acid. The nucleic acid also includes a label or other moiety, *e.g.*, a moiety that allows separation or localization.

In embodiments, the nucleic acid is between 20 and 1,000, 30 and 900, 40 and 800, 50 and 700, 60 and 600, 70 and 500, 80 and 400, 90 and 300, or 100 and 200 nucleotides in length (with or without heterologous sequences). In one embodiment, the nucleic acid is between 40 and 1,000, 50 and 900, 60 and 800, 70 and 700, 80 and 600, 90 and 500, 100 and 400, 110 and 300, or 120 and 200 nucleotides in length (with or without heterologous sequences). In another embodiment, the nucleic acid is between 50 and 1,000, 50 and 900, 50 and 800, 50 and 700, 50 and 600, 50 and 500, 50 and 400, 50 and 300, or 50 and 200 nucleotides in length (with or without heterologous sequences). In embodiments, the nucleic acid is of sufficient length to allow sequencing (*e.g.*, by chemical sequencing or by determining a difference in T_m between mutant and reference preparations) but is optionally less than 100, 200, 300, 400, or 500 nucleotides in length (with or without heterologous sequences).

Such preparations can be used to sequence nucleic acid from a sample, *e.g.*, a neoplastic or tumor sample. In an embodiment the purified preparation is provided by *in situ* amplification of a nucleic acid provided on a substrate. In embodiments the purified preparation is spatially distinct from other nucleic acids, *e.g.*, other amplified nucleic acids, on a substrate.

In an embodiment, the purified or isolated preparation of nucleic acid is derived from a neoplasm or tumor of a type described herein, *e.g.*, neoplasm and/or cancer, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cancer. In one embodiment, the fusion nucleic acid is derived from a histiocytoses, *e.g.*, a non-Langerhans cell histiocytosis.

Such preparations can be used to determine if a sample comprises mutant sequence, *e.g.*, a translocation as described herein. In one embodiment, the translocation includes a breakpoint, *e.g.*, a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g.*, SEQ ID NO:5). In other embodiments, the translocation includes an LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g.*, SEQ ID NO:8). Nucleic acids that include the aforesaid breakpoints, *e.g.*, a DCTN1-ALK breakpoint

IDT Exhibit 1020.2959

or an LMNA-NTRK1 breakpoint, are collectively referred to herein as fusion nucleic acids.

In another aspect, the invention features, a method of determining the sequence of an interrogation position for a mutation described herein, comprising:

providing a purified or isolated preparations of nucleic acid or fusion nucleic acid, *e.g.*, DNA, *e.g.*, genomic DNA or cDNA, or RNA, containing an interrogation position described herein,

sequencing, by a method that breaks or forms a chemical bond, *e.g.*, a covalent or non-covalent chemical bond, *e.g.*, in a detection reagent or a target sequence, the nucleic acid so as to determine the identity of the nucleotide at an interrogation position. The method allows determining if a mutation described herein is present.

In an embodiment, sequencing comprises contacting the fusion nucleic acid with a detection reagent described herein.

In an embodiment, sequencing comprises determining a physical property, *e.g.*, stability of a duplex form of the fusion nucleic acid, *e.g.*, T_m, that can distinguish mutant from reference sequence.

In an embodiment, the fusion nucleic acid is derived from a neoplasm or a tumor of a type described herein, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cancer. In one embodiment, the fusion nucleic acid is derived from a histiocytoses, *e.g.*, a non-Langerhans cell histiocytosis.

Reaction Mixtures and Devices

In another aspect, the invention features, purified or isolated preparations of a fusion nucleic acid, *e.g.*, DNA, *e.g.*, genomic DNA or cDNA, or RNA, containing an interrogation position described herein, useful for determining if a mutation disclosed herein is present, disposed in sequencing device, or a sample holder for use in such a device. In an embodiment, the fusion nucleic acid is derived from a neoplasm or a tumor of a type described herein, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cancer. In one embodiment, the fusion nucleic acid is derived from a histiocytoses, *e.g.*, a non-Langerhans cell histiocytosis.

In another aspect, the invention features, purified or isolated preparations of a fusion nucleic acid, *e.g.*, DNA, *e.g.*, genomic DNA or cDNA, or RNA, containing an

IDT Exhibit 1020.2960

interrogation position described herein, useful for determining if a mutation disclosed herein is present, disposed in a device for determining a physical or chemical property, *e.g.*, stability of a duplex, e.g., T_m or a sample holder for use in such a device. In an embodiment, the device is a calorimeter. In an embodiment the fusion nucleic acid is derived from a neoplasm or a tumor of a type described herein, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cancer. In one embodiment, the fusion nucleic acid is derived from a histiocytoses, *e.g.*, a non-Langerhans cell histiocytosis.

The detection reagents described herein can be used to determine if a mutation described herein is present in a sample. In embodiments, the sample comprises a nucleic acid that is derived from a neoplastic or a tumor cell. The cell can be from a neoplastic or a tumor sample, *e.g.*, a biopsy taken from the neoplasm or the tumor; from circulating tumor cells, *e.g.*, from peripheral blood; or from a blood or plasma sample. In an embodiment, the fusion nucleic acid is derived from a neoplasm or a tumor of a type described herein, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cancer. In one embodiment, the fusion nucleic acid is derived from a histiocytoses, *e.g.*, a non-Langerhans cell histiocytosis.

Accordingly, in one aspect, the invention features a method of making a reaction mixture, comprising:

combining a detection reagent, or purified or isolated preparation thereof, described herein with a target nucleic acid derived from a neoplastic or a tumor cell which comprises a sequence having an interrogation position for a mutation described herein.

In another aspect, the invention features a reaction mixture, comprising:

a detection reagent, or purified or isolated preparation thereof, described herein; and

a target nucleic acid derived from a neoplastic or tumor cell which comprises a sequence having an interrogation position for a mutation described herein.

In an embodiment of the reaction mixture, or the method of making the reaction mixture:

the detection reagent comprises a nucleic acid, *e.g.*, a DNA, RNA or mixed DNA/RNA, molecule which is complementary with a nucleic acid sequence on a

IDT Exhibit 1020.2961

target nucleic acid (the detection reagent binding site) wherein the detection reagent binding site is disposed in relationship to the interrogation position such that binding of the detection reagent to the detection reagent binding site allows differentiation of mutant and reference sequences for a mutant described herein.

In an embodiment of the reaction mixture, or the method of making the reaction mixture, the neoplasm or a tumor is as described herein.

In an embodiment of the reaction mixture, or the method of making the reaction mixture:

the mutation is a mutation described herein, including: a translocation event, e.g., a translocation as described herein. In one embodiment, the mutation is a breakpoint, *e.g.*, a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g.*, SEQ ID NO:5); or an LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g.*, SEQ ID NO:8). In another embodiment, the mutation is a DCTN1-ALK fusion nucleic acid or polypeptide (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:6); or DCTN1-ALK polypeptide (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:7). In another embodiment, the mutation is an LMNA-NTRK1 fusion nucleic acid or polypeptide (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:9); or LMNA-NTRK1 polypeptide (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:10)).

A mutation described herein, can be distinguished from a reference, *e.g.*, a non-mutant or wildtype sequence, by reaction with an enzyme that reacts differentially with the mutation and the reference. *E.g.*, they can be distinguished by cleavage with a restriction enzyme that has differing activity for the mutant and reference sequences. *E.g.*, the invention includes a method of contacting a nucleic acid comprising a mutation described herein with such an enzyme and determining if a product of that cleavage which can distinguish mutant form reference sequence is present.

In one aspect the inventions provides, a purified preparation of a restriction enzyme cleavage product which can distinguish between mutant and reference sequence, wherein one end of the cleavage product is defined by an enzyme that cleaves differentially between mutant and reference sequence. In an embodiment, the cleavage product includes the interrogation position.

IDT Exhibit 1020.2962

Protein-based Detection Reagents, Methods, Reaction Mixtures and Devices

A mutant protein described herein can be distinguished from a reference, *e.g.*, a non-mutant or wild-type protein, by reaction with a reagent, *e.g.*, a substrate, *e.g*, a substrate for catalytic activity, *e.g.*, phosphorylation or other ALK or NTRK1 activity, or an antibody, that reacts differentially with the mutant and reference protein. In one aspect, the invention includes a method of contacting a sample comprising a mutant protein described herein with such reagent and determining if the mutant protein is present in the sample.

In another embodiment, the invention features, an antibody that can distinguish a mutant protein described herein, e.g., a mutant protein corresponding to a junction shown FIG. 2 or FIG. 4, or an associated mutation from a reference, *e.g.*, a non-mutant or wildtype protein.

Accordingly, in one aspect, the invention features a method of making a reaction mixture comprising:

combining a detection reagent, or purified or isolated preparation thereof, *e.g.*, a substrate, *e.g.*, a substrate for phosphorylation or other activity, or an antibody, described herein with a target fusion protein derived from a neoplastic or a tumor cell which comprises a sequence having an interrogation position for a mutation described herein.

In another aspect, the invention features, a reaction mixture comprising:

a detection reagent, or purified or isolated preparation thereof, *e.g.*, a substrate, e.g., a substrate for phosphorylation or other activity, or an antibody, described herein; and

a target fusion protein derived from a neoplastic or a tumor cell which comprises a sequence having an interrogation position for a mutation described herein.

In an embodiment of the reaction mixture, or the method of making the reaction mixture:

the detection reagent comprises an antibody specific for a mutant fusion protein described herein.

In an embodiment of the reaction mixture, or the method of making the reaction mixture:

IDT Exhibit 1020.2963

the neoplastic or a tumor cell is a cell described herein, *e.g.*, a melanocytic neoplasm, melanoma or metastatic cell.

In an embodiment of the reaction mixture, or the method of making the reaction mixture:

the mutation is a mutation described herein, including: a translocation event, e.g., a translocation as described herein. In one embodiment, the mutation is a breakpoint, *e.g.*, a DCTN1-ALK breakpoint, *e.g.*, as identified in FIG. 1 (*e.g.*, SEQ ID NO:5); or an LMNA-NTRK1 breakpoint, *e.g.*, as identified in FIG. 6 (*e.g.*, SEQ ID NO:8). In another embodiment, the mutation is a DCTN1-ALK fusion nucleic acid or polypeptide (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:6); or DCTN1-ALK polypeptide (*e.g.*, as described herein in FIG. 2 (SEQ ID NO:7). In another embodiment, the mutation is an LMNA-NTRK1 fusion nucleic acid or polypeptide (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:9); or LMNA-NTRK1 polypeptide (*e.g.*, as described herein in FIG. 7 (SEQ ID NO:10)).

Kits

In another aspect, the invention features a kit comprising a detection reagent as described herein.

Methods Reducing a DCTN1-ALK Activity

In another aspect, the invention features a method of reducing an activity of a DCTN1-ALK fusion. The method includes contacting the DCTN1-ALK fusion, or a DCTN1-ALK -expressing cell, with an agent that inhibits an activity or expression of DCTN1-ALK (*e.g.*, a kinase inhibitor). In one embodiment, the contacting step can be effected *in vitro*, *e.g.*, in a cell lysate or in a reconstituted system. Alternatively, the method can be performed on cells in culture, *e.g.*, *in vitro* or *ex vivo*. In other embodiments, the method can be performed on DCTN1-ALK -expressing cells present in a subject, *e.g.*, as part of an *in vivo* (*e.g.*, therapeutic or prophylactic) protocol. In an embodiment the method is practiced on an animal subject (*e.g.*, an *in vivo* animal model). In certain embodiments, the DCTN1-ALK fusion is a nucleic acid molecule or a polypeptide as described herein.

IDT Exhibit 1020.2964

In a related aspect, a method of inhibiting, reducing, or treating a hyperproliferative disorder, *e.g.*, a neoplasm (including benign, pre-malignant or malignant (*e.g.*, a cancer), in a subject is provided. The method includes administering to the subject a preselected therapeutic agent, *e.g.*, an anti-cancer agent (*e.g.*, a kinase inhibitor), as a single agent, or in combination, in an amount sufficient to reduce, inhibit or treat the activity or expression of DCTN1-ALK (*e.g.*, a DCTN1-ALK fusion described herein), thereby inhibiting, reducing, or treating the hyperproliferative disorder in the subject. "Treatment" as used herein includes, but is not limited to, inhibiting tumor growth, reducing tumor mass, reducing size or number of metastatic lesions, inhibiting the development of new metastatic lesions, prolonged survival, prolonged progression-free survival, prolonged time to progression, and/or enhanced quality of life.

In one embodiment, the kinase inhibitor is administered based on a determination that a DCTN1-ALK fusion is present in a subject, *e.g.,* based on its present in a subject's sample. Thus, treatment can be combined with a DCTN1-ALK detection or evaluation method, *e.g.*, as described herein, or administered in response to a determination made by a DCTN1-ALK detection or evaluation method, *e.g.,* as described herein. In certain embodiments, the kinase inhibitor is administered responsive to acquiring knowledge or information of the presence of the DCTN1-ALK fusion in a subject. In one embodiment, the kinase inhibitor is administered responsive to acquiring knowledge or information on the subject's genotype, *e.g.*, acquiring knowledge or information that the patient's genotype has a DCTN1-ALK fusion. In other embodiments, the kinase inhibitor is administered responsive to receiving a communication (*e.g.*, a report) of the presence of the DCTN1-ALK fusion in a subject (*e.g.*, a subject's sample). In yet other embodiments, the kinase inhibitor is administered responsive to information obtained from a collaboration with another party that identifies the presence of the DCTN1-ALK fusion in a subject (*e.g.*, a subject's sample). In other embodiments, the kinase inhibitor is administered responsive to a determination that the DCTN1-ALK fusion is present in a subject. In one embodiment, the determination of the presence of the DCTN1-ALK fusion is carried out using one or more of the methods, *e.g.*, the sequencing methods, described herein. In other embodiments, the determination of the presence of the DCTN1-ALK

IDT Exhibit 1020.2965

fusion includes receiving information on the subject's DCTN1-ALK fusion genotype, *e.g.*, from another party or source.

The methods can, optionally, further include the step(s) of identifying (*e.g.*, evaluating, diagnosing, screening, and/or selecting) a subject at risk of having, or having, a DCTN1-ALK fusion. In one embodiment, the method further includes one or more of: acquiring knowledge or information of the presence of the DCTN1-ALK fusion in a subject (*e.g.*, a subject's sample); acquiring knowledge or information on the subject's genotype, *e.g.*, acquiring knowledge or information that the patient's genotype has a DCTN1-ALK fusion; receiving a communication (*e.g.*, a report) of the presence of the DCTN1-ALK fusion in a subject (*e.g.*, a subject's sample); or collaborating with another party that identifies the presence of the DCTN1-ALK fusion in a subject.

In one embodiment, the subject treated has a DCTN1-ALK fusion; *e.g.*, the subject has a tumor or cancer harboring a DCTN1-ALK fusion. In other embodiments, the subject has been previously identified as having a DCTN1-ALK fusion. In yet other embodiments, the subject has been previously identified as being likely or unlikely to respond to treatment with a protein kinase inhibitor, *e.g.*, a subject that has previously participated in a clinical trial. In other embodiments, the subject has been previously identified as being likely or unlikely to respond to treatment with a protein kinase inhibitor, based on the presence of the DCTN1-ALK fusion. In one embodiment, the subject is a mammal, *e.g.*, a human. In one embodiment, the subject has, or at risk of having a cancer at any stage of disease. In other embodiments, the subject is a patient, *e.g.*, a cancer patient.

In other embodiments, the subject treated is a cancer patient who has participated in a clinical trial. For example, the subject participated in a clinical trial that evaluated a kinase inhibitor (*e.g.*, a multikinase inhibitor, an ALK kinase inhibitor). In other embodiment, the subject participated in a clinical trial that evaluates upstream or downstream targets of ALK. In one embodiment, said cancer patient responded to the kinase inhibitor evaluated.

In certain embodiments, the neoplasm or neoplastic cell is a benign, pre-malignant, malignant (cancer) or metastasis. In certain embodiments, the cancer is a solid tumor, a soft tissue tumor, or a metastatic lesion. In one embodiment, the cancer

IDT Exhibit 1020.2966

is chosen from a lung cancer, a pancreatic cancer, melanoma, a colorectal cancer, an esophageal-gastric cancer, a thyroid cancer, or an adenocarcinoma. In other embodiment, the lung cancer is chosen from one or more of the following: non-small cell lung cancer (NSCLC), small cell lung cancer (SCLC), squamous cell carcinoma (SCC), adenocarcinoma of the lung, bronchogenic carcinoma, or a combination thereof. In one embodiment, the lung cancer is NSCLC or SCC. In one embodiment, the neoplasm or cancer is a melanocytic neoplasm, a Spitz nevi, a Spitz tumor, a Spitzoid melanoma, a metastatic Spitz tumor, or a melanoma. In one embodiment, the Spitz tumor is metastatic, *e.g.*, localized to lymph nodes or widespread disease.

