9) Sequencing the resulting copies of the at least one intact cDNA strand.

[0095] Embodiment 10: Consequently, it is also possible to include a sequencing step if the first-stranded included the at least one modified dNTP and the intact second-strand cDNA is amplified, as in embodiment 2:

1) Providing at least one single-stranded RNA molecule;
2) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP;
3) Generating a first cDNA strand;
4) Replacing the at least one modified dNTP with the corresponding dNTPs or analogues of the corresponding dNTPs;
5) Performing second-strand synthesis on the first cDNA strand;
6) Ligating polarity-specific adaptors to one or both ends of the cDNA double strand;
7) Specifically modifying or removing the first cDNA strand containing the at least one modified dNTP; and
8) Optionally, amplifying the at least one remaining intact cDNA strand by performing PCR.
9) Sequencing the resulting copies of the at least one intact cDNA strand.

[0096] Embodiment 11: After sequencing it may be desirable to determine the polarity of the resulting sequence(s). This may serve for several analytic or diagnostic purposes (see below) or for aligning RNA transcripts to gene sequences known from DNA sequencing projects. If the remaining intact first strand cDNA was sequenced the order of the steps of the inventive method read as follows:

1) Providing at least one single-stranded RNA molecule;
2) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof;
3) Generating a first cDNA strand;
4) Replacing at least one dNTP with at least one corresponding modified dNTP;
5) Performing second-strand synthesis on the first cDNA strand;
6) Ligating polarity-specific adaptors to one or both ends of the cDNA double strand;
7) Specifically modifying or removing the second cDNA strand containing the at least one modified dNTP; and
8) Optionally, amplifying the at least one remaining intact cDNA strand by performing PCR.
9) Sequencing the resulting copies of the at least one intact cDNA strand.
10) Aligning of sequencing reads to the reference genome or reconstructing of the RNA sequence from individual sequencing reads.
11) Determining the polarity of the at least one RNA transcript.

[0097] Embodiment 12: Consequently, such a determination of the polarity and optionally an alignment to a gene structure known from a sequencing project can also be performed if the remaining intact second strand cDNA was amplified and sequenced. As in this case the antisense sequence is resulting the sequence has to be inverted, preferably by means of a computer programme, into the correct reading sequence.

[0098] Thus the order of the steps of the inventive method is:

1) Providing at least one single-stranded RNA molecule;
2) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP;
3) Generating a first cDNA strand;
4) Replacing the at least one modified dNTP with the corresponding dNTPs or analogues of the corresponding dNTPs;
5) Performing second-strand synthesis on the first cDNA strand;
6) Ligating polarity-specific adaptors to one or both ends of the cDNA double strand;
7) Specifically modifying or removing the first cDNA strand containing the at least one modified dNTP; and
8) Optionally, amplifying the at least one remaining intact cDNA strand by performing PCR;
9) Sequencing the resulting copies of the at least one intact cDNA strand;
10) Aligning of sequencing reads to the reference genome or reconstructing of the RNA sequence from individual sequencing reads;
11) Determining the polarity of the at least one RNA transcript

[0099] Embodiment 13: In analogy, it may be also desirable to sequence the amplified copies of a single-stranded DNA if the sequence isn't known yet. Thus a sequencing step ensues to the order of the steps of the inventive method of embodiment 3:

IDT Exhibit 1020.9423

1) Providing at least one single-stranded DNA molecule;
2) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP used for the strand specific differentiation;
3) Generating a cDNA strand;
4) Ligating polarity-specific adaptors to one or both ends of the resulting DNA double strand;
5) Specifically modifying or removing the cDNA strand containing the at least one modified dNTP; and
6) Optionally, amplifying the at least one intact DNA strand by performing PCR.
7) Sequencing the resulting copies of the at least one intact DNA strand.

[0100] Embodiment 14: Also for single-stranded DNA it may be desirable to determine the polarity of the resulting copies after amplification and sequencing. As for single-stranded DNA by needs only one cDNA strand is generated including the at least one modified dNTP the resulting sequence must be inverted, as in embodiment 12.