In one embodiment, the anti-cancer agent is a kinase inhibitor. For example, the kinase inhibitor is a multi-kinase inhibitor or an ALK-specific inhibitor. In one embodiment, the kinase inhibitor is an ALK inhibitor including, but not limited to, TAE-684 (also referred to herein as "NVP-TAE694"), PF02341066 (also referred to herein as "crizotinib" or "1066"), AF-802, LDK-378, ASP-3026, CEP-37440, CEP-28122, CEP-108050 and AP26113. Additional examples of ALK kinase inhibitors are described in examples 3-39 of WO 2005016894 by Garcia-Echeverria C, *et al.*

In other embodiments, the anti-cancer agent is a DCTN1-ALK antagonist inhibits the expression of nucleic acid encoding DCTN1-ALK. Examples of such DCTN1-ALK antagonists include nucleic acid molecules, for example, antisense molecules, ribozymes, RNAi, triple helix molecules that hybridize to a nucleic acid encoding DCTN1-ALK, or a transcription regulatory region, and blocks or reduces mRNA expression of DCTN1-ALK.

In other embodiments, the kinase inhibitor is administered in combination with a second therapeutic agent or a different therapeutic modality, *e.g.*, anti-cancer agents, and/or in combination with surgical and/or radiation procedures. For example, the second therapeutic agent can be a cytotoxic or a cytostatic agent. Exemplary cytotoxic agents include antimicrotubule agents, topoisomerase inhibitors, or taxanes, antimetabolites, mitotic inhibitors, alkylating agents, intercalating agents, agents capable of interfering with a signal transduction pathway, agents that promote apoptosis and radiation. In yet other embodiments, the methods can be used in combination with immunodulatory agents, *e.g.*, IL-1, 2, 4, 6, or 12, or interferon alpha or gamma, or immune cell growth factors such as GM-CSF.

IDT Exhibit 1020.2967

Screening Methods

In another aspect, the invention features a method, or assay, for screening for agents that modulate, *e.g.*, inhibit, the expression or activity of a DCTN1-ALK fusion, *e.g.*, a DCTN1-ALK fusion as described herein. The method includes contacting a DCTN1-ALK fusion, or a cell expressing a DCTN1-ALK fusion, with a candidate agent; and detecting a change in a parameter associated with a DCTN1-ALK fusion, *e.g.*, a change in the expression or an activity of the DCTN1-ALK fusion. The method can, optionally, include comparing the treated parameter to a reference value, *e.g.*, a control sample (*e.g.*, comparing a parameter obtained from a sample with the candidate agent to a parameter obtained from a sample without the candidate agent). In one embodiment, if a decrease in expression or activity of the DCTN1-ALK fusion is detected, the candidate agent is identified as an inhibitor. In another embodiment, if an increase in expression or activity of the DCTN1-ALK fusion is detected, the candidate agent is identified as an activator. In certain embodiments, the DCTN1-ALK fusion is a nucleic acid molecule or a polypeptide as described herein.

In one embodiment, the contacting step is effected in a cell-free system, *e.g.*, a cell lysate or in a reconstituted system. In other embodiments, the contacting step is effected in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In yet other embodiments, the contacting step is effected in a cell *in vivo* (a DCTN1-ALK -expressing cell present in a subject, *e.g.*, an animal subject (*e.g.*, an *in vivo* animal model).

Exemplary parameters evaluated include one or more of:

(i) a change in binding activity, *e.g.*, direct binding of the candidate agent to a DCTN1-ALK fusion polypeptide; a binding competition between a known ligand and the candidate agent to a DCTN1-ALK fusion polypeptide;

(ii) a change in kinase activity, *e.g.*, phosphorylation levels of a DCTN1-ALK fusion polypeptide (*e.g.*, an increased or decreased autophosphorylation); or a change in phosphorylation of a target of an ALK kinase, *e.g.*, Akt, In certain embodiments, a change in kinase activity, *e.g.*, phosphorylation, is detected by any of Western blot (*e.g.*, using an anti-DCTN1 or anti-ALK antibody; a phosphor-specific antibody, detecting a

IDT Exhibit 1020.2968

shift in the molecular weight of a DCTN1-ALK fusion polypeptide), mass spectrometry, immunoprecipitation, immunohistochemistry, immunomagnetic beads, among others;

(iii) a change in an activity of a cell containing a DCTN1-ALK fusion (*e.g.*, a tumor cell or a recombinant cell), *e.g.*, a change in proliferation, morphology or tumorigenicity of the cell;

(iv) a change in tumor present in an animal subject, *e.g.*, size, appearance, proliferation, of the tumor; or

(v) a change in the level, *e.g.*, expression level, of a DCTN1-ALK fusion polypeptide or nucleic acid molecule.

In one embodiment, a change in a cell free assay in the presence of a candidate agent is evaluated. For example, an activity of a DCTN1-ALK fusion, or interaction of a DCTN1-ALK fusion with a downstream ligand can be detected. In one embodiment, a DCTN1-ALK fusion polypeptide is contacted with a ligand, *e.g.*, in solution, and a candidate agent is monitored for an ability to modulate, *e.g.*, inhibit, an interaction, *e.g.*, binding, between the DCTN1-ALK fusion polypeptide and the ligand.

In other embodiments, a change in an activity of a cell is detected in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In one embodiment, the cell is a recombinant cell that is modified to express a DCTN1-ALK fusion nucleic acid, *e.g.*, is a recombinant cell transfected with a DCTN1-ALK fusion nucleic acid. The transfected cell can show a change in response to the expressed DCTN1-ALK fusion, *e.g.*, increased proliferation, changes in morphology, increased tumorigenicity, and/or acquired a transformed phenotype. A change in any of the activities of the cell, *e.g.*, the recombinant cell, in the presence of the candidate agent can be detected. For example, a decrease in one or more of: proliferation, tumorigenicity, transformed morphology, in the presence of the candidate agent can be indicative of an inhibitor of a DCTN1-ALK fusion. In other embodiments, a change in binding activity or phosphorylation as described herein is detected.

In yet other embodiment, a change in a tumor present in an animal subject (*e.g.*, an *in vivo* animal model) is detected. In one embodiment, the animal model is a tumor containing animal or a xenograft comprising cells expressing a DCTN1-ALK fusion (*e.g.*, tumorigenic cells expressing a DCTN1-ALK fusion). The candidate agent can be

IDT Exhibit 1020.2969

administered to the animal subject and a change in the tumor is detected. In one embodiment, the change in the tumor includes one or more of a tumor growth, tumor size, tumor burden, survival, is evaluated. A decrease in one or more of tumor growth, tumor size, tumor burden, or an increased survival is indicative that the candidate agent is an inhibitor.

In other embodiments, a change in expression of a DCTN1-ALK fusion can be monitored by detecting the nucleic acid or protein levels, *e.g.*, using the methods described herein.

In certain embodiments, the screening methods described herein can be repeated and/or combined. In one embodiment, a candidate agent that is evaluated in a cell-free or cell-based described herein can be further tested in an animal subject.

In one embodiment, the candidate agent is a small molecule compound, e.g., a kinase inhibitor, a nucleic acid (e.g., antisense, siRNA, aptamer, ribozymes, microRNA), an antibody molecule (*e.g.*, a full antibody or antigen binding fragment thereof that binds to DCTN1 or ALK). The candidate agent can be obtained from a library (*e.g.*, a commercial library of kinase inhibitors) or rationally designed (*e.g.*, based on the ALK kinase domain).

Methods for Detecting Fusions

In another aspect, the invention features a method of determining the presence of a fusion as described herein, *e.g.*, a DCTN1-ALK fusion or an LMNA-NTRK1 as described herein. In one embodiment, the fusion (*e.g.*, DCTN1-ALK or LMNA-NTRK1) is detected in a nucleic acid molecule or a polypeptide. The method includes detecting whether a fusion nucleic acid molecule or polypeptide is present in a cell (*e.g.*, a circulating cell), a tissue (*e.g.*, a tumor), or a sample, *e.g.*, a tumor sample, from a subject. In one embodiment, the sample is a nucleic acid sample. In one embodiment, the nucleic acid sample comprises DNA, *e.g.*, genomic DNA or cDNA, or RNA, *e.g.*, mRNA. In other embodiments, the sample is a protein sample.

In one embodiment, the sample is, or has been, classified as non-malignant using other diagnostic techniques, *e.g.*, immunohistochemistry.

In one embodiment, the sample is acquired from a subject (*e.g.*, a subject having or at risk of having a cancer, *e.g.*, a patient), or alternatively, the method further includes

IDT Exhibit 1020.2970

acquiring a sample from the subject. The sample can be chosen from one or more of: tissue, e.g., cancerous tissue (*e.g.*, a tissue biopsy), whole blood, serum, plasma, buccal scrape, sputum, saliva, cerebrospinal fluid, urine, stool, circulating tumor cells, circulating nucleic acids, or bone marrow. In certain embodiments, the sample is a tissue (*e.g.*, a tumor biopsy), a circulating tumor cell or nucleic acid.

In embodiments, the tumor is from a cancer described herein, *e.g.*, is chosen from a lung cancer, a colorectal cancer, an esophageal-gastric cancer, a thyroid cancer, an adenocarcinoma or a melanoma. In one embodiment, the tumor is from a lung cancer, *e.g.*, a NSCLC, a SCLC, a SCC, or a combination thereof.

In one embodiment, the subject is at risk of having, or has a cancer (*e.g.*, a patient with a cancer described herein).

In other embodiments, the DCTN1-ALK fusion or LMNA-NTRK1 fusion is detected in a nucleic acid molecule by a method chosen from one or more of: nucleic acid hybridization assay, amplification-based assays (*e.g.*, polymerase chain reaction (PCR)), PCR-RFLP assay, real-time PCR, sequencing, screening analysis (including metaphase cytogenetic analysis by standard karyotype methods, FISH (*e.g.*, break away FISH), spectral karyotyping or MFISH, comparative genomic hybridization), *in situ* hybridization, SSP, HPLC or mass-spectrometric genotyping.

In one embodiment, the method includes: contacting a nucleic acid sample, *e.g.*, a genomic DNA sample (*e.g.*, a chromosomal sample or a fractionated, enriched or otherwise pre-treated sample) or a gene product (mRNA, cDNA), obtained from the subject, with a nucleic acid fragment (*e.g.*, a probe or primer as described herein (*e.g.*, an exon-specific probe or primer) under conditions suitable for hybridization, and determining the presence or absence of the DCTN1-ALK or LMNA-NTRK1 nucleic acid molecule. The method can, optionally, include enriching a sample for the gene or gene product.

In a related aspect, a method for determining the presence of a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule is provided. The method includes: acquiring a sequence for a position in a nucleic acid molecule, *e.g.*, by sequencing at least one nucleotide of the nucleic acid molecule (*e.g.*, sequencing at least one nucleotide in the nucleic acid molecule that comprises the fusion), thereby determining that the DCTN1-ALK or LMNA-NTRK1 fusion is present in the nucleic acid molecule.

IDT Exhibit 1020.2971

Optionally, the sequence acquired is compared to a reference sequence, or a wild type reference sequence. In one embodiment, the nucleic acid molecule is from a cell (*e.g.,* a circulating cell), a tissue (*e.g.*, a tumor), or any sample from a subject (*e.g.*, blood or plasma sample). In other embodiments, the nucleic acid molecule from a tumor sample (*e.g.*, a tumor or cancer sample) is sequenced. In one embodiment, the sequence is determined by a next generation sequencing method. The method further can further include acquiring, *e.g.*, directly or indirectly acquiring, a sample, *e.g.*, a tumor or cancer sample, from a subject (*e.g.*, a patient). In certain embodiments, the cancer is chosen from a lung cancer, colorectal cancer, esophageal-gastric cancer or melanoma.

In another aspect, the invention features a method of analyzing a tumor or a circulating tumor cell. The method includes acquiring a nucleic acid sample from the tumor or the circulating cell; and sequencing, *e.g.*, by a next generation sequencing method, a nucleic acid molecule, *e.g.*, a nucleic acid molecule that includes a DCTN1-ALK fusion or a LMNA-NTRK1 as described herein.

In yet other embodiment, a fusion polypeptide is detected. The method includes: contacting a protein sample with a reagent which specifically binds to a fusion polypeptide (*e.g.,* DCTN1-ALK or LMNA-NTRK1); and detecting the formation of a complex of the fusion polypeptide and the reagent. In one embodiment, the reagent is labeled with a detectable group to facilitate detection of the bound and unbound reagent. In one embodiment, the reagent is an antibody molecule, *e.g.*, is selected from the group consisting of an antibody, and antibody derivative, and an antibody fragment.

In yet another embodiment, the level (*e.g.,* expression level) or activity the DCTN1-ALK or LMNA-NTRK1 fusion is evaluated. For example, the level (*e.g.,* expression level) or activity of the DCTN1-ALK fusion (*e.g.*, mRNA or polypeptide) is detected and (optionally) compared to a pre-determined value, *e.g.*, a reference value (*e.g.*, a control sample). Similarly, the level (*e.g.,* expression level) or activity of the LMNA-NTRK1 fusion (*e.g.,* mRNA or polypeptide) is detected and (optionally compared to a pre-determined value, *e.g.,* a reference value (*e.g.,* a control sample).

In yet another embodiment, the fusion is detected prior to initiating, during, or after, a treatment in a subject having an ALK fusion, such as a DCTN1-ALK fusion, *e.g.*, treatment with an ALK kinase inhibitor.

IDT Exhibit 1020.2972

In one embodiment, the DCTN1-ALK fusion or LMNA-NTRK1 fusion is detected at the time of diagnosis with a cancer. In other embodiment, the fusion is detected at a pre-determined interval, *e.g.*, a first point in time and at least at a subsequent point in time.

In certain embodiments, responsive to a determination of the presence of the DCTN1-ALK fusion, the method further includes one or more of:

(1) stratifying a patient population (*e.g.*, assigning a subject, *e.g.*, a patient, to a group or class);

(2) identifying or selecting the subject as likely or unlikely to respond to a treatment, *e.g.*, a kinase inhibitor treatment as described herein;

(3) selecting a treatment option, *e.g.,* administering or not administering a preselected therapeutic agent, *e.g.*, a kinase inhibitor as described herein; or

(4) prognosticating the time course of the disease in the subject (*e.g.*, evaluating the likelihood of increased or decreased patient survival).

In certain embodiments, the kinase inhibitor is a multi-kinase inhibitor or an ALK-specific inhibitor. In one embodiment, the kinase inhibitor an ALK inhibitor including, but not limited to, TAE-684 (also referred to herein as "NVP-TAE694"), PF02341066 (also referred to herein as "crizotinib" or "1066"), AF-802, LDK-378, ASP-3026, CEP-37440, CEP-28122, CEP-108050 and AP26113. Additional examples of ALK kinase inhibitors are described, *e.g.*, in examples 3-39 of WO 2005016894 by Garcia-Echeverria C, *et al.*

In certain embodiments, responsive to the determination of the presence of a DCTN1-ALK fusion, the subject is classified as a candidate to receive treatment with a kinase inhibitor, *e.g.*, a kinase inhibitor as described herein. In one embodiment, responsive to the determination of the presence of a DCTN1-ALK fusion, the subject, *e.g.*, a patient, can further be assigned to a particular class if a fusion is identified in a sample of the patient. For example, a patient identified as having a DCTN1-ALK fusion can be classified as a candidate to receive treatment with a kinase inhibitor, *e.g.*, an ALK kinase inhibitor as described herein. In one embodiment, the subject, *e.g.*, a patient, is assigned to a second class if the mutation is not present. For example, a patient who has a lung tumor that does not contain a DCTN1-ALK fusion, may be determined as not

IDT Exhibit 1020.2973

being a candidate to receive a kinase inhibitor, *e.g.*, an ALK kinase inhibitor as described herein.

In another embodiment, responsive to the determination of the presence of the DCTN1-ALK fusion, the subject is identified as likely to respond to a treatment that comprises a kinase inhibitor *e.g.*, a kinase inhibitor as described herein.

In yet another embodiment, responsive to the determination of the presence of the DCTN1-ALK fusion, the method includes administering a kinase inhibitor, *e.g.*, a kinase inhibitor as described herein, to the subject.

Method of Evaluating a Tumor or a Subject

In another aspect, the invention features a method of evaluating a subject (*e.g.*, a patient), *e.g.*, for risk of having or developing a cancer, *e.g.*, a lung cancer, colorectal cancer or skin cancer. The method includes: acquiring information or knowledge of the presence of a fusion as described herein (*e.g.*, DCTN1-ALK or LMNA-NTRK1) in a subject (*e.g.*, acquiring genotype information of the subject that identifies a fusion as being present in the subject); acquiring a sequence for a nucleic acid molecule identified herein (*e.g.*, a nucleic acid molecule that includes a DCTN1-ALK or LMNA-NTRK1 fusion sequence); or detecting the presence of a fusion nucleic acid or polypeptide in the subject), wherein the presence of the fusion (*e.g.*, DCTN1-ALK, or LMNA-NTRK1) is positively correlated with increased risk for, or having, a cancer associated with such a fusion.

The method can further include acquiring, *e.g.*, directly or indirectly, a sample from a patient and evaluating the sample for the present of a DCTN1-ALK fusion or a LMNA-NTRK1 fusion as described herein.

The method can further include the step(s) of identifying (*e.g.*, evaluating, diagnosing, screening, and/or selecting) the subject as being positively correlated with increased risk for, or having, a cancer associated with the DCTN1-ALK fusion or the LMNA-NTRK1 fusion.

In another embodiment, a subject identified has having a DCTN1-ALK fusion is identified or selected as likely or unlikely to respond to a treatment, *e.g.*, a kinase inhibitor treatment as described herein. The method can further include treating the subject with a kinase inhibitor, *e.g.*, a kinase inhibitor as described herein.

IDT Exhibit 1020.2974

In certain embodiments, the subject is a patient or patient population that has participated in a clinical trial. In one embodiment, the subject has participated in a clinical trial for evaluating a kinase inhibitor (*e.g.*, a multi-kinase inhibitor or an ALK inhibitor). In one embodiment, the clinical trial is discontinued or terminated. In other embodiments, the subject has participated in a clinical trial that evaluates an ALK kinase, a DCTN1 inhibitor (*e.g.*, a kinesin inhibitor), an upstream or downstream component of DCTN1 or ALK. In one embodiment, the subject responded favorably to the clinical trial, *e.g.*, experience an improvement in at least one symptom of a cancer (*e.g.*, decreased in tumor size, rate of tumor growth, increased survival). In other embodiments, the subject did not respond in a detectable way to the clinical trial.

In a related aspect, a method of evaluating a patient or a patient population is provided. The method includes: identifying, selecting, or obtaining information or knowledge that the patient or patient population has participated in a clinical trial; acquiring information or knowledge of the presence of a DCTN1-ALK fusion in the patient or patient population (*e.g.*, acquiring genotype information of the subject that identifies a DCTN1-ALK fusion as being present in the subject); acquiring a sequence for a nucleic acid molecule identified herein (*e.g.*, a nucleic acid molecule that includes a fusion sequence); or detecting the presence of a DCTN1-ALK fusion nucleic acid or polypeptide in the subject), wherein the presence of the fusion identifies the patient or patient population as having an increased risk for, or having, a cancer associated with the DCTN1-ALK.

In some embodiments, the method further includes treating the subject with a kinase inhibitor, *e.g.*, a kinase inhibitor as described herein.

Reporting

Methods described herein can include providing a report, such as, in electronic, web-based, or paper form, to the patient or to another person or entity, *e.g.*, a caregiver, *e.g.*, a physician, *e.g.*, an oncologist, a hospital, clinic, third-party payor, insurance company or government office. The report can include output from the method, *e.g.*, the identification of nucleotide values, the indication of presence or absence of a DCTN1-ALK fusion or an LMNA-NTRK1 fusion as described herein, or wildtype sequence. In one embodiment, a report is generated, such as in paper or electronic form,

IDT Exhibit 1020.2975

which identifies the presence or absence of an alteration described herein, and optionally includes an identifier for the patient from which the sequence was obtained.

The report can also include information on the role of a DCTN1-ALK fusion as described herein, or wildtype sequence, in disease. Such information can include information on prognosis, resistance, or potential or suggested therapeutic options. The report can include information on the likely effectiveness of a therapeutic option, the acceptability of a therapeutic option, or the advisability of applying the therapeutic option to a patient, *e.g.*, a patient having a sequence, alteration or mutation identified in the test, and in embodiments, identified in the report. For example, the report can include information, or a recommendation on, the administration of a drug, *e.g.*, the administration at a preselected dosage or in a preselected treatment regimen, *e.g.*, in combination with other drugs, to the patient. In an embodiment, not all mutations identified in the method are identified in the report. For example, the report can be limited to mutations in genes having a preselected level of correlation with the occurrence, prognosis, stage, or susceptibility of the cancer to treatment, *e.g.*, with a preselected therapeutic option. The report can be delivered, *e.g.*, to an entity described herein, within 7, 14, or 21 days from receipt of the sample by the entity practicing the method.

In another aspect, the invention features a method for generating a report, *e.g.*, a personalized cancer treatment report, by obtaining a sample, *e.g.*, a tumor sample, from a subject, detecting a DCTN1-ALK fusion as described herein in the sample, and selecting a treatment based on the mutation identified. In one embodiment, a report is generated that annotates the selected treatment, or that lists, *e.g.*, in order of preference, two or more treatment options based on the mutation identified. In another embodiment, the subject, *e.g.*, a patient, is further administered the selected method of treatment.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of

IDT Exhibit 1020.2976

conflict, the present specification, including definitions, will control. In addition, the materials, methods, and the example are illustrative only and not intended to be limiting.

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

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1 is a schematic of a balanced translocation with breakpoints in intron 26 of DCTN1 and intron 19 of ALK. The sequence at the breakpoint region is indicated (SEQ ID NO:5). The intron 19 sequence of ALK is indicated by underlining, and the fusion junction is indicated by a vertical line. The sequence of SEQ ID NO:5 is shown in the reference genome orientation.

FIGs. 2A-2F is the mRNA and amino acid sequence of the DCTN1-ALK protein fusion (SEQ ID NO:6 (nucleotide) and SEQ ID NO:7 (amino acid)). The amino acid and mRNA sequences of DCTN1 are underlined. The ALK kinase domain (amino acids 1116-1362) is indicated in bold letters.

FIG. 3 is a schematic diagram of the reciprocal translocation shown in FIG. 1, which creates an ALK-DCTN1 fusion. As a result of the translocation, ALK intron 19 is fused to DCTN1 intron 26, and the sequence at the fusion junction is shown (SEQ ID NO:13). The sequence of SEQ ID NO:13 is shown in the reference genome orientation.

FIGs. 4A-4D are the sequences of the ALK-DCTN1 fusion mRNA (SEQ ID NO:11, nucleotide) and protein (SEQ ID NO:12). The DCTN1 mRNA and protein sequences are underlined.

FIG. 5 is a snapshot of the sequencing reads illustrating the read pairs arising from the two fused constructs, mapped to the reference genome. Inward-pointing reads are pairs corresponding to the DCTN1-ALK fusion, whereas outward-pointing reads are pairs corresponding to the ALK-DCTN1 fusion.

FIG. 6 is a schematic of a ~750 bp deletion with breakpoints in intron 2 of LMNA and intron 10 of NTRK1. The sequence at the breakpoint region is indicated (SEQ ID NO:8). The intron 10 sequence of NTRK1 is indicated by underlining, and

IDT Exhibit 1020.2977

the fusion junction is indicated by a vertical line. The sequence of SEQ ID NO:8 is shown in the reference genome orientation.

FIGs. 7A and 7B are the sequences of the LMNA-NTRK1 fusion mRNA (SEQ ID NO:9, nucleotide) and protein (SEQ ID NO:10, amino acid). The LMNA mRNA and protein sequences are underlined. The NTRK1 kinase domain (amino acids 289-560) is indicated in bold letters.

FIG. 8 is a snapshot of the sequencing reads arising from the fused construct, including members of LMNA-NTRK1 read-pairs mapping to LMNA and clipped reads in read-pairs mapping within NTRK1 that capture the fusion breakpoint.

DETAILED DESCRIPTION

The invention is based, at least in part, on the discovery of novel translocation and deletion events, and their association with cancer, *e.g.*, melanoma. In one embodiment, balanced translocation on chromosome 2 that results in an in-frame fusion of a fragment of a DCTN1 gene and a fragment of an ALK gene was discovered. In another embodiment, a chromosomal deletion resulting in an in-frame fusion of the LMNA gene and the NRTK1 gene was discovered.

DCTN1-ALK fusions

The term "DCTN1-ALK" or "DCTN1-ALK fusion" is used generically herein, and includes any fusion molecule (*e.g.*, gene, gene product (*e.g.*, cDNA, mRNA, polypeptide), and variant thereof) that includes a fragment of DCTN1 and a fragment of ALK, in any configuration, including, *e.g.*, a 5'DCTN1-3'ALK or a 5'ALK-3'DCTN1 fusion molecule.

In one embodiment, a DCTN1-ALK fusion includes an in-frame fusion of an exon of DCTN1 (*e.g.*, one more of exons 1-26) and an exon of ALK (*e.g.*, one or more exons encoding an ALK tyrosine kinase domain or a fragment thereof). In one embodiment, the fusion is produced by a translocation on chromosome 2 having a breakpoint in intron 26 of DCTN1 and in intron 19 of ALK. For example, the DCTN1-ALK fusion can include an in-frame fusion of at least exon 26 of DCTN1 or a fragment thereof (*e.g.*, exons 1-26 of DCTN1 or a fragment thereof) with at least exon 10 of ALK or a fragment thereof (*e.g.*, exons 20-34 of ALK or a fragment thereof). In certain

IDT Exhibit 1020.2978

embodiments, the DCTN1-ALK fusion is in a 5'- DCTN1 to 3'-ALK configuration referred to herein as "5'DCTN1-3'ALK."

The ALK receptor tyrosine kinase is known to be associated with cancerous phenotypes including inflammatory myofibroblastic tumors, neuroblastoma, lung cancer, non-Hodgkin's lymphoma, and anaplastic large cell lymphoma, among others. For example, a chromosomal rearrangement that generates a fusion gene resulting in the juxtaposition of the N-terminal region of nucleophosmin (NPM) with the kinase domain of ALK are known to be associated with non-Hodgkin's lymphoma (Morris, SW (1994) *Science* 263:1281-1284). Thus, the DCTN1-ALK fusions disclosed herein (*e.g.*, the 5'-DCTN1 to 3'-ALK fusions that include an ALK tyrosine kinase domain) are likely to be associated with cancers, *e.g.*, lung cancer.

Accordingly, the invention provides, at least in part, isolated DCTN1-ALK nucleic acid molecules, nucleic acid constructs, host cells containing the nucleic acid molecules; purified DCTN1-ALK polypeptides and binding agents; detection reagents (*e.g.*, probes, primers, antibodies, kits); screening assays for identifying novel kinase inhibitors; as well as methods, assays and kits for evaluating, identifying, assessing and/or treating a subject having a cancer, *e.g.*, a cancer having a DCTN1-ALK fusion disclosed herein. The compositions and methods identified herein can be used, for example, to identify new DCTN1-ALK inhibitors; to treat or prevent a cancer; as well as in methods or assays for evaluating a cancer (*e.g.*, evaluating one or more of: cancer progression, cancer treatment response or resistance to cancer treatment; selecting a treatment option, stratifying a patient population, and/or more effectively monitoring, treating or preventing a cancer).

LMNA-NTRK1 fusions

The term "LMNA-NTRK1" or "LMNA-NTRK1 fusion" is used generically herein, and includes any fusion molecule (*e.g.*, gene, gene product (*e.g.*, cDNA, mRNA, polypeptide), and variant thereof) that includes a fragment of LMNA and a fragment of NTRK1, in any configuration, including, *e.g.*, a 5'LMNA-3'NTRK1 or a 5'NTRK1-3'LMNA fusion molecule.

In one embodiment, a LMNA-NTRK1 fusion includes an in-frame fusion of an exon of LMNA (*e.g.*, one more exons of LMNA, such as exons 1 or 2, or a fragment

IDT Exhibit 1020.2979

thereof) and an exon of NTRK1 (*e.g.*, one or more exons encoding an NTRK1 tyrosine kinase domain or a fragment thereof). In one embodiment, the fusion is produced by a translocation having a breakpoint in intron 2 of LMNA and in intron 10 of NTRK1. For example, the LMNA-NTRK1 fusion can include an in-frame fusion of at least exon 2 of LMNA or a fragment thereof (*e.g.*, one or more of exons 1-2 of LMNA or a fragment thereof) with at least exon 11 of LMNA or a fragment thereof (*e.g.*, one or more of exons 11-41 of NTRK1 or a fragment thereof).

In certain embodiments, the LMNA-NTRK1 fusion is in a 5'- LMNA to 3'-NTRK1 configuration referred to herein as "5'LMNA-3'NTRK."

The NTRK1 receptor tyrosine kinase (also referred to as the TrkA receptor) is known to be associated with cancerous phenotypes including colon tumors and papillary thyroid carcinomas, among others. For example, a chromosomal rearrangement that generates a fusion gene resulting in the juxtaposition of the TFG gene with the kinase binding domain of NTRK1 are known to be associated with papillary thyroid carcinomas. Other NTRK1 rearrangements resulting in fusions include, *e.g.,* NTRK1-TPM3 and TPR-NTRK1. NTRK1 fusion genes are reviewed in *e.g.,* Greco, A. *et al.*, *Mol Cell Endocrinol* (2010) 321(1):44-49.

Accordingly, the invention provides, at least in part, isolated LMNA-NTRK1 nucleic acid molecules, nucleic acid constructs, host cells containing the nucleic acid molecules; purified LMNA-NTRK1 polypeptides and binding agents; detection reagents (*e.g.*, probes, primers, antibodies, kits); screening assays for identifying novel kinase inhibitors; as well as methods, assays and kits for evaluating, identifying, and/or assessing a subject having a cancer, *e.g.*, a cancer having a LMNA-NTRK1 fusion disclosed herein.

Certain terms are first defined. Additional terms are defined throughout the specification.

As used herein, the articles "a" and "an" refer to one or to more than one (*e.g.*, to at least one) of the grammatical object of the article.

The term "or" is used herein to mean, and is used interchangeably with, the term "and/or", unless context clearly indicates otherwise.

"About" and "approximately" shall generally mean an acceptable degree of error for the quantity measured given the nature or precision of the measurements. Exemplary degrees

IDT Exhibit 1020.2980

of error are within 20 percent (%), typically, within 10%, and more typically, within 5% of a given value or range of values.

"Acquire" or "acquiring" as the terms are used herein, refer to obtaining possession of a physical entity, or a value, *e.g.*, a numerical value, by "directly acquiring" or "indirectly acquiring" the physical entity or value. "Directly acquiring" means performing a process (*e.g.*, performing a synthetic or analytical method) to obtain the physical entity or value. "Indirectly acquiring" refers to receiving the physical entity or value from another party or source (*e.g.*, a third party laboratory that directly acquired the physical entity or value). Directly acquiring a physical entity includes performing a process that includes a physical change in a physical substance, *e.g.*, a starting material. Exemplary changes include making a physical entity from two or more starting materials, shearing or fragmenting a substance, separating or purifying a substance, combining two or more separate entities into a mixture, performing a chemical reaction that includes breaking or forming a covalent or non-covalent bond. Directly acquiring a value includes performing a process that includes a physical change in a sample or another substance, *e.g.*, performing an analytical process which includes a physical change in a substance, *e.g.*, a sample, analyte, or reagent (sometimes referred to herein as "physical analysis"), performing an analytical method, *e.g.*, a method which includes one or more of the following: separating or purifying a substance, *e.g.*, an analyte, or a fragment or other derivative thereof, from another substance; combining an analyte, or fragment or other derivative thereof, with another substance, *e.g.*, a buffer, solvent, or reactant; or changing the structure of an analyte, or a fragment or other derivative thereof, *e.g.*, by breaking or forming a covalent or non-covalent bond, between a first and a second atom of the analyte; or by changing the structure of a reagent, or a fragment or other derivative thereof, *e.g.*, by breaking or forming a covalent or non-covalent bond, between a first and a second atom of the reagent.

"Acquiring a sequence" as the term is used herein, refers to obtaining possession of a nucleotide sequence or amino acid sequence, by "directly acquiring" or "indirectly acquiring" the sequence. "Directly acquiring a sequence" means performing a process (*e.g.*, performing a synthetic or analytical method) to obtain the sequence, such as performing a sequencing method (*e.g.*, a Next Generation Sequencing (NGS) method). "Indirectly acquiring a sequence" refers to receiving information or knowledge of, or

IDT Exhibit 1020.2981

receiving, the sequence from another party or source (*e.g.*, a third party laboratory that directly acquired the sequence). The sequence acquired need not be a full sequence, *e.g.*, sequencing of at least one nucleotide, or obtaining information or knowledge, that identifies a DCTN1-ALK fusion or LMNA-NTRK1 fusion disclosed herein as being present in a subject constitutes acquiring a sequence.

Directly acquiring a sequence includes performing a process that includes a physical change in a physical substance, *e.g.*, a starting material, such as a tissue sample, *e.g.*, a biopsy, or an isolated nucleic acid (*e.g.*, DNA or RNA) sample. Exemplary changes include making a physical entity from two or more starting materials, shearing or fragmenting a substance, such as a genomic DNA fragment; separating or purifying a substance (*e.g.*, isolating a nucleic acid sample from a tissue); combining two or more separate entities into a mixture, performing a chemical reaction that includes breaking or forming a covalent or non-covalent bond. Directly acquiring a value includes performing a process that includes a physical change in a sample or another substance as described above.