1) Providing at least one single-stranded DNA molecule;
2) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP used for the strand specific differentiation;
3) Generating a cDNA strand;
4) Ligating polarity-specific adaptors to one or both ends of the resulting DNA double strand;
5) Specifically modifying or removing the cDNA strand containing the at least one modified dNTP; and
6) Optionally, amplifying the at least one intact DNA strand by performing PCR;
7) Sequencing the resulting copies of the at least one intact DNA strand;
8) Aligning of sequencing reads to the reference genome or reconstructing of the DNA sequence from individual sequencing reads;
9) Determining the polarity of the at least one DNA molecule.

[0101] In another preferred embodiment 15 the original polynucleotide strand is a double-stranded DNA. Herein it may be necessary to set a nick into the particular strand of the double-stranded DNA which shall be substituted. A possible sequence of the steps according to this embodiment is:

1) Providing at least one double-stranded DNA;
2) Setting a nick in one particular strand of the double-stranded DNA by using a nicking endonuclease;
3) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP used for the strand specific differentiation;
4) Performing nick-translation of the nicked DNA strand by means of a DNA polymerase with 5' to 3' exonuclease activity;
5) Specifically modifying or removing the cDNA strand containing the at least one modified dNTP.

[0102] In a further preferred embodiment 16 the original polynucleotide strand is a double-stranded DNA obtained as a result of PCR reaction. A special modified nucleotide is introduced into one PCR primer to set a nick into the particular strand of the double-stranded DNA which shall be substituted. A possible sequence of the steps according to this embodiment is:

1) Providing at least one double-stranded DNA obtained in PCR reaction where one of PCR-primers contains dUTP;
2) Setting a nick in one particular strand of the double-stranded DNA by UNG treatment and APE 1 digestion or UNG treatment and alkalic hydrolisis;
3) Adding sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified dNTP used for the strand specific differentiation;
4) Performing nick-translation of the nicked DNA strand by means of a DNA polymerase with 5' to 3' exonuclease activity;
5) Specifically modifying or removing the cDNA strand containing the at least one modified dNTP.

[0103] It is obvious to a person skilled in the art that the different steps mentioned before in the respective embodiments can be combined for solving similar or different tasks provided the order and the selection of the different steps results in a meaningful order of the steps. All these possible combinations shall fall under the scope and the spirit of the invention.

[0104] By way of example, the method according to the invention was applied on the yeast transcriptome (Example 1). In a similar fashion, it was applied on the transcriptome of the mouse brain (Example 2). These two examples were chosen because comparative data from the Applicants' laboratory and other laboratories are available. Also complete genomes of these organisms are published. By way of comparing the results from conventional RNA-Seq and the

IDT Exhibit 1020.9424

inventive method according to the invention the advantages of the inventive method become apparent. As shown e.g. in Example 1 and Fig. 6, the inventive method allows a differentiation in sense and antisense transcripts. In this plot sense transcripts are marked in blue and antisense transcripts in red. Thus the method according to the invention can use the maintained information of the polarity of a transcript to generate more precise gene expression maps. Practical applications are discussed further below.

[0105] The reliability of the method according to the invention was validated in several control experiments. First, it was tested to which degree the inventive method can reproduce the results of conventional RNA sequencing. Herein the amplified sense and antisense transcripts according to the invention were pooled for allowing a correlation analysis. The results such as shown in Example 3 and Fig. 7A show a high correlation between the two methods. A high correlation corresponds to a distribution close to unity. A low correlation would be represented by a widely scattered distribution pattern. Thus the inventive method does not disturb the overall expression and transcriptome pattern in a sample. So the results of both methods can be compared to one another.

[0106] Next, two different strand-specific expression profilings were performed according to the invention on the same sample. Herein, only sense transcripts were included. The results were plotted again as described before. An excellent reproducibility of the inventive method is given, as can be seen in Example 4 and Fig. 7B.

[0107] In a further control experiment, the combination of sense and antisense transcripts obtained were compared to the results from another laboratory. There is still a very high correlation between the results, although with a somewhat broader scatter pattern (see Example 4 and Fig. 7C). Such a variation nearly always occurs when comparing the results from different laboratories. This experiment shows that this method does not work only under the specific conditions of the Applicants' laboratory, but can be universally used.

[0108] In order to test whether sense and antisense transcripts obtained according to the inventive method can be appropriately differentiated sense and antisense transcripts were compared and depicted in such an aforesaid scatter plot. As can be seen in Fig. 7D there is nearly no correlation. Thus sense and antisense transcripts obtained according to the inventive method are unequivocally different. This is a proof that the inventive method is highly suitable for differentiating accurately between sense and antisense transcripts.