"Acquiring a sample" as the term is used herein, refers to obtaining possession of a sample, *e.g.*, a tissue sample or nucleic acid sample, by "directly acquiring" or "indirectly acquiring" the sample. "Directly acquiring a sample" means performing a process (*e.g.*, performing a physical method such as a surgery or extraction) to obtain the sample. "Indirectly acquiring a sample" refers to receiving the sample from another party or source (*e.g.*, a third party laboratory that directly acquired the sample). Directly acquiring a sample includes performing a process that includes a physical change in a physical substance, *e.g.*, a starting material, such as a tissue, *e.g.*, a tissue in a human patient or a tissue that has was previously isolated from a patient. Exemplary changes include making a physical entity from a starting material, dissecting or scraping a tissue; separating or purifying a substance (*e.g.*, a sample tissue or a nucleic acid sample); combining two or more separate entities into a mixture; performing a chemical reaction that includes breaking or forming a covalent or non-covalent bond. Directly acquiring a sample includes performing a process that includes a physical change in a sample or another substance, *e.g.*, as described above.

"Binding entity" means any molecule to which molecular tags can be directly or indirectly attached that is capable of specifically binding to an analyte. The binding

IDT Exhibit 1020.2982

entity can be an affinity tag on a nucleic acid sequence. In certain embodiments, the binding entity allows for separation of the nucleic acid from a mixture, such as an avidin molecule, or an antibody that binds to the hapten or an antigen-binding fragment thereof. Exemplary binding entities include, but are not limited to, a biotin molecule, a hapten, an antibody, an antibody binding fragment, a peptide, and a protein.

"Complementary" refers to sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds ("base pairing") with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. In certain embodiments, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. In other embodiments, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

The term "cancer" or "tumor" is used interchangeably herein. These terms refer to the presence of cells possessing characteristics typical of cancer-causing cells, such as uncontrolled proliferation, immortality, metastatic potential, rapid growth and proliferation rate, and certain characteristic morphological features. Cancer cells are often in the form of a tumor, but such cells can exist alone within an animal, or can be a non-tumorigenic cancer cell, such as a leukemia cell. These terms include a solid tumor, a soft tissue tumor, or a metastatic lesion. As used herein, the term "cancer" includes premalignant, as well as malignant cancers. In certain embodiments, the cancer is a solid tumor, a soft tissue tumor, or a metastatic lesion.

IDT Exhibit 1020.2983

The term “neoplasm” or “neoplastic” cell refers to an abnormal proliferative stage, *e.g.*, a hyperproliferative stage, in a cell or tissue that can include a benign, pre-malignant, malignant (cancer) or metastatic stage.

Cancer is "inhibited" if at least one symptom of the cancer is alleviated, terminated, slowed, or prevented. As used herein, cancer is also "inhibited" if recurrence or metastasis of the cancer is reduced, slowed, delayed, or prevented.

"Chemotherapeutic agent" means a chemical substance, such as a cytotoxic or cytostatic agent, that is used to treat a condition, particularly cancer.

As used herein, “cancer therapy” and “cancer treatment” are synonymous terms.

As used herein, “chemotherapy” and “chemotherapeutic” and “chemotherapeutic agent” are synonymous terms.

The terms “homology” or “identity,” as used interchangeably herein, refer to sequence similarity between two polynucleotide sequences or between two polypeptide sequences, with identity being a more strict comparison. The phrases "percent identity or homology" and "% identity or homology" refer to the percentage of sequence similarity found in a comparison of two or more polynucleotide sequences or two or more polypeptide sequences. "Sequence similarity" refers to the percent similarity in base pair sequence (as determined by any suitable method) between two or more polynucleotide sequences. Two or more sequences can be anywhere from 0-100% similar, or any integer value there between. Identity or similarity can be determined by comparing a position in each sequence that can be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same nucleotide base or amino acid, then the molecules are identical at that position. A degree of similarity or identity between polynucleotide sequences is a function of the number of identical or matching nucleotides at positions shared by the polynucleotide sequences. A degree of identity of polypeptide sequences is a function of the number of identical amino acids at positions shared by the polypeptide sequences. A degree of homology or similarity of polypeptide sequences is a function of the number of amino acids at positions shared by the polypeptide sequences. The term “substantially identical,” as used herein, refers to an identity or homology of at

IDT Exhibit 1020.2984

least 75%, at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more.

"Likely to" or "increased likelihood," as used herein, refers to an increased probability that an item, object, thing or person will occur. Thus, in one example, a subject that is likely to respond to treatment with a kinase inhibitor, alone or in combination, has an increased probability of responding to treatment with the inhibitor alone or in combination, relative to a reference subject or group of subjects.

"Unlikely to" refers to a decreased probability that an event, item, object, thing or person will occur with respect to a reference. Thus, a subject that is unlikely to respond to treatment with a kinase inhibitor, alone or in combination, has a decreased probability of responding to treatment with a kinase inhibitor, alone or in combination, relative to a reference subject or group of subjects.

"Sequencing" a nucleic acid molecule requires determining the identity of at least 1 nucleotide in the molecule. In embodiments, the identity of less than all of the nucleotides in a molecule are determined. In other embodiments, the identity of a majority or all of the nucleotides in the molecule is determined.

"Next-generation sequencing or NGS or NG sequencing" as used herein, refers to any sequencing method that determines the nucleotide sequence of either individual nucleic acid molecules (*e.g.*, in single molecule sequencing) or clonally expanded proxies for individual nucleic acid molecules in a highly parallel fashion (*e.g.*, greater than 10^5 molecules are sequenced simultaneously). In one embodiment, the relative abundance of the nucleic acid species in the library can be estimated by counting the relative number of occurrences of their cognate sequences in the data generated by the sequencing experiment. Next generation sequencing methods are known in the art, and are described, *e.g.*, in Metzker, M. (2010) *Nature Biotechnology Reviews* 11:31-46, incorporated herein by reference. Next generation sequencing can detect a variant present in less than 5% of the nucleic acids in a sample.

"Sample," "tissue sample," "patient sample," "patient cell or tissue sample" or "specimen" each refers to a collection of similar cells obtained from a tissue of a subject or patient. The source of the tissue sample can be solid tissue as from a fresh, frozen and/or preserved organ, tissue sample, biopsy, or aspirate; blood or any blood constituents; bodily fluids such as cerebral spinal fluid, amniotic fluid, peritoneal fluid or

IDT Exhibit 1020.2985

interstitial fluid; or cells from any time in gestation or development of the subject. The tissue sample can contain compounds that are not naturally intermixed with the tissue in nature such as preservatives, anticoagulants, buffers, fixatives, nutrients, antibiotics or the like. In one embodiment, the sample is preserved as a frozen sample or as formaldehyde- or paraformaldehyde-fixed paraffin-embedded (FFPE) tissue preparation. For example, the sample can be embedded in a matrix, *e.g.*, an FFPE block or a frozen sample.

A "tumor nucleic acid sample" as used herein, refers to nucleic acid molecules from a tumor or cancer sample. Typically, it is DNA, *e.g.*, genomic DNA, or cDNA derived from RNA, from a tumor or cancer sample. In certain embodiments, the tumor nucleic acid sample is purified or isolated (*e.g.*, it is removed from its natural state).

A "control" or "reference" "nucleic acid sample" as used herein, refers to nucleic acid molecules from a control or reference sample. Typically, it is DNA, *e.g.*, genomic DNA, or cDNA derived from RNA, not containing the alteration or variation in the gene or gene product, *e.g.*, not containing a DCTN1-ALK fusion. In certain embodiments, the reference or control nucleic acid sample is a wild type or a non-mutated sequence. In certain embodiments, the reference nucleic acid sample is purified or isolated (*e.g.*, it is removed from its natural state). In other embodiments, the reference nucleic acid sample is from a non-tumor sample, *e.g.*, a blood control, a normal adjacent tumor (NAT), or any other non-cancerous sample from the same or a different subject.

"Adjacent to the interrogation position," as used herein, means that a site sufficiently close such that a detection reagent complementary with the site can be used to distinguish between a mutation, e.g., a mutation described herein, and a reference sequence, e.g., a non-mutant or wild-type sequence, in a target nucleic acid. Directly adjacent, as used herein, is where 2 nucleotides have no intervening nucleotides between them.

"Associated mutation," as used herein, refers to a mutation within a preselected distance, in terms of nucleotide or primary amino acid sequence, from a definitional mutation, *e.g.*, a mutant as described herein, *e.g.*, a translocation, breakpoint or fusion molecule described herein. In embodiments, the associated mutation is within n, wherein n is 2, 5, 10, 20, 30, 50, 100, or 200 nucleotides from the definitional mutation (n does not include the nucleotides defining the associated

IDT Exhibit 1020.2986

and definitional mutations). In embodiments, the associated mutation is a translocation mutation.

"Interrogation position," as used herein, comprises at least one nucleotide (or, in the case of polypeptides, an amino acid residue) which corresponds to a nucleotide (or amino acid residue) that is mutated in a mutation of interest, *e.g.*, a mutation being identified, or in a nucleic acid (or protein) being analyzed, *e.g.*, sequenced, or recovered. By way of example, the interrogation position in the breakpoint shown in FIG. 1 (SEQ ID NO:5), includes one, two, or more nucleotide positions at the junction site. In other embodiments, the interrogation position in the breakpoint shown in FIG. 3 (SEQ ID NO:13), includes one, two, or more nucleotide positions at the junction site.

A "reference sequence," as used herein, *e.g.*, as a comparator for a mutant sequence, is a sequence which has a different nucleotide or amino acid at an interrogation position than does the mutant(s) being analyzed. In an embodiment, the reference sequence is wild-type for at least the interrogation position.

Headings, *e.g.*, (a), (b), (i) etc, are presented merely for ease of reading the specification and claims. The use of headings in the specification or claims does not require the steps or elements be performed in alphabetical or numerical order or the order in which they are presented.

Various aspects featured in the invention are described in further detail below. Additional definitions are set out throughout the specification.

Isolated Nucleic Acid Molecules

One aspect featured in the invention pertains to isolated nucleic acid molecules that include a DCTN1-ALK fusion or an LMNA-NTRK1 fusion, including nucleic acids which encode a DCTN1-ALK or LMNA-NTRK1 fusion polypeptide or a portion of such a polypeptide. The nucleic acid molecules include those nucleic acid molecules which reside in genomic regions identified herein. As used herein, the term "nucleic acid molecule" includes DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded; in certain embodiments the nucleic acid molecule is double-stranded DNA.

IDT Exhibit 1020.2987

Isolated nucleic acid molecules also include nucleic acid molecules sufficient for use as hybridization probes or primers to identify nucleic acid molecules that correspond to a DCTN1-ALK fusion or an LMNA-NTRK1 fusion, *e.g.*, those suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. In certain embodiments, an "isolated" nucleic acid molecule is free of sequences (such as protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, less than about 4 kB, less than about 3 kB, less than about 2 kB, less than about 1 kB, less than about 0.5 kB or less than about 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

The language "substantially free of other cellular material or culture medium" includes preparations of nucleic acid molecule in which the molecule is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, nucleic acid molecule that is substantially free of cellular material includes preparations of nucleic acid molecule having less than about 30%, less than about 20%, less than about 10%, or less than about 5% (by dry weight) of other cellular material or culture medium.

A fusion nucleic acid molecule can be isolated using standard molecular biology techniques and the sequence information in the database records described herein. Using all or a portion of such nucleic acid sequences, fusion nucleic acid molecules as described herein can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual, 2nd ed.,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

IDT Exhibit 1020.2988

A fusion nucleic acid molecule (*e.g.*, DCTN1-ALK or LMNA-NTRK1 fusion) can be amplified using cDNA, mRNA, or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid molecules so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule featured in the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In another embodiment, a fusion nucleic acid molecule (*e.g.*, DCTN1-ALK or LMNA-NTRK1 fusion) comprises a nucleic acid molecule which has a nucleotide sequence complementary to the nucleotide sequence of the fusion nucleic acid molecule or to the nucleotide sequence of a nucleic acid encoding a fusion protein. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence or which encodes a DCTN1-ALK or LMNA-NTRK1 fusion polypeptide. Such nucleic acid molecules can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, at least about 15, at least about 25, at least about 50, at least about 75, at least about 100, at least about 125, at least about 150, at least about 175, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 500, at least about 600, at least about 700, at least about 800, at least about 900, at least about 1 kb, at least about 2 kb, at least about 3 kb, at least about 4 kb, at least about 5 kb, at least about 6 kb, at least about 7 kb, at least about 8 kb, at least about 9 kb, at least about 10 kb, at least about 15 kb, at least about 20 kb, at least about 25 kb, at least about 30 kb, at least about 35 kb, at least about 40 kb, at least about 45 kb, at least about 50 kb, at least about 60 kb, at least about 70 kb, at least about 80 kb, at least about 90 kb, at least about 100 kb, at least about 200 kb, at least about 300 kb, at least about 400 kb,

IDT Exhibit 1020.2989

at least about 500 kb, at least about 600 kb, at least about 700 kb, at least about 800 kb, at least about 900 kb, at least about 1 mb, at least about 2 mb, at least about 3 mb, at least about 4 mb, at least about 5 mb, at least about 6 mb, at least about 7 mb, at least about 8 mb, at least about 9 mb, at least about 10 mb or more consecutive nucleotides of a DCTN1-ALK fusion nucleic acid.

The invention further encompasses nucleic acid molecules that are substantially identical to the gene mutations and/or gene products described herein, *e.g.*, DCTN1-ALK fusion having a nucleotide sequence of SEQ ID NO:6, or an amino acid sequence of SEQ ID NO:7, such that they are at least 70%, at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or greater. The invention further encompasses nucleic acid molecules that are substantially identical to the gene mutations and/or gene products described herein, *e.g.*, LMNA-NTRK1 fusion having a nucleotide sequence of SEQ ID NO:9, or an amino acid sequence of SEQ ID NO:10, such that they are at least 70%, at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or greater.

In other embodiments, the invention further encompasses nucleic acid molecules that are substantially homologous to the DCTN1-ALK or LMNA-NTRK1 fusion gene mutations and/or gene products described herein, such that they differ by only or at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600 nucleotides or any range in between.

In another embodiment, an isolated DCTN1-ALK fusion nucleic acid molecule or LMNA-NTRK1 fusion nucleic acid molecule is at least 7, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 90, at least 95, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250,

IDT Exhibit 1020.2990

at least 300, at least 350, at least 400, at least 450, at least 550, at least 650, at least 700, at least 800, at least 900, at least 1000, at least 1200, at least 1400, at least 1600, at least 1800, at least 2000, at least 2200, at least 2400, at least 2600, at least 2800, at least 3000, or more nucleotides in length and hybridizes under stringent conditions to a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule or to a nucleic acid molecule encoding a protein corresponding to a marker featured in the invention.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, or at least 85% identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). Another, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C.

The invention also includes molecular beacon nucleic acid molecules having at least one region which is complementary to a DCTN1-ALK fusion or LMNA-NTRK1 fusion nucleic acid molecule, such that the molecular beacon is useful for quantitating the presence of the nucleic acid molecule featured in the invention in a sample. A "molecular beacon" nucleic acid is a nucleic acid molecule comprising a pair of complementary regions and having a fluorophore and a fluorescent quencher associated therewith. The fluorophore and quencher are associated with different portions of the nucleic acid in such an orientation that when the complementary regions are annealed with one another, fluorescence of the fluorophore is quenched by the quencher. When the complementary regions of the nucleic acid molecules are not annealed with one another, fluorescence of the fluorophore is quenched to a lesser degree. Molecular beacon nucleic acid molecules are described, for example, in U.S. Patent 5,876,930.

IDT Exhibit 1020.2991

Probes

The invention also provides isolated nucleic acid molecules useful as probes. Such nucleic acid probes can be designed based on the sequence of a DCTN1-ALK fusion or a LMNA-NTRK1 fusion.

Probes based on the sequence of a fusion nucleic acid molecule as described herein can be used to detect transcripts or genomic sequences corresponding to one or more markers featured in the invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a test kit for identifying cells or tissues which express the fusion protein (*e.g.*, DCTN1-ALK or LMNA-NTRK1), such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, *e.g.*, detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

Probes featured in the invention include those that will specifically hybridize to a gene sequence described in the Example, *e.g.*, a DCTN1-ALK fusion or a LMNA-NTRK1 fusion. Typically these probes are 12 to 20, *e.g.*, 17 to 20 nucleotides in length (longer for large insertions) and have the nucleotide sequence corresponding to the region of the mutations at their respective nucleotide locations on the gene sequence. Such molecules can be labeled according to any technique known in the art, such as with radiolabels, fluorescent labels, enzymatic labels, sequence tags, biotin, other ligands, etc. As used herein, a probe that "specifically hybridizes" to a fusion gene sequence will hybridize under high stringency conditions.

A probe will typically contain one or more of the specific mutations described herein. Typically, a nucleic acid probe will encompass only one mutation. Such molecules may be labeled and can be used as allele-specific probes to detect the mutation of interest.

In one aspect, the invention features a probe or probe set that specifically hybridizes to a nucleic acid comprising an inversion resulting in a DCTN1-ALK fusion. In another aspect, the invention features a probe or probe set that specifically hybridizes to a nucleic acid comprising a deletions resulting in an LMNA-NTRK1 fusion.

IDT Exhibit 1020.2992

Isolated pairs of allele specific oligonucleotide probes are also provided, where the first probe of the pair specifically hybridizes to the mutant allele, and the second probe of the pair specifically hybridizes to the wildtype allele. For example, in one exemplary probe pair, one probe will recognize the fusion junction in the DCTN1-ALK fusion, and the other probe will recognize a sequence downstream or upstream of DCTN1 or ALK, neither of which includes the fusion junction. These allele-specific probes are useful in detecting an ALK somatic mutation in a tumor sample, *e.g.*, melanoma sample. In a similar manner, probe pairs can be designed and produced for the LMNA-NTRK1 fusion and are useful in detecting an NTRK1 somatic mutation in a tumor sample.

Primers

The invention also provides isolated nucleic acid molecules useful as primers.

The term "primer" as used herein refers to a sequence comprising two or more deoxyribonucleotides or ribonucleotides, *e.g.*, more than three, and more than eight, or at least 20 nucleotides of a gene described in the Example, where the sequence corresponds to a sequence flanking one of the mutations or a wild type sequence of a gene identified in the Example, e.g., a DCTN1, ALK, LMNA, or NRTK1 gene. Primers may be used to initiate DNA synthesis via the PCR (polymerase chain reaction) or a sequencing method. Primers featured in the invention include the sequences recited and complementary sequences which would anneal to the opposite DNA strand of the sample target. Since both strands of DNA are complementary and mirror images of each other, the same segment of DNA will be amplified.

Primers can be used to sequence a nucleic acid, *e.g.*, an isolated nucleic acid described herein, such as by an NGS method, or to amplify a gene described in the Example, such as by PCR. The primers can specifically hybridize, for example, to the ends of the exons or to the introns flanking the exons. The amplified segment can then be further analyzed for the presence of the mutation such as by a sequencing method. The primers are useful in directing amplification of a target polynucleotide prior to sequencing. In another aspect, the invention features a pair of oligonucleotide primers that amplify a region that contains or is adjacent to a fusion junction identified in the Example. Such primers are useful in directing amplification of a target region that includes a fusion junction identified in the Example, *e.g.*, prior to sequencing. The

IDT Exhibit 1020.2993

primer typically contains 12 to 20, or 17 to 20, or more nucleotides, although a primer may contain fewer nucleotides.

A primer is typically single stranded, *e.g.*, for use in sequencing or amplification methods, but may be double stranded. If double stranded, the primer may first be treated to separate its strands before being used to prepare extension products. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including applications (*e.g.*, amplification method), temperature, buffer, and nucleotide composition. A primer typically contains 12-20 or more nucleotides, although a primer may contain fewer nucleotides.

Primers are typically designed to be "substantially" complementary to each strand of a genomic locus to be amplified. Thus, the primers must be sufficiently complementary to specifically hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5' and 3' sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

The term "substantially complementary to" or "substantially the sequence" refers to sequences that hybridize to the sequences provided under stringent conditions and/or sequences having sufficient homology with a sequence comprising a fusion junction identified in the Example, or the wildtype counterpart sequence, such that the allele specific oligonucleotides hybridize to the sequence. In one embodiment, a sequence is substantially complementary to a fusion junction in an inversion event, *e.g.*, to a fusion junction in SEQ ID NO:6. "Substantially the same" as it refers to oligonucleotide sequences also refers to the functional ability to hybridize or anneal with sufficient specificity to distinguish between the presence or absence of the mutation. This is measurable by the temperature of melting being sufficiently different to permit easy identification of whether the oligonucleotide is binding to the normal or mutant gene sequence identified in the Example.

In one aspect, the invention features a primer or primer set for amplifying a nucleic acid comprising an inversion resulting in a DCTN1-ALK fusion. In another aspect, the invention features a primer or primer set for amplifying a nucleic acid comprising a deletion resulting in an LMNA-NTRK1 fusion.

IDT Exhibit 1020.2994

Isolated pairs of allele specific oligonucleotide primer are also provided, where the first primer of the pair specifically hybridizes to the mutant allele, and the second primer of the pair specifically hybridizes to a sequence upstream or downstream of a mutation, or a fusion junction resulting from, e.g., an inversion, duplication, deletion, insertion or translocation. For example, in one exemplary primer pair, one probe will recognize a DCTN1-ALK translocation, such as by hybridizing to a sequence at the fusion junction between the DCTN1 and ALK transcripts, and the other primer will recognize a sequence upstream or downstream of the fusion junction. These allele-specific primers are useful for amplifying a DCTN1-ALK fusion sequence from a tumor sample, *e.g.*, a skin biopsy, such as a skin biopsy from a suspected melanoma. Similarly, in one exemplary primer pair, one probe will recognize a LMNA-NTRK1 fusion, such as by hybridizing to a sequence at the fusion junction between the LMNA and NTRK1 transcripts, and the other primer will recognize a sequence upstream or downstream of the fusion junction. These allele-specific primers are useful for amplifying a LMNA-NTRK1fusion sequence from a tumor sample, *e.g.*, a skin biopsy, such as a skin biopsy from a suspected melanoma.

In another exemplary primer pair, one primer can recognize an ALK-DCTN1 translocation (*e.g.*, the reciprocal of the DCTN1-ALK translocation), such as by hybridizing to a sequence at the fusion junction between the ALK and DCTN1 transcripts, and the other primer will recognize a sequence upstream or downstream of the fusion junction. These allele-specific primers are useful for amplifying a ALK-DCTN1 fusion sequence from a tumor sample, *e.g.*, a melanoma or skin biopsy sample.

In addition, an exemplary primer pair can be designed such that one primer recognizes an NTRK1-LMNA fusion (*e.g.,* the reciprocal of the LMNA-NTRK1 fusion), such as by hybridizing to a sequence at the fusion junction between the NTRK1 and LMNA transcripts, and the other primer will recognize a sequence upstream or downstream of the fusion junction. These allele-specific primers are useful for amplifying a NTRK1-LMNA fusion sequence from a tumor sample, *e.g.*, a melanoma or skin biopsy sample.

Primers can be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and

IDT Exhibit 1020.2995

may be synthesized as described by Beaucage, *et al.*, Tetrahedron Letters, 22:1859-1862, (1981). One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Pat. No. 4,458,066.

An oligonucleotide probe or primer that hybridizes to a mutant or wildtype allele is said to be the complement of the allele. As used herein, a probe exhibits "complete complementarity" when every nucleotide of the probe is complementary to the corresponding nucleotide of the allele. Two polynucleotides are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the polynucleotides are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are known to those skilled in the art and can be found, for example in *Molecular Cloning: A Laboratory Manual*, 3rd edition Volumes 1, 2, and 3. J. F. Sambrook, D. W. Russell, and N. Irwin, Cold Spring Harbor Laboratory Press, 2000.

Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of a probe to hybridize to an allele. Thus, in order for a polynucleotide to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed. Appropriate stringency conditions which promote DNA hybridization are, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 50°C. Such conditions are known to those skilled in the art and can be found, for example in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). Salt concentration and temperature in the wash step can be adjusted to alter hybridization stringency. For example, conditions may vary from low stringency of about 2.0 X SSC at 40°C to moderately stringent conditions of about 2.0 X SSC at 50°C to high stringency conditions of about 0.2 X SSC at 50°C.

Fusion Proteins and Antibodies

One aspect featured in the invention pertains to purified fusion polypeptides, and biologically active portions thereof. In one embodiment, the native DCTN1-ALK

IDT Exhibit 1020.2996

or NTRK1-LMNA fusion polypeptide can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, a DCTN1-ALK or NTRK1-LMNA fusion polypeptide is produced by recombinant DNA techniques. Alternative to recombinant expression, a DCTN1-ALK or NTRK1-LMNA fusion polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about 30%, less than about 20%, less than about 10%, or less than about 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it can be substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, less than about 10%, or less than about 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it can substantially be free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, less than about 20%, less than about 10%, less than about 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

Biologically active portions of a fusion polypeptide include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the fusion protein, which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein, *e.g.*, a kinase activity *e.g.*, an ALK kinase or NTRK1 kinase activity. A biologically active portion of a protein featured in the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover,

IDT Exhibit 1020.2997

other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide.

In certain embodiments, the DCTN1-ALK fusion polypeptide or the NTRK1-LMNA fusion polypeptide has an amino acid sequence of a protein encoded by a nucleic acid molecule disclosed herein. Other useful proteins are substantially identical (*e.g.*, at least 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 86, at least 87, at least 88, at least 89, at least 90, at least 91, at least 92, at least 93, at least 94, at least 95, at least 96, at least 97, at least 98, at least 99, at least 99.5% or greater) to one of these sequences and retain the functional activity of the protein of the corresponding full-length protein yet differ in amino acid sequence.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Another, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules featured in the invention. BLAST protein

IDT Exhibit 1020.2998

searches can be performed with the XBLAST program, score = 50, word length = 3 to obtain amino acid sequences homologous to protein molecules featured in the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov. Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *Comput Appl Biosci,* 4:11-7. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

An isolated fusion polypeptide (*e.g.,* DCTN1-ALK or LMNA-NTRK1), or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length fusion polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein featured in the invention comprises at least 8 (or at least 10, at least 15, at least 20, or at least 30 or more) amino acid residues of the amino acid sequence of one of the polypeptides featured in the invention, and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with a marker featured in the invention to which the protein corresponds. Exemplary

IDT Exhibit 1020.2999

epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, *e.g.*, hydrophilic regions. Hydrophobicity sequence analysis, hydrophilicity sequence analysis, or similar analyses can be used to identify hydrophilic regions.

An immunogen typically is used to prepare antibodies by immunizing a suitable (*i.e.*, immunocompetent) subject such as a rabbit, goat, mouse, or other mammal or vertebrate. An appropriate immunogenic preparation can contain, for example, recombinantly-expressed or chemically-synthesized polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar immunostimulatory agent.

Accordingly, another aspect featured in the invention pertains to antibodies directed against a DCTN1-ALK fusion polypeptide. In one embodiment, the antibody molecule specifically binds to DCTN1-ALK fusion, *e.g.*, specifically binds to an epitope formed by the DCTN1-ALK fusion. In embodiments the antibody can distinguish wild type ALK (or DCTN1) from DCTN1-ALK.

Another aspect featured in the invention provides antibodies directed against a LMNA-NTRK1 fusion polypeptide are contemplated. In one embodiment, the antibody molecule specifically binds to LMNA-NTRK1fusion, *e.g.*, specifically binds to an epitope formed by the LMNA-NTRK1fusion. In embodiments the antibody can distinguish wild type NTRK1 (or LNMA) from LMNA-NTRK1.

The terms "antibody" and "antibody molecule" as used interchangeably herein refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds an antigen, such as a polypeptide featured in the invention. A molecule which specifically binds to a given polypeptide featured in the invention is a molecule which binds the polypeptide, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the polypeptide. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term “monoclonal antibody” or “monoclonal antibody composition,” as used herein, refers to a population of antibody molecules that contain only one

IDT Exhibit 1020.3000

species of an antigen binding site capable of immunoreacting with a particular epitope.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a fusion polypeptide as an immunogen. Antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (see Kozbor *et al.,* 1983, *Immunol. Today* 4:72), the EBV-hybridoma technique (see Cole *et al.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology,* Coligan *et al.* ed., John Wiley & Sons, New York, 1994). Hybridoma cells producing a monoclonal antibody are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275- 1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions can be made using standard recombinant DNA techniques. Such chimeric and humanized

IDT Exhibit 1020.3001

monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521- 3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

Completely human antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, *e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

An antibody directed against a DCTN1-ALK fusion polypeptide or a LMNA-NTRK1 fusion polypeptide (*e.g.*, a monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the marker (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the level and pattern of expression of the marker. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include, but are not limited to, various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include, but are not limited to, horseradish peroxidase, alkaline phosphatase,

IDT Exhibit 1020.3002

β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include, but are not limited to, streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include, but are not limited to, umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes, but is not limited to, luminol; examples of bioluminescent materials include, but are not limited to, luciferase, luciferin, and aequorin, and examples of suitable radioactive materials include, but are not limited to, ^{125}I, ^{131}I, ^{35}S or ^{3}H.

An antibody directed against a DCTN1-ALK fusion polypeptide can also be used diagnostically to monitor protein levels in tissues or body fluids (*e.g.,* in a tumor cell-containing body fluid) as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen.

Antigens and Vaccines

Embodiments featured in the invention include preparations, *e.g.,* antigenic preparations, of the entire fusion or a fragment thereof, *e.g.*, a fragment capable of raising antibodies specific to the fusion protein, e.g., a fusion junction containing fragment (collectively referred to herein as a fusion specific polypeptides or FSP). The preparation can include an adjuvant or other component.

An FSP can be used as an antigen or vaccine. For example, an FSP can be used as an antigen to immunize an animal, *e.g.,* a rodent, *e.g.,* a mouse or rat, rabbit, horse, goat, dog, or non-human primate, to obtain antibodies, *e.g.,* fusion protein specific antibodies. In an embodiment a fusion specific antibody molecule is an antibody molecule described herein, *e.g.,* a polyclonal. In other embodiments a fusion specific antibody molecule is monospecific, *e.g.*, monoclonal, human, humanized, chimeric or other monospecific antibody molecule. An anti-DCTN1-ALK fusion protein specific antibody molecule can be used to treat a subject having a cancer, *e.g.,* a cancer described herein, such as a melanoma.

Embodiments featured in the invention include vaccine preparations that comprise an FSP capable of stimulating an immune response in a subject, *e.g.*, by raising, in the subject, antibodies specific to the fusion protein. The vaccine

IDT Exhibit 1020.3003

preparation can include other components, *e.g.*, an adjuvant. The vaccine preparations can be used to treat a subject having cancer, *e.g.*, a cancer described herein.

Expression Vectors, Host Cells and Recombinant Cells

In another aspect, the invention includes vectors (*e.g.*, expression vectors), containing a nucleic acid encoding a DCTN1-ALK fusion polypeptide or encoding an LMNA-NTRK1 fusion polypeptide as described herein. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors include, *e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses.

A vector can include a fusion nucleic acid in a form suitable for expression of the nucleic acid in a host cell. Preferably the recombinant expression vector includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors can be introduced into host cells to thereby produce a fusion polypeptide, including fusion proteins or polypeptides encoded by nucleic acids as described herein, mutant forms thereof, and the like).

The term "recombinant host cell" (or simply "host cell" or "recombinant cell"), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell, but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

The recombinant expression vectors can be designed for expression of a fusion polypeptide (*e.g.*, LMNA-NTRK1 or DCTN1-ALK) in prokaryotic or eukaryotic cells. For

IDT Exhibit 1020.3004

example, polypeptides featured in the invention can be expressed in *E. coli*, insect cells (e.g., using baculovirus expression vectors), yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, (1990) *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified DCTN1-ALK or LMNA-NTRK1fusion polypeptides can be used in activity assays (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for DCTN1-ALK or LMNA-NTRK1 fusion polypeptides.

To maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., (1990) *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences can be carried out by standard DNA synthesis techniques.

IDT Exhibit 1020.3005

The fusion polypeptide expression vector can be a yeast expression vector, a vector for expression in insect cells, *e.g.,* a baculovirus expression vector or a vector suitable for expression in mammalian cells.

When used in mammalian cells, the expression vector's control functions can be provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40.

In another embodiment, the promoter is an inducible promoter, *e.g.,* a promoter regulated by a steroid hormone, by a polypeptide hormone (*e.g.,* by means of a signal transduction pathway), or by a heterologous polypeptide (*e.g.,* the tetracycline-inducible systems, "Tet-On" and "Tet-Off"; see, e.g., Clontech Inc., CA, Gossen and Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547, and Paillard (1989) *Human Gene Therapy* 9:983).

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.,* tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.,* the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.,* milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule featured in the invention cloned into the expression vector in an antisense orientation. Regulatory sequences (*e.g.,* viral promoters and/or enhancers) operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of

IDT Exhibit 1020.3006

cell types. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus.

Another aspect the invention provides a host cell which includes a nucleic acid molecule described herein, *e.g.*, a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule within a recombinant expression vector or a DCTN1-ALK or LMNA-NTRK1 fusion nucleic acid molecule containing sequences which allow it to homologous recombination into a specific site of the host cell's genome.