[0109] The stability of the inventive method in respect of a complete exchange of at least one dNTP by at least one corresponding degradable analogue was tested. For this purpose in one reaction dTTP was exchanged by 100% dUTP. In another reaction dTTP was exchanged only by 75% dUTP, with 25% dTTP remaining in the reaction batch. The results from both experiments were compared for the sense transcripts and antisense transcripts (see Example 5 and Fig. 8). There is no significant deterioration in correlation when the exchange of the at least one dNTP was incomplete. When extrapolating these results it appears that only a part of the at least one dNTP must be replaced by the at least one corresponding degradable analogue to keep the method going with acceptable results, as shown in the correlation coefficient.

[0110] Actinomycin D is a substance which selectively inhibits DNA synthesis from DNA templates, but not from RNA templates. In a further experiment it could be shown that the presence of actinomycin D does not deteriorate significantly the results from the inventive protocol, neither for sense nor for antisense transcripts (see Example 6 and Fig. 9).

[0111] Practical applications of the inventive method include but are not limited to the provision of suitable kits for performing the inventive method according to the different protocols (embodiments) disclosed above.

[0112] Fig. 10 presents examples where information about transcription orientation helps to resolve overlaps of annotated and new genes in both yeast and mouse transcriptomes.

[0113] A kit in molecular biology is a package which includes all necessary ingredients for performing a certain method or singular step. Standard chemicals as present in any standard molecular biology laboratory are normally not included. Nevertheless some of these standard chemicals may be indispensable to carry out the inventive method properly. It is understood that all ingredients are provided in quantities that allow for a proper execution of the inventive method for the majority of scientific, diagnostic and industrial applications.

[0114] Often, but not always, these ingredients are provided in already prepared solutions ready- or close to ready-for-use. There may be also combinations of different ingredients already added together. A further advantage is that such kits use to be verified. Therefore the operator doesn't have to prove again the viability of the method and can save on at least some control experiments. Therefore kits are a very popular tool in molecular biology laboratories in research, diagnostics and industry.

[0115] For carrying out the key steps of the method according to the invention such a kit shall include the following components:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt

IDT Exhibit 1020.9425

for removing the strand containing the modified dNTPs.

[0116] If the protocol of the inventive method that shall be carried out includes an amplification step another embodiment of the kit should include:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt for removing the strand containing the modified dNTPs.
E) One or more polarity-specific adaptor.

[0117] If the protocol of the inventive method that shall be carried out includes an amplification step and the original polynucleotide is a RNA another embodiment of the kit should include:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt for removing the strand containing the modified dNTPs.
E) One or more polarity-specific adaptor;
F) One or more Reverse transcriptases.

[0118] In any of the aforegoing kits according to the invention it may be preferable to include additionally components for removal of unmodified nucleotides by gel-filtration or purification on silica-column. Thus the kits should include the following:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt for removing the strand containing the modified dNTPs;
E) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

or

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt for removing the strand containing the modified dNTPs.
E) One or more polarity-specific adaptor.
F) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

or

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific differentiation;
B) Primers
C) One or more DNA polymerases;
D) One or more degrading agents corresponding to the modified dNTPs and/or one or more removing agents apt for removing the strand containing the modified dNTPs.

IDT Exhibit 1020.9426

E) One or more polarity-specific adaptor;
F) One or more Reverse transcriptases.
G) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

[0119] Another application is the use of the results of this method for improved quantitative and qualitative diagnosis tests in selected diseases, as a part of the treatment or for prophylactic purposes. In general every disease is eligible which includes a temporary or permanent change in RNA expression levels. It is also possible to obtain a better monitoring of the effects of pharmaceutical agents used for therapy and/or prophylaxis of diseases provided these pharmaceutical agents produce a characteristic expression pattern different from that in untreated patients (humans or animals). Also improved gene expression patterns of plant, fungi, bacteria and viruses can be obtained by this method when the polarity of transcripts is available.

[0120] A further application is the production of improved and more selective microarrays. When the polarity of transcripts is known only the transcripts (respectively the antisense sequence) with a correct reading direction are meaningfully included in the microarray so that the limited space on the microarray is more economically used. Respectively, microarrays used for the generation of characteristic patterns instead of individual results may obtain a higher resolution by omitting non-characteristic transcripts of the incorrect reading direction.

[0121] As shown before, the advantages of the procedure according to the invention are providing a simple, inexpensive and reliable method for strand-specific cDNA cloning. This method is a novel and easy approach for determining the polarity of transcripts and using this information for purposes as outlined above.

[0122] Moreover, the method according to the invention has a higher accuracy in measuring expression levels of genes compared to RNA-Seq. Ignoring orientation of transcripts leads for example to overestimation of the expression level for about 10% of the mouse and about 15% of the yeast genes with noticeable expression level.

Figures

[0123]

Fig. 1: Directional double-stranded DNA
(A) original single-stranded polynucleotide; (B) directional double- stranded DNA with distinguishable adaptors (state of the art); (C) directional double-stranded DNA with distinguishable strands; (D) directional double-stranded DNA with distinguishable strands and distinguishable adapters.

Fig. 2A: Conversion of distinguishable strands into distinguishable adaptors
(A1) Y-shaped adaptors are ligated to double-stranded DNA with distinguishable strands; modified strands are separated from unmodified ones; after synthesis of complementary strands or PCR- amplification double-stranded DNA with distinguishable adapters is generated.
(A2) directional adaptors are ligated to double-stranded DNA with distinguishable strands; modified strands are separated from unmodified ones; after PCR-amplification double-stranded DNA with distinguishable adapters is generated.

Fig. 2B: Conversion of distinguishable adaptors into distinguishable strands
(B1) nick introduced into one of the adaptors using nicking endonuclease;
(B2) dUTP in one of the adaptors is converted into a nick: Uracil-N- Glycosylase converts dUTP into an abasic site and APE 1 nuclease converts the abasic site into a nick;
(B3) double-stranded DNA is destabilized in one of the Uracil-N- Glycosylase; new primer is annealed in the destabilized region;
(B4) position for annealing of the primer is organized by action 5' to 3' exonuclease; another adapter is protected from action of 5' to 3' exonuclease by thio-nucleotides; nick-translation is performed by DNA polymerase I (or Taq polymerase) in presence of modified nucleotides.

Fig. 3A: Preparation of single-stranded DNA according to the inventive method using dUTP as a modified nucleotide
* double-stranded DNA is a product of PCR reaction where one of the PCR-primers has dUTP at 3'-part;
* dUTP in the PCR primer is converted into a nick: Uracil-N-Glycosylase converts dUTP into an abasic site and APE 1 nuclease converts the abasic site into a nick;
* nick-translation is performed by DNA polymerase I (or Taq polymerase) in presence of dUTP;
* dUTP-containing strand is destroyed by Uracil-N-Glycosylase digestion and alkali hydrolysis;
* single-stranded DNA is purified by silica-based purification or gel- filtration or electrophoretic purification.

IDT Exhibit 1020.9427

Fig. 3B: Preparation of single-stranded DNA according to the inventive method using biotinilated nucleotides
* the double-stranded DNA is a product of PCR reaction where one of the PCR-primers has dUTP at 3'-part;
* dUTP in the PCR primer is converted into a nick: Uracil-N-Glycosylase converts dUTP into an abasic site and APE 1 nuclease converts the abasic site into a nick;
* nick-translation is performed by DNA polymerase I (or Taq polymerase) in presence of biotin-associated dTTPs and/or biotin- associated dUTPs;
* strands are separated using streptavidine-coated particles;

Fig. 4: Flowchart of the inventive method
RNA is shown in red colour, DNA in green. Arrows are in 5' to 3' direction.

Fig. 5: Schematic view of the inventive method on the level of base pairs

Fig. 6: Expression profile of a yeast gene
(A) A conventional transcriptional profile of the yeast YGR203W gene without orientation is shown. Vertical lines mark the boundaries of the YGR203W gene as they were determined previously
(B) A transcriptional profile of the yeast YGR203W gene with orientation according to the inventive method is depicted. Reads mapped in the forward direction are shown in blue, in the reverse direction in red.