A host cell can be any prokaryotic or eukaryotic cell. For example, a fusion polypeptide can be expressed in bacterial cells (such as *E. coli*), insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells, *e.g.*, COS-7 cells, CV-1 origin SV40 cells; Gluzman (1981) *Cell* 23:175-182). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation.

A host cell can be used to produce (*e.g.*, express) a fusion polypeptide (*e.g.*, DCTN1-ALK or LMNA-NTRK1). Accordingly, the invention further provides methods for producing a fusion polypeptide using the host cells. In one embodiment, the method includes culturing the host cell (into which a recombinant expression vector encoding a polypeptide has been introduced) in a suitable medium such that the fusion polypeptide is produced. In another embodiment, the method further includes isolating a fusion polypeptide from the medium or the host cell.

In another aspect, the invention features, a cell or purified preparation of cells which include a DCTN1-ALK fusion transgene, or which otherwise misexpress DCTN1-ALK fusion. In another aspect, the invention features, a cell or purified preparation of cells which include a LMNA-NTRK1fusion transgene, or which otherwise misexpress LMNA-NTRK1fusion.

The cell preparation can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In embodiments, the cell or cells include a DCTN1-ALK fusion transgene, *e.g.*, a heterologous form of a DCTN1-ALK fusion, *e.g.*, a

IDT Exhibit 1020.3007

gene derived from humans (in the case of a non-human cell) or a LMNA-NTRK1 fusion transgene, *e.g.,* a heterologous form of a LMNA-NTRK1 fusion. The fusion transgene can be misexpressed, *e.g.,* overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene that mis-expresses an endogenous fusion, *e.g.,* a gene the expression of which is disrupted, *e.g.,* a knockout. Such cells can serve as a model for studying disorders that are related to mutated or mis-expressed fusion alleles (*e.g.*, cancers) or for use in drug screening, as described herein.

Therapeutic Methods

Alternatively, or in combination with the methods described herein, the invention features a method of treating a neoplasm, a cancer or a tumor harboring a DCTN1-ALK fusion described herein. The methods include administering an anti-cancer agent, *e.g.*, a kinase inhibitor, alone or in combination, *e.g.*, in combination with other chemotherapeutic agents or procedures, in an amount sufficient to reduce or inhibit the tumor cell growth, and/or treat or prevent the cancer(s), in the subject.

"Treat," "treatment," and other forms of this word refer to the administration of a kinase inhibitor, alone or in combination with a second agent to impede growth of a cancer, to cause a cancer to shrink by weight or volume, to extend the expected survival time of the subject and or time to progression of the tumor or the like. In those subjects, treatment can include, but is not limited to, inhibiting tumor growth, reducing tumor mass, reducing size or number of metastatic lesions, inhibiting the development of new metastatic lesions, prolonged survival, prolonged progression-free survival, prolonged time to progression, and/or enhanced quality of life.

As used herein, unless otherwise specified, the terms "prevent," "preventing" and "prevention" contemplate an action that occurs before a subject begins to suffer from the re-growth of the cancer and/or which inhibits or reduces the severity of the cancer.

As used herein, and unless otherwise specified, a "therapeutically effective amount" of a compound is an amount sufficient to provide a therapeutic benefit in the treatment or management of the cancer, or to delay or minimize one or more symptoms associated with the cancer. A therapeutically effective amount of a compound means an amount of therapeutic agent, alone or in combination with other

IDT Exhibit 1020.3008

therapeutic agents, which provides a therapeutic benefit in the treatment or management of the cancer. The term "therapeutically effective amount" can encompass an amount that improves overall therapy, reduces or avoids symptoms or causes of the cancer, or enhances the therapeutic efficacy of another therapeutic agent.

As used herein, and unless otherwise specified, a "prophylactically effective amount" of a compound is an amount sufficient to prevent re-growth of the cancer, or one or more symptoms associated with the cancer, or prevent its recurrence. A prophylactically effective amount of a compound means an amount of the compound, alone or in combination with other therapeutic agents, which provides a prophylactic benefit in the prevention of the cancer. The term "prophylactically effective amount" can encompass an amount that improves overall prophylaxis or enhances the prophylactic efficacy of another prophylactic agent.

As used herein, the term "patient" or "subject" refers to an animal, typically a human (*i.e.,* a male or female of any age group, *e.g.,* a pediatric patient (*e.g,* infant, child, adolescent) or adult patient (*e.g.,* young adult, middle-aged adult or senior adult) or other mammal, such as a primate (*e.g.,* cynomolgus monkey, rhesus monkey); commercially relevant mammals such as cattle, pigs, horses, sheep, goats, cats, and/or dogs; and/or birds, including commercially relevant birds such as chickens, ducks, geese, and/or turkeys, that will be or has been the object of treatment, observation, and/or experiment. When the term is used in conjunction with administration of a compound or drug, then the patient has been the object of treatment, observation, and/or administration of the compound or drug.

In certain embodiments, the cancer includes, but is not limited to, a solid tumor, a soft tissue tumor, and a metastatic lesion (*e.g.*, a cancer as described herein). In one embodiment, the cancer is chosen from a lung cancer, a pancreatic cancer, melanoma, a colorectal cancer, an esophageal-gastric cancer, a thyroid cancer, or an adenocarcinoma. In other embodiment, the lung cancer is chosen from one or more of the following: non-small cell lung cancer (NSCLC), small cell lung cancer (SCLC), squamous cell carcinoma (SCC), adenocarcinoma of the lung, bronchogenic carcinoma, or a combination thereof. In one embodiment, the lung cancer is NSCLC or SCC.

IDT Exhibit 1020.3009

In other embodiments, the cancer is chosen from lung cancer, thyroid cancer, colorectal cancer, adenocarcinoma, melanoma, B cell cancer, breast cancer, bronchus cancer, cancer of the oral cavity or pharynx, cancer of hematological tissues, cervical cancer, colon cancer, esophageal cancer, esophageal-gastric cancer, gastric cancer, kidney cancer, liver cancer, multiple myeloma, ovarian cancer, pancreatic cancer, prostate cancer, salivary gland cancer, small bowel or appendix cancer, stomach cancer, testicular cancer, urinary bladder cancer, uterine or endometrial cancer, inflammatory myofibroblastic tumors, gastrointestinal stromal tumor (GIST), and the like.

In certain embodiments, the neoplasm or neoplastic cell is a benign, pre-malignant, malignant (cancer) or metastasis. In one embodiment, the neoplasm or cancer is a melanocytic neoplasm, a Spitz nevi, a Spitz tumor, a Spitzoid melanoma, a metastatic Spitz tumor, or a melanoma. In one embodiment, the Spitz tumor is metastatic, *e.g.*, localized to lymph nodes or widespread disease.

In one embodiment, the anti-cancer agent is a kinase inhibitor. For example, the kinase inhibitor is a multi-kinase inhibitor or an ALK-specific inhibitor. Exemplary kinase inhibitors include, but are not limited to, axitinib (AG013736), bosutinib (SKI-606), cediranib (RECENTINTM, AZD2171), dasatinib (SPRYCEL®, BMS-354825), erlotinib (TARCEVA®), gefitinib (IRESSA®), imatinib (Gleevec®, CGP57148B, STI-571), lapatinib (TYKERB®, TYVERB®), lestaurtinib (CEP-701), neratinib (HKI-272), nilotinib (TASIGNA®), semaxanib (semaxinib, SU5416), sunitinib (SUTENT®, SU11248), toceranib (PALLADIA®), vandetanib, vatalanib (PTK787, PTK/ZK), sorafenib (NEXAVAR®), ENMD-2076, PCI-32765, AC220, dovitinib lactate (TKI258, CHIR-258), BIBW 2992 (TOVOKTM), SGX523, PF-04217903, PF-02341066, PF-299804, BMS-777607, ABT-869, MP470, BIBF 1120 (VARGATEF®), AP24534, JNJ-26483327, MGCD265, DCC-2036, BMS-690154, CEP-11981, tivozanib (AV-951), OSI-930, MM-121, XL-184, XL-647, and XL228. DCTN1-ALK

In one embodiment, the kinase inhibitor is an ALK inhibitor including, but not limited to, TAE-684 (also referred to herein as "NVP-TAE694"), PF02341066 (also referred to herein as "crizotinib" or "1066"), AF-802, LDK-378, ASP-3026, CEP-37440, CEP-28122, CEP-108050, and AP26113. Additional examples of ALK kinase

IDT Exhibit 1020.3010

inhibitors are described in examples 3-39 of WO 2005016894 by Garcia-Echeverria C, *et al.*

In other embodiments, the anti-cancer agent is a DCTN1-ALK antagonist inhibits the expression of nucleic acid encoding DCTN1-ALK. Examples of such DCTN1-ALK antagonists include nucleic acid molecules, for example, antisense molecules, ribozymes, RNAi, triple helix molecules that hybridize to a nucleic acid encoding DCTN1-ALK, or a transcription regulatory region, and blocks or reduces mRNA expression of DCTN1-ALK.

In other embodiments, the kinase inhibitor is administered in combination with a second therapeutic agent or a different therapeutic modality, *e.g.*, anti-cancer agents, and/or in combination with surgical and/or radiation procedures.

By "in combination with," it is not intended to imply that the therapy or the therapeutic agents must be administered at the same time and/or formulated for delivery together, although these methods of delivery are within the scope of the invention. The pharmaceutical compositions can be administered concurrently with, prior to, or subsequent to, one or more other additional therapies or therapeutic agents. In general, each agent will be administered at a dose and/or on a time schedule determined for that agent. In will further be appreciated that the additional therapeutic agent utilized in this combination can be administered together in a single composition or administered separately in different compositions. The particular combination to employ in a regimen will take into account compatibility of the inventive pharmaceutical composition with the additional therapeutically active agent and/or the desired therapeutic effect to be achieved.

For example, the second therapeutic agent can be a cytotoxic or a cytostatic agent. Exemplary cytotoxic agents include antimicrotubule agents, topoisomerase inhibitors, or taxanes, antimetabolites, mitotic inhibitors, alkylating agents, intercalating agents, agents capable of interfering with a signal transduction pathway, agents that promote apoptosis and radiation. In yet other embodiments, the methods can be used in combination with immunodulatory agents, *e.g.*, IL-1, 2, 4, 6, or 12, or interferon alpha or gamma, or immune cell growth factors such as GM-CSF.

Anti-cancer agents, *e.g.*, kinase inhibitors, used in therapeutic methods can be evaluated using the screening assays described herein. In one embodiment, the

IDT Exhibit 1020.3011

anti-cancer agents are evaluated in a cell-free system, *e.g.,* a cell lysate or in a reconstituted system. In other embodiments, the anti-cancer agents are evaluated in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In yet other embodiments, the anti-cancer agents are evaluated cell *in vivo* (a DCTN1-ALK -expressing cell present in a subject, *e.g.*, an animal subject (*e.g.*, an *in vivo* animal model).

Exemplary parameters evaluated include one or more of:

(i) a change in binding activity, *e.g.*, direct binding of the candidate agent to a DCTN1-ALK fusion polypeptide; a binding competition between a known ligand and the candidate agent to a DCTN1-ALK fusion polypeptide;

(ii) a change in kinase activity, *e.g.*, phosphorylation levels of a DCTN1-ALK fusion polypeptide (*e.g.,* an increased or decreased autophosphorylation); or a change in phosphorylation of a target of an ALK kinase, *e.g.*, Akt/Protein kinase B;

(iii) a change in an activity of a cell containing a DCTN1-ALK fusion (*e.g.*, a tumor cell or a recombinant cell), *e.g.*, a change in proliferation, morphology or tumorigenicity of the cell;

(iv) a change in tumor present in an animal subject, *e.g.*, size, appearance, proliferation, of the tumor; or

(v) a change in the level, *e.g.*, expression level, of a DCTN1-ALK fusion polypeptide or nucleic acid molecule.

In one embodiment, a change in a cell free assay in the presence of a candidate agent is evaluated. For example, an activity of a DCTN1-ALK fusion, or interaction of a DCTN1-ALK fusion with a downstream ligand can be detected.

In other embodiments, a change in an activity of a cell is detected in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In one embodiment, the cell is a recombinant cell that is modified to express a DCTN1-ALK fusion nucleic acid, *e.g.*, is a recombinant cell transfected with a DCTN1-ALK fusion nucleic acid. The transfected cell can show a change in response to the expressed DCTN1-ALK fusion, *e.g.*, increased proliferation, changes in morphology, increased tumorigenicity, and/or acquired a transformed phenotype. A change in any of the activities of the cell, *e.g.*, the recombinant cell, in the presence of the candidate agent can be detected. For example, a decrease in one or more

IDT Exhibit 1020.3012

of: proliferation, tumorigenicity, transformed morphology, in the presence of the candidate agent can be indicative of an inhibitor of a DCTN1-ALK fusion. In other embodiments, a change in binding activity or phosphorylation as described herein is detected.

In yet other embodiment, a change in a tumor present in an animal subject (*e.g.*, an *in vivo* animal model) is detected. In one embodiment, the animal model is a tumor containing animal or a xenograft comprising cells expressing a DCTN1-ALK fusion (*e.g.*, tumorigenic cells expressing a DCTN1-ALK fusion). The anti-cancer agents can be administered to the animal subject and a change in the tumor is detected. In one embodiment, the change in the tumor includes one or more of a tumor growth, tumor size, tumor burden, survival, is evaluated. A decrease in one or more of tumor growth, tumor size, tumor burden, or an increased survival is indicative that the candidate agent is an inhibitor.

The screening methods and assays are described in more detail herein below.

Screening Methods

In another aspect, the invention features a method, or assay, for screening for agents that modulate, *e.g.*, inhibit, the expression or activity of a DCTN1-ALK fusion, *e.g.*, a DCTN1-ALK fusion as described herein. The method includes contacting a DCTN1-ALK fusion, or a cell expressing a DCTN1-ALK fusion, with a candidate agent; and detecting a change in a parameter associated with a DCTN1-ALK fusion, *e.g.*, a change in the expression or an activity of the DCTN1-ALK fusion. The method can, optionally, include comparing the treated parameter to a reference value, *e.g.*, a control sample (*e.g.*, comparing a parameter obtained from a sample with the candidate agent to a parameter obtained from a sample without the candidate agent). In one embodiment, if a decrease in expression or activity of the DCTN1-ALK fusion is detected, the candidate agent is identified as an inhibitor. In another embodiment, if an increase in expression or activity of the DCTN1-ALK fusion is detected, the candidate agent is identified as an activator. In certain embodiments, the DCTN1-ALK fusion is a nucleic acid molecule or a polypeptide as described herein.

In one embodiment, the contacting step is effected in a cell-free system, *e.g.*, a cell lysate or in a reconstituted system. In other embodiments, the contacting step is

IDT Exhibit 1020.3013

effected in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In yet other embodiments, the contacting step is effected in a cell *in vivo* (a DCTN1-ALK -expressing cell present in a subject, *e.g.*, an animal subject (*e.g.*, an *in vivo* animal model).

Exemplary parameters evaluated include one or more of:

(i) a change in binding activity, *e.g.*, direct binding of the candidate agent to a DCTN1-ALK fusion polypeptide; a binding competition between a known ligand and the candidate agent to a DCTN1-ALK fusion polypeptide;

(ii) a change in kinase activity, *e.g.*, phosphorylation levels of a DCTN1-ALK fusion polypeptide (*e.g.,* an increased or decreased autophosphorylation); or a change in phosphorylation of a target of an ALK kinase, *e.g.*, Protein kinase B/Akt. In certain embodiments, a change in kinase activity, *e.g.*, phosphorylation, is detected by any of Western blot (*e.g.*, using an anti-DCTN1 or anti-ALK antibody; a phosphor-specific antibody, detecting a shift in the molecular weight of a DCTN1-ALK fusion polypeptide), mass spectrometry, immunoprecipitation, immunohistochemistry, immunomagnetic beads, among others;

(iii) a change in an activity of a cell containing a DCTN1-ALK fusion (*e.g.*, a tumor cell or a recombinant cell), *e.g.*, a change in proliferation, morphology or tumorigenicity of the cell;

(iv) a change in tumor present in an animal subject, *e.g.*, size, appearance, proliferation, of the tumor; or

(v) a change in the level, *e.g.*, expression level, of a DCTN1-ALK fusion polypeptide or nucleic acid molecule.

In one embodiment, a change in a cell free assay in the presence of a candidate agent is evaluated. For example, an activity of a DCTN1-ALK fusion, or interaction of a DCTN1-ALK fusion with a downstream ligand can be detected. In one embodiment, a DCTN1-ALK fusion polypeptide is contacted with a ligand, *e.g.*, in solution, and a candidate agent is monitored for an ability to modulate, *e.g.*, inhibit, an interaction, *e.g.*, binding, between the DCTN1-ALK fusion polypeptide and the ligand. In one exemplary assay, purified DCTN1-ALK fusion protein is contacted with a ligand, *e.g.,* in solution, and a candidate agent is monitored for an ability to inhibit interaction of the fusion protein with the ligand, or to inhibit phosphorylation of the ligand by the fusion protein.