Fig. 7: Scatter plots comparing mouse mRNA expression data (number of reads in annotated genes)
(A) The same mouse liver sample, strand-specific (according to the method, X-axis) and strand-unspecific (RNA-Seq, Y-axis) protocols (Pearson correlation coefficient (cc) = 0.999). A distribution close to unity would corresponds to a high correlation in these scatter plots.
(B) results according to the method for two biological replicas (mouse whole brain mRNA); cc=0.990.
(C) Mouse whole brain expression data according to the method (X- axis) and data from Mortazavi et al. 2008 (Y-axis); cc=0.817.
(D) Sense (X-axis) and antisense (Y-axis) expression in mouse brain. Thus the inventive method allows for a good differentiation of sense and antisense transcripts.

Fig. 8: Comparison of ssRNA-Seq results for protocols using 75%/25% dUTP/dTTP mixture and 100% dUTP
(A) Scatter plot of levels of sense transcription.
(B) Scatter plot of levels of antisense transcription.
Still a good correlation can be reached. Thus the inventive method does not depend on a complete exchange of at least one dNTP by a degradable analogue.

Fig. 9: Comparison of ssRNA-Seq results for protocols with and without actinomycin D added to the reverse transcription reaction
(A) Scatter plot of levels of sense transcription.
(B) Scatter plot of levels of antisense transcription.
X-Axe: with actinomycin D; Y-Axe: without actinomycin D. The use of actinomycin D does not bias significantly the correlation when using sense strands.

Fig. 10: Examples of overlapping genes and novel gene-candidates
A. *S. cerevisiae* transcriptome. Overlapping tails of two genes:
B. *S. cerevisiae* transcriptome. Overlapping of three genes:
C. *S. cerevisiae* transcriptome. Novel gene-candidate between
D. Mouse whole brain transcriptome. Overlapping genes: Ncaph2 and Ecgf1.
E. Mouse whole brain transcriptome. genes: Mrpl24 and BC023814.
F. Mouse whole brain transcriptome. One unannotated and one incorrectly annotated exon of Cdc42bpa mouse gene.
G. Mouse whole brain transcriptome. Novel exon of Chd3 gene.
H. Mouse whole brain transcriptome. Novel gene-candidate.

ssRNA-Seq stands for single-strand RNA sequencing carried out according to the method.

IDT Exhibit 1020.9428

Examples

Example 1: Transcriptome of Saccharomyces cerevisiae

[0124] In this example an application of the method according to the invention in a preferred embodiment is presented. Strand-specific sequencing of the yeast transcriptome is performed. Subsequently, the sequence of each clone is determined and matched with already existing DNA sequence of the yeast genome.

RNA isolation:

[0125] Yeast strain BY4741 (MATa; *his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*) was grown in rich medium (YPD; BD Company) at 30°C overnight, diluted to an OD_{600} of 0.15 and grown until reaching an OD_{600} of 0.87. The cells were harvested by centrifugation at room temperature, washed once with 1xPBS, and frozen in liquid nitrogen. Total RNA was extracted using the RiboPure™-Yeast kit (Ambion) and analyzed by an Agilent 2100 bioanalyzer (Agilent Technologies).

polyA+ RNA purification:

[0126] polyA+ RNA was purified with the Dynabeads mRNA purification kit (Invitrogen) following manufacturer's instructions and treated for 30 min at 37°C with 0.2 u TURBO™ DNase (Ambion) per 1 μg of RNA.

First strand synthesis (FSS):

[0127] A typical FSS reaction was prepared by mixing 0.5 μg of polyA+ RNA, 40 ng of $(dN)_6$ primers (Invitrogen) and 25 pmol of oligo(dT) primer (Invitrogen) in 8.5 μl of 1x reverse transcription buffer (Invitrogen), 0.5 mM dNTP's, 5 mM $MgCl_2$ and 10 mM DTT. The mixture was incubated at 98°C for 1 min to melt RNA secondary structures, then at 70°C for 5 min and was ramped to 15°C at 0.1°C/sec. Slow temperature ramp was used to make secondary structures of RNA and primers annealing as reproducible as possible. At 15°C 0.5 μl of actinomycin D solution (120 ng/μl), 0.5 μl of RNase OUT (40 u/μl, Invitrogen) and 0.5 μl of Superscript III polymerase (200 u/μl, Invitrogen) were added to the reaction. Reverse transcription reaction temperature was increased gradually as a compromise between survival of the enzyme, stability of the primers and denaturation of RNA secondary structures: ramp from 15 to 25°C at 0.1 °C/sec; incubation at 25°C for 10 min; ramp from 25 to 42°C at 0.1 °C/sec; incubation at 42°C for 45 min; ramp from 42 to 50°C at 0.1°C/sec; incubation at 50°C for 25 min. Superscript III polymerase was finally inactivated at 75°C for 15 min.