IDT Exhibit 1020.3014

An effect on an interaction between the fusion protein and a ligand can be monitored by methods known in the art, such as by absorbance, and an effect on phosphorylation of the ligand can be assayed, *e.g.*, by Western blot, immunoprecipitation, or immunomagnetic beads.

In other embodiments, a change in an activity of a cell is detected in a cell in culture, *e.g.*, a cell expressing a DCTN1-ALK fusion (*e.g.*, a mammalian cell, a tumor cell or cell line, a recombinant cell). In one embodiment, the cell is a recombinant cell that is modified to express a DCTN1-ALK fusion nucleic acid, *e.g.*, is a recombinant cell transfected with a DCTN1-ALK fusion nucleic acid. The transfected cell can show a change in response to the expressed K DCTN1-ALK fusion, *e.g.*, increased proliferation, changes in morphology, increased tumorigenicity, and/or acquired a transformed phenotype. A change in any of the activities of the cell, *e.g.*, the recombinant cell, in the presence of the candidate agent can be detected. For example, a decrease in one or more of: proliferation, tumorigenicity, transformed morphology, in the presence of the candidate agent can be indicative of an inhibitor of a DCTN1-ALK fusion. In other embodiments, a change in binding activity or phosphorylation as described herein is detected.

In an exemplary cell-based assay, a nucleic acid comprising a DCTN1-ALK fusion can be expressed in a cell, such as a cell (*e.g.*, a mammalian cell) in culture. The cell containing a nucleic acid expressing the DCTN1-ALK fusion can be contacted with a candidate agent, and the cell is monitored for an effect of the candidate agent. A candidate agent that causes decreased cell proliferation or cell death can be determined to be a candidate for treating a tumor (*e.g.*, a cancer) that carries a DCTN1-ALK fusion.

In one embodiment, a cell containing a nucleic acid expressing a DCTN1-ALK fusion can be monitored for expression of the DCTN1-ALK fusion protein. Protein expression can be monitored by methods known in the art, such as by, *e.g.*, mass spectrometry (*e.g.*, tandem mass spectrometry), a reporter assay (*e.g.*, a fluorescence-based assay), Western blot, and immunohistochemistry. By one method, decreased DCTN1-ALK expression is detected. A candidate agent that causes decreased expression of the DCTN1-ALK fusion protein as compared to a cell that does not contain the DCTN1-ALK nucleic acid fusion can be determined to be a candidate for treating a tumor (*e.g.*, a cancer) that carries a DCTN1-ALK fusion.

IDT Exhibit 1020.3015

A cell containing a nucleic acid expressing a DCTN1-ALK fusion can be monitored for altered ALK kinase activity. Kinase activity can be assayed by measuring the effect of a candidate agent on a known ALK kinase target protein, such as *e.g.*, protein kinase B/Akt.

In yet other embodiment, a change in a tumor present in an animal subject (*e.g.*, an *in vivo* animal model) is detected. In one embodiment, the animal model is a tumor containing animal or a xenograft comprising cells expressing a DCTN1-ALK fusion (*e.g.*, tumorigenic cells expressing a DCTN1-ALK fusion). The candidate agent can be administered to the animal subject and a change in the tumor is detected. In one embodiment, the change in the tumor includes one or more of a tumor growth, tumor size, tumor burden, survival, is evaluated. A decrease in one or more of tumor growth, tumor size, tumor burden, or an increased survival is indicative that the candidate agent is an inhibitor.

In one exemplary animal model, a xenograft is created by injecting cells into mouse. A candidate agent is administered to the mouse, *e.g.,* by injection (such as subcutaneous, intraperitoneal, or tail vein injection, or by injection directly into the tumor) or oral delivery, and the tumor is observed to determine an effect of the candidate anti-cancer agent. The health of the animal is also monitored, such as to determine if an animal treated with a candidate agent survives longer. A candidate agent that causes growth of the tumor to slow or stop, or causes the tumor to shrink in size, or causes decreased tumor burden, or increases survival time, can be considered to be a candidate for treating a tumor (*e.g.*, a cancer) that carries a DCTN1-ALK fusion.

In another exemplary animal assay, cells expressing a DCTN1-ALK fusion are injected into the tail vein, *e.g.*, of a mouse, to induce metastasis. A candidate agent is administered to the mouse, *e.g.,* by injection (such as subcutaneous, intraperitoneal, or tail vein injection, or by injection directly into the tumor) or oral delivery, and the tumor is observed to determine an effect of the candidate anti-cancer agent. A candidate agent that inhibits or prevents or reduces metastasis, or increases survival time, can be considered to be a candidate for treating a tumor (*e.g.*, a cancer) that carries a DCTN1-ALK fusion.

Cell proliferation can be measured by methods known in the art, such as PCNA (Proliferating cell nuclear antigen) assay, 5-bromodeoxyuridine (BrdUrd) incorporation,

IDT Exhibit 1020.3016

Ki-67 assay, mitochondrial respiration, or propidium iodide staining. Cells can also be measured for apoptosis, such as by use of a TUNEL (Terminal Deoxynucleotide Transferase dUTP Nick End Labeling) assay. Cells can also be assayed for presence of angiogenesis using methods known in the art, such as by measuring endothelial tube formation or by measuring the growth of blood vessels from subcutaneous tissue, such as into a solid gel of basement membrane.

In other embodiments, a change in expression of a DCTN1-ALK fusion can be monitored by detecting the nucleic acid or protein levels, *e.g.*, using the methods described herein.

In certain embodiments, the screening methods described herein can be repeated and/or combined. In one embodiment, a candidate agent that is evaluated in a cell-free or cell-based described herein can be further tested in an animal subject.

In one embodiment, the candidate agent is identified and re-tested in the same or a different assay. For example, a test compound is identified in an *in vitro* or cell-free system, and re-tested in an animal model or a cell-based assay. Any order or combination of assays can be used. For example, a high throughput assay can be used in combination with an animal model or tissue culture.

Candidate agents suitable for use in the screening assays described herein include, *e.g.*, small molecule compounds, nucleic acids (*e.g.*, siRNA, aptamers, short hairpin RNAs, antisense oligonucleotides, ribozymes, antagomirs, microRNA mimics or DNA, *e.g.*, for gene therapy) or polypeptides, *e.g.*, antibodies (*e.g.*, full length antibodies or antigen-binding fragments thereof, Fab fragments, or scFv fragments). The candidate anti-cancer agents can be obtained from a library (*e.g.*, a commercial library), or can be rationally designed, such as to target an active site in a functional domain of ALK (*e.g.*, the kinase domain of ALK), or a functional domain of DCTN1 (*e.g.*, the dynein or kinesin II binding domain).

In other embodiments, the method, or assay, includes providing a step based on proximity-dependent signal generation, *e.g.*, a two-hybrid assay that includes a first fusion protein (e.g., a DCTN1-ALK fusion protein), and a second fusion protein (*e.g.*, a ligand), contacting the two-hybrid assay with a test compound, under conditions wherein said two hybrid assay detects a change in the formation and/or stability of the complex, *e.g.,* the formation of the complex initiates transcription activation of a reporter gene.

IDT Exhibit 1020.3017

In one non-limiting example, the three-dimensional structure of the active site of DCTN1-ALK fusion is determined by crystallizing the complex formed by the DCTN1-ALK fusion and a known inhibitor. Rational drug design is then used to identify new test agents by making alterations in the structure of a known inhibitor or by designing small molecule compounds that bind to the active site of the DCTN1-ALK fusion.

The candidate agents can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; *see, e.g.*, Zuckermann, R.N. *et al.* (1994) *J. Med. Chem.* 37:2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.,* Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner U.S. Patent No. 5,223,409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici (1991) *J. Mol. Biol.* 222:301-310; Ladner *supra.*).

The interaction between two molecules can also be detected, *e.g.,* using fluorescence energy transfer (FET) (see, for example, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos, *et al.*, U.S. Patent No. 4,868,103). A fluorophore label on the first, 'donor'

IDT Exhibit 1020.3018

molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means known in the art (*e.g.*, using a fluorimeter).

In another embodiment, determining the ability of the DCTN1-ALK fusion protein to bind to a target molecule can be accomplished using real-time Biomolecular Interaction Analysis (BIA) (*see, e.g.*, Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705). "Surface plasmon resonance" or "BIA" detects biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

Nucleic Acid Inhibitors

In yet another embodiment, the DCTN1-ALK fusion inhibitor inhibits the expression of nucleic acid encoding the fusion. Examples of such fusion inhibitors include nucleic acid molecules, for example, antisense molecules, ribozymes, siRNA, triple helix molecules that hybridize to a nucleic acid encoding a DCTN1-ALK fusion, or a transcription regulatory region, and blocks or reduces mRNA expression of the fusion.

In one embodiment, the nucleic acid antagonist is a siRNA that targets mRNA encoding a DCTN1-ALK fusion. Other types of antagonistic nucleic acids can also be used, *e.g.*, a dsRNA, a ribozyme, a triple-helix former, or an antisense nucleic acid. Accordingly,

IDT Exhibit 1020.3019

isolated nucleic acid molecules that are nucleic acid inhibitors, *e.g.*, antisense, RNAi, to a DCTN1-ALK fusion-encoding nucleic acid molecule are provided.

An "antisense" nucleic acid can include a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire fusion coding strand, or to only a portion thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding fusion (e.g., the 5' and 3' untranslated regions). Anti-sense agents can include, for example, from about 8 to about 80 nucleobases (*i.e.*, from about 8 to about 80 nucleotides), *e.g.*, about 8 to about 50 nucleobases, or about 12 to about 30 nucleobases. Anti-sense compounds include ribozymes, external guide sequence (EGS) oligonucleotides (oligozymes), and other short catalytic RNAs or catalytic oligonucleotides which hybridize to the target nucleic acid and modulate its expression. Anti-sense compounds can include a stretch of at least eight consecutive nucleobases that are complementary to a sequence in the target gene. An oligonucleotide need not be 100% complementary to its target nucleic acid sequence to be specifically hybridizable. An oligonucleotide is specifically hybridizable when binding of the oligonucleotide to the target interferes with the normal function of the target molecule to cause a loss of utility, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide to non-target sequences under conditions in which specific binding is desired, *i.e.*, under physiological conditions in the case of in vivo assays or therapeutic treatment or, in the case of in vitro assays, under conditions in which the assays are conducted.

Hybridization of antisense oligonucleotides with mRNA can interfere with one or more of the normal functions of mRNA. The functions of mRNA to be interfered with include all key functions such as, for example, translocation of the RNA to the site of protein translation, translation of protein from the RNA, splicing of the RNA to yield one or more mRNA species, and catalytic activity which may be engaged in by the RNA. Binding of specific protein(s) to the RNA may also be interfered with by antisense oligonucleotide hybridization to the RNA.

Exemplary antisense compounds include DNA or RNA sequences that specifically hybridize to the target nucleic acid, *e.g.*, the mRNA encoding DCTN1-ALK fusion. The

IDT Exhibit 1020.3020

complementary region can extend for between about 8 to about 80 nucleobases. The compounds can include one or more modified nucleobases. Modified nucleobases are known in the art. Descriptions of modified nucleic acid agents are also available. *See, e.g.*, U.S. Patent Nos. 4,987,071; 5,116,742; and 5,093,246; Woolf et al. (1992) *Proc Natl Acad Sci USA*; *Antisense RNA and DNA*, D.A. Melton, Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1988); 89:7305-9; Haselhoff and Gerlach (1988) *Nature* 334:585-59; Helene, C. (1991) *Anticancer Drug Des.* 6:569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14:807-15.

The antisense nucleic acid molecules are typically administered to a subject (*e.g.*, by direct injection at a tissue site), or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a DCTN1-NTRK1 fusion to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then be administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

siRNAs are small double stranded RNAs (dsRNAs) that optionally include overhangs. For example, the duplex region of an siRNA is about 18 to 25 nucleotides in length, *e.g.*, about 19, 20, 21, 22, 23, or 24 nucleotides in length. Typically, the siRNA sequences are exactly complementary to the target mRNA. dsRNAs and siRNAs in particular can be used to silence gene expression in mammalian cells (e.g., human cells). siRNAs also

IDT Exhibit 1020.3021

include short hairpin RNAs (shRNAs) with 29-base-pair stems and 2-nucleotide 3' overhangs. See, e.g., Clemens *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97:6499-6503; Billy *et al.* (2001) *Proc. Natl. Sci. USA* 98:14428-14433; Elbashir *et al.* (2001) *Nature*. 411:494-8; Yang *et al.* (2002) *Proc. Natl. Acad. Sci. USA* 99:9942-9947; Siolas *et al.* (2005), *Nat. Biotechnol.* 23(2):227-31; 20040086884; U.S. 20030166282; 20030143204; 20040038278; and 20030224432.

In still another embodiment, an antisense nucleic acid featured in the invention is a ribozyme. A ribozyme having specificity for a DCTN1-ALK fusion-encoding nucleic acid can include one or more sequences complementary to the nucleotide sequence of a fusion cDNA disclosed herein (*i.e.,* SEQ ID NO:6), and a sequence having known catalytic sequence responsible for mRNA cleavage (*see* U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach (1988) *Nature* 334:585-591). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a DCTN1-ALK fusion -encoding mRNA. See, e.g., Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, fusion mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Inhibition of a DCTN1-ALK fusion gene can be accomplished by targeting nucleotide sequences complementary to the regulatory region of the fusion to form triple helical structures that prevent transcription of the DCTN1-ALK fusion gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6:569-84; Helene, C. i (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14:807-15. The potential sequences that can be targeted for triple helix formation can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

The invention also provides detectably labeled oligonucleotide primer and probe molecules. Typically, such labels are chemiluminescent, fluorescent, radioactive, or colorimetric.

IDT Exhibit 1020.3022

A fusion nucleic acid molecule can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For non-limiting examples of synthetic oligonucleotides with modifications see Toulmé (2001) *Nature Biotech.* 19:17 and Faria *et al.* (2001) *Nature Biotech.* 19:40-44. Such phosphoramidite oligonucleotides can be effective antisense agents.

For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4: 5-23). As used herein, the terms "peptide nucleic acid" or "PNA" refers to a nucleic acid mimic, *e.g.,* a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra* and Perry-O'Keefe *et al. Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of DCTN1-ALK fusion nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of fusion nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.,* by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.,* S1 nucleases (Hyrup B. *et al.* (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.,* for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; W088/09810) or the blood-brain barrier (see, e.g., W0 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, *e.g.,* Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents (*See, e.g.,* Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.,* a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

IDT Exhibit 1020.3023

In some embodiments, a nucleic acid inhibitor described herein is provided for the inhibition of expression of an LMNA1-NTRK1 nucleic acid *in vitro*.

Evaluation of Subjects

Subjects, *e.g.*, patients, can be evaluated for the presence of a DCTN1-ALK fusion or a LMNA-NTRK1 fusion. A patient can be evaluated, for example, by determining the genomic sequence of the patient, *e.g.*, by an NGS method. Alternatively, or in addition, evaluation of a patient can include directly assaying for the presence of a DCTN1-ALK fusion or an LMNA-NTRK1 fusion in the patient, such as by an assay to detect a fusion nucleic acid (*e.g.*, DNA or RNA), such as by, Southern blot, Northern blot, or RT-PCR, e.g., qRT-PCR. Alternatively, or in addition, a patient can be evaluated for the presence of a protein fusion, such as by immunohistochemistry, Western blot, immunoprecipitation, or immunomagnetic bead assay.

Evaluation of a patient can also include a cytogenetic assay, such as by fluorescence in situ hybridization (FISH), to identify the chromosomal rearrangement resulting in the DCTN1-ALK or LMNA-NTRK1 fusion. For example, to perform FISH, at least a first probe tagged with a first detectable label can be designed to target DCTN1, such as in one or more exons of DCTN1 and at least a second probe tagged with a second detectable label can be designed to target ALK, such as in one or more exons of ALK (*e.g.*, the exons containing the part of the protein that includes the tyrosine kinase domain). The at least one first probe and the at least one second probe will be closer together in patients who carry the DCTN1-ALK fusion than in patients who do not carry the DCTN1-ALK fusion. These methods can be utilized in a similar manner for the LMNA-NTRK1 fusion.

Additional methods for fusion detection are provided below.

In one aspect, the results of a clinical trial, *e.g.*, a successful or unsuccessful clinical trial, can be repurposed to identify agents that target a DCTN1-ALK fusion. By one exemplary method, a candidate agent used in a clinical trial can be reevaluated to determine if the agent in the trial targets a fusion, or is effective to treat a tumor containing a particular fusion. For example, subjects who participated in a clinical trial for an agent, such as a kinase inhibitor, can be identified. Patients who experienced an improvement in symptoms, *e.g.*, cancer (*e.g.*, lung cancer) symptoms, such as decreased

IDT Exhibit 1020.3024

tumor size, or decreased rate of tumor growth, can be evaluated for the presence of a DCTN1-ALK fusion. Patients who did not experience an improvement in cancer symptoms can also be evaluated for the presence of a DCTN1-ALK fusion. Where patients carrying a DCTN1-ALK fusion are found to have been more likely to respond to the test agent than patients who did not carry such a fusion, then the agent is determined to be an appropriate treatment option for a patient carrying the fusion.