Removal of dNTPs:

[0128] 20 μl of EB (10 mM TrisCl, pH 8.5, Qiagen) was added to the reaction. dNTPs were removed by purification of the first strand mixture on a self-made 200 μl G-50 gel-filtration spin-column equilibrated with 1 mM TrisCl, pH 7.0.

Second strand synthesis (SSS):

[0129] The Invitrogen kit for second strand synthesis was used. Therefore the first strand synthesis buffer was restored after gel-filtration column. It is possible to prepare self-made buffer and reduce the number of buffers. Water was added to the purified FSS reaction to bring the final volume to 52.5 μl. The mixture was cooled on ice. Then 22.5 μl of the "second strand mixture" (1 μl of 1O x reverse transcription buffer (Invitrogen); 0.5 μl of 100 mM $MgCl_2$; 1 μl of 0.1 M DTT; 2 μl of 10 mM mixture of each: dATP, dGTP, dCTP, dUTP; 15 μl of 5x SSS buffer (Invitrogen); 0.5 μl of E.coli ligase (10 u/μl, NEB); 2 μl of DNA polymerase I (10 u/μl, NEB) and 0.5 μl RNase H (2u/μl, Invitrogen)) were added. SSS reaction was incubated at 16°C for 2 hours. ds cDNA was purified on QIAquick columns (Qiagen) according to the manufacturer's instructions.

DNA fragmentation:

[0130] About 250 ng of double-stranded cDNA was fragmented by sonication on a UTR200 (Hielscher Ultrasonics GmbH, Germany) under the following conditions: 1 hour, 50% pulse, 100% power, and continuous cooling by 0°C water flow-through.

Preparation of libraries for Illumina sequencing platform:

[0131] Libraries were prepared using the DNA sample kit (#FC-102-1002, Illumina), as described in (Sultan et al.

IDT Exhibit 1020.9429

2008), but with the following modifications: just before library amplification uridine digestion was performed at 37°C for 15 min in 5 μl of 1 x TE buffer, pH 7.5 with 1 u UNG (Applied Biosystems).

[0132] Procedure of paired-end sequencing library preparation was the same except that other ligation adapters and PCR primers were used (#PE-102-1002, Illumina).

Sequencing:

[0133] Amplified material was loaded onto flowcell at 4 pM concentration. Sequencing was carried out on the Illumina 1 G Genome Analyser by running 36 cycles according to the manufacturer's instructions.

Data analysis:

[0134] Images deconvolution, quality values calculation, mapping of exonic reads and exon junctions were performed as described in the art. Sequencing reads were aligned to *S. cerevisiae* (UCSC sacCerl) genomes using modification of Eland software (Gerald module v.1.27, Illumina). The mapping criteria of Eland are the following: sequencing read should be uniquely match to the genome allowing up to two mismatches, without insertions or deletions. The following recursive modification of Eland procedure was applied: first 32 bp of reads (trimming the last 4 bp of 36 bp reads due to Eland limitations) were aligned, then reads which do not match on reference by Eland criteria were trimmed to 31 bp, and aligned again. This 3'-end trimming of unmatched reads was done recursively down to 25 bp length. The modified procedure increases gain of uniquely aligned fragments typically by 20-50%, because the sequencing errors, which prevent successful alignment by Eland criteria, are located mostly on reads ends, which are gradually trimmed. Under these conditions, ~60% of the reads obtained here are matched to unique locations of the reference genome, whereas ~25% of the reads map to more than one genomic position and about 15% do not map.

Mapping end tags:

[0135] Unmapped sequencing reads with 1-11 nt long leading oligo(dT) stretches were used for mapping of the 3'-gene boundaries. Leading oligo(dT) stretches were removed, and if the remaining fragment length was longer or equal to 25 bp, it was aligned on reference genome.

Search for novel genes:

[0136] The search for novel genes in yeast using RNA-Seq was previously performed. Such an algorithm was not used herein, since it was not accurately described in the art. To estimate roughly how the knowledge of RNA orientation influences the number of novel gene-candidates the following procedure