"Reevaluation" of patients can include, for example, determining the genomic sequence of the patients, or a subset of the clinical trial patients, *e.g.,* by an NGS method. Alternatively, or in addition, reevaluation of the patients can include directly assaying for the presence of a DCTN1-ALK fusion in the patient, such as by an assay to detect a DCTN1-ALK nucleic acid (*e.g.,* RNA), such as by RT-PCR, e.g., qRT-PCR. Alternatively, or in addition, a patient can be evaluated for the presence of a DCTN1-ALK protein fusion, such as by immunohistochemistry, Western blot, immunoprecipitation, or immunomagnetic bead assay.

Clinical trials suitable for repurposing as described above include trials that tested ALK inhibitors, tyrosine kinase inhibitors, multikinase inhibitors, and drugs purported to act upstream or downstream of ALK in a pathway involving ALK. Other clinical trials suitable for repurposing as described above include trials that tested DCTN1 inhibitors, kinesin inhibitors, inhibitors of cell trafficking and drugs purported to act upstream or downstream of DCTN1 in a pathway involving DCTN1.

Methods for Detection of Fusion Nucleic Acids and Polypeptides

Methods for evaluating a fusion gene, mutations and/or gene products are known to those of skill in the art. In one embodiment, the fusion is detected in a nucleic acid molecule by a method chosen from one or more of: nucleic acid hybridization assay, amplification-based assays (*e.g.*, polymerase chain reaction (PCR)), PCR-RFLP assay, real-time PCR, sequencing, screening analysis (including metaphase cytogenetic analysis by standard karyotype methods, FISH (*e.g.*, break away FISH), spectral karyotyping or MFISH, comparative genomic hybridization), *in situ* hybridization, SSP, HPLC or mass-spectrometric genotyping.

Additional exemplary methods include, traditional "direct probe" methods such as Southern blots or *in situ* hybridization (*e.g.,* fluorescence in situ hybridization

IDT Exhibit 1020.3025

(FISH) and FISH plus SKY), and "comparative probe" methods such as comparative genomic hybridization (CGH), *e.g.,* cDNA-based or oligonucleotide-based CGH, can be used. The methods can be used in a wide variety of formats including, but not limited to, substrate (e.g., membrane or glass) bound methods or array-based approaches.

In certain embodiments, the evaluation methods include the probes/primers described herein.

In one embodiment, probes/primers can be designed to detect a DCTN1-ALK fusion or a reciprocal thereof. The DCTN1 probes/primers can be from nucleotides 47-78 of SEQ ID NO:6 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the DCTN1 protein). These probes/primers are suitable, *e.g.,* for FISH or PCR amplification. The ALK probes/primers can be from nucleotides 3196-4907 of SEQ ID NO:6 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the ALK protein). These probes/primers are suitable, *e.g.*, for FISH or PCR amplification.

For PCR, *e.g.*, to amply a region including a DCTN1-ALK fusion junction, forward primers can be designed to hybridize to a DCTN1 sequence from nucleotides corresponding to DCTN1 in SEQ ID NO:6, and reverse primers can be designed to hybridize to an ALK sequence from nucleotides corresponding to ALK in SEQ ID NO:6.

For example, probes/primers can be designed to detect a DCTN1-ALK fusion or a reciprocal thereof. The ALK probes/primers can be from nucleotides 3196-4907 of SEQ ID NO:6 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the ALK protein). The DCNT1 probes/primers can be from nucleotides 1-3195 of SEQ ID NO:6 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the DCTN1 protein). These probes/primers are suitable, *e.g.*, for FISH or PCR amplification.

For PCR, *e.g.*, to amply a region including an LMNA-NTRK1 fusion junction (or reciprocal thereof), forward primers can be designed to hybridize to an LMNA sequence from nucleotides corresponding to LMNA in SEQ ID NO:9, and reverse primers can be designed to hybridize to an NTRK1 sequence from nucleotides corresponding to NTRK1 in SEQ ID NO:9.

IDT Exhibit 1020.3026

For example, probes/primers can be designed to detect an LMNA-NTRK1 fusion junction (or a reciprocal thereof). The NTRK1 probes/primers can be from nucleotides 514-1740 of SEQ ID NO:9 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the NTRK1 protein). The LMNA probes/primers can be from nucleotides 1-513 of SEQ ID NO:9 (*e.g.*, can hybridize to the nucleotides encoding one or more exons of the LMNA protein). These probes/primers are suitable, *e.g.*, for FISH or PCR amplification.

In one embodiment, FISH analysis is used to identify the chromosomal rearrangement resulting in the fusions as described above. For example, to perform FISH, at least a first probe tagged with a first detectable label can be designed to target DCTN1, such as in one or more exons of DCTN1 and at least a second probe tagged with a second detectable label can be designed to target ALK, such as in one or more exons of ALK (*e.g.*, the exons containing the part of the protein that includes the tyrosine kinase domain). The at least one first probe and the at least one second probe will be closer together in a subject who carries the ALK-DCTN1 fusion compared to a subject who does not carry the fusion.

In one approach, a variation of a FISH assay, *e.g.*, "break-away FISH", is used to evaluate a patient. By this method, at least one probe targeting the ALK intron 19/ DCTN1 intron 26 junction and at least one probe targeting DCNT1 (or ALK), *e.g.*, at one or more exons and or introns of DCTN1 or ALK, are utilized. In normal cells, both probes will be observed (or a secondary color will be observed due to the close proximity of the DCTN1 and ALK genes), and only the DCTN1 probe will be observed when the translocation occurs. Other variations of the FISH method known in the art are suitable for evaluating a patient.

The FISH methods described herein above use DCTN1-ALK as an example, and such methods can be readily applied to the LMNA-NTRK1 fusion by one of skill in the art.

Probes are used that contain DNA segments that are essentially complementary to DNA base sequences existing in different portions of chromosomes. Examples of probes useful according to the invention, and labeling and hybridization of probes to samples are described in two U.S. patents to Vysis, Inc. U.S. Patent Nos. 5,491,224 and 6,277,569 to Bittner, et al.

IDT Exhibit 1020.3027

Additional protocols for FISH detection are described below.

Chromosomal probes are typically about 50 to about 10^5 nucleotides in length. Longer probes typically comprise smaller fragments of about 100 to about 500 nucleotides in length. Probes that hybridize with centromeric DNA and locus-specific DNA are available commercially, for example, from Vysis, Inc. (Downers Grove, Ill.), Molecular Probes, Inc. (Eugene, Oreg.) or from Cytocell (Oxfordshire, UK). Alternatively, probes can be made non-commercially from chromosomal or genomic DNA through standard techniques. For example, sources of DNA that can be used include genomic DNA, cloned DNA sequences, somatic cell hybrids that contain one, or a part of one, chromosome (*e.g.*, human chromosome) along with the normal chromosome complement of the host, and chromosomes purified by flow cytometry or microdissection. The region of interest can be isolated through cloning, or by site-specific amplification via the polymerase chain reaction (PCR). See, for example, Nath and Johnson, Biotechnic Histochem., 1998, 73(1):6-22, Wheeless et al., Cytometry 1994, 17:319-326, and U.S. Patent No. 5,491,224.

The probes to be used hybridize to a specific region of a chromosome to determine whether a cytogenetic abnormality is present in this region. One type of cytogenetic abnormality is a deletion. Although deletions can be of one or more entire chromosomes, deletions normally involve loss of part of one or more chromosomes. If the entire region of a chromosome that is contained in a probe is deleted from a cell, hybridization of that probe to the DNA from the cell will normally not occur and no signal will be present on that chromosome. If the region of a chromosome that is partially contained within a probe is deleted from a cell, hybridization of that probe to the DNA from the cell can still occur, but less of a signal can be present. For example, the loss of a signal is compared to probe hybridization to DNA from control cells that do not contain the genetic abnormalities which the probes are intended to detect. In some embodiments, at least 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, or more cells are enumerated for presence of the cytogenetic abnormality.

Cytogenetic abnormalities to be detected can include, but are not limited to, non-reciprocal translocations, balanced translocations, intra-chromosomal inversions, point mutations, deletions, gene copy number changes, gene expression level changes,

IDT Exhibit 1020.3028

and germ line mutations. In particular, one type of cytogenetic abnormality is a duplication. Duplications can be of entire chromosomes, or of regions smaller than an entire chromosome. If the region of a chromosome that is contained in a probe is duplicated in a cell, hybridization of that probe to the DNA from the cell will normally produce at least one additional signal as compared to the number of signals present in control cells with no abnormality of the chromosomal region contained in the probe.

Chromosomal probes are labeled so that the chromosomal region to which they hybridize can be detected. Probes typically are directly labeled with a fluorophore, an organic molecule that fluoresces after absorbing light of lower wavelength/higher energy. The fluorophore allows the probe to be visualized without a secondary detection molecule. After covalently attaching a fluorophore to a nucleotide, the nucleotide can be directly incorporated into the probe with standard techniques such as nick translation, random priming, and PCR labeling. Alternatively, deoxycytidine nucleotides within the probe can be transaminated with a linker. The fluorophore then is covalently attached to the transaminated deoxycytidine nucleotides. See, U.S. Patent No. 5,491,224.

U.S. Patent No. 5,491,224 describes probe labeling as a number of the cytosine residues having a fluorescent label covalently bonded thereto. The number of fluorescently labeled cytosine bases is sufficient to generate a detectable fluorescent signal while the individual so labeled DNA segments essentially retain their specific complementary binding (hybridizing) properties with respect to the chromosome or chromosome region to be detected. Such probes are made by taking the unlabeled DNA probe segment, transaminating with a linking group a number of deoxycytidine nucleotides in the segment, covalently bonding a fluorescent label to at least a portion of the transaminated deoxycytidine bases.

Probes can also be labeled by nick translation, random primer labeling or PCR labeling. Labeling is done using either fluorescent (direct)-or haptene (indirect)-labeled nucleotides. Representative, non-limiting examples of labels include: AMCA-6-dUTP, CascadeBlue-4-dUTP, Fluorescein-12-dUTP, Rhodamine-6-dUTP, TexasRed-6-dUTP, Cy3-6-dUTP, Cy5-dUTP, Biotin(BIO)-11-dUTP, Digoxygenin(DIG)-11-dUTP or Dinitrophenyl (DNP)-11-dUTP.

IDT Exhibit 1020.3029

Probes also can be indirectly labeled with biotin or digoxygenin, or labeled with radioactive isotopes such as ^{32}p and .^{3}H, although secondary detection molecules or further processing then is required to visualize the probes. For example, a probe labeled with biotin can be detected by avidin conjugated to a detectable marker. For example, avidin can be conjugated to an enzymatic marker such as alkaline phosphatase or horseradish peroxidase. Enzymatic markers can be detected in standard colorimetric reactions using a substrate and/or a catalyst for the enzyme. Catalysts for alkaline phosphatase include 5-bromo-4-chloro-3-indolylphosphate and nitro blue tetrazolium. Diaminobenzoate can be used as a catalyst for horseradish peroxidase.

Probes can also be prepared such that a fluorescent or other label is not part of the DNA before or during the hybridization, and is added after hybridization to detect the probe hybridized to a chromosome. For example, probes can be used that have antigenic molecules incorporated into the DNA. After hybridization, these antigenic molecules are detected using specific antibodies reactive with the antigenic molecules. Such antibodies can themselves incorporate a fluorochrome, or can be detected using a second antibody with a bound fluorochrome.

However treated or modified, the probe DNA is commonly purified in order to remove unreacted, residual products (*e.g.,* fluorochrome molecules not incorporated into the DNA) before use in hybridization.

Prior to hybridization, chromosomal probes are denatured according to methods well known in the art. Probes can be hybridized or annealed to the chromosomal DNA under hybridizing conditions. "Hybridizing conditions" are conditions that facilitate annealing between a probe and target chromosomal DNA. Since annealing of different probes will vary depending on probe length, base concentration and the like, annealing is facilitated by varying probe concentration, hybridization temperature, salt concentration and other factors well known in the art.

Hybridization conditions are facilitated by varying the concentrations, base compositions, complexities, and lengths of the probes, as well as salt concentrations, temperatures, and length of incubation. For example, *in situ* hybridizations are typically performed in hybridization buffer containing 1-2x SSC, 50-65% formamide and blocking DNA to suppress non-specific hybridization. In general, hybridization

IDT Exhibit 1020.3030

conditions, as described above, include temperatures of about 25° C to about 55° C, and incubation lengths of about 0.5 hours to about 96 hours.

Non-specific binding of chromosomal probes to DNA outside of the target region can be removed by a series of washes. Temperature and concentration of salt in each wash are varied to control stringency of the washes. For example, for high stringency conditions, washes can be carried out at about 65° C to about 80° C, using 0.2x to about 2x SSC, and about 0.1% to about 1% of a non-ionic detergent such as Nonidet P-40 (NP40). Stringency can be lowered by decreasing the temperature of the washes or by increasing the concentration of salt in the washes. In some applications it is necessary to block the hybridization capacity of repetitive sequences. Thus, in some embodiments, tRNA, human genomic DNA, or Cot-I DNA is used to block non-specific hybridization. After washing, the slide is allowed to drain and air dry, then mounting medium, a counterstain such as DAPI, and a coverslip are applied to the slide. Slides can be viewed immediately or stored at -20° C. before examination.

For fluorescent probes used in fluorescence in situ hybridization (FISH) techniques, fluorescence can be viewed with a fluorescence microscope equipped with an appropriate filter for each fluorophore, or by using dual or triple band-pass filter sets to observe multiple fluorophores. See, for example, U.S. Patent No. 5,776,688. Alternatively, techniques such as flow cytometry can be used to examine the hybridization pattern of the chromosomal probes.

In CGH methods, a first collection of nucleic acids (*e.g.*, from a sample, *e.g.,* a possible tumor) is labeled with a first label, while a second collection of nucleic acids (*e.g.*, a control, *e.g.,* from a healthy cell/tissue) is labeled with a second label. The ratio of hybridization of the nucleic acids is determined by the ratio of the two (first and second) labels binding to each fiber in the array. Where there are chromosomal deletions or multiplications, differences in the ratio of the signals from the two labels will be detected and the ratio will provide a measure of the copy number. Array-based CGH can also be performed with single-color labeling (as opposed to labeling the control and the possible tumor sample with two different dyes and mixing them prior to hybridization, which will yield a ratio due to competitive hybridization of probes on the arrays). In single color CGH, the control is labeled and hybridized to

IDT Exhibit 1020.3031

one array and absolute signals are read, and the possible tumor sample is labeled and hybridized to a second array (with identical content) and absolute signals are read. Copy number difference is calculated based on absolute signals from the two arrays.

Hybridization protocols suitable for use with the methods featured in the invention are described, *e.g.,* in Albertson (1984) *EMBO J.* 3: 1227-1234; Pinkel (1988) *Proc. Natl. Acad. Sci. USA* 85: 9138-9142; EPO Pub. No. 430,402; *Methods in Molecular Biology*, Vol. 33: *In situ* Hybridization Protocols, Choo, ed., Humana Press, Totowa, N.J. (1994), *etc.* In one embodiment, the hybridization protocol of Pinkel, *et al.* (1998) *Nature Genetics* 20: 207-211, or of Kallioniemi (1992) *Proc. Natl Acad Sci USA* 89:5321-5325 (1992) is used. Array-based CGH is described in U.S. Patent No. 6,455,258, the contents of each of which are incorporated herein by reference.

In still another embodiment, amplification-based assays can be used to measure presence/absence and copy number. In such amplification-based assays, the nucleic acid sequences act as a template in an amplification reaction (*e.g.,* Polymerase Chain Reaction (PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls, *e.g.,* healthy tissue, provides a measure of the copy number.

Methods of "quantitative" amplification are well known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that can be used to calibrate the PCR reaction. Detailed protocols for quantitative PCR are provided in Innis, *et al.* (1990) PCR Protocols, *A Guide to Methods and Applications*, Academic Press, Inc. N.Y.). Measurement of DNA copy number at microsatellite loci using quantitative PCR analysis is described in Ginzonger, *et al.* (2000) *Cancer Research* 60:5405-5409. The known nucleic acid sequence for the genes is sufficient to enable one of skill in the art to routinely select primers to amplify any portion of the gene. Fluorogenic quantitative PCR can also be used. In fluorogenic quantitative PCR, quantitation is based on amount of fluorescence signals, *e.g.,* TaqMan and sybr green.

IDT Exhibit 1020.3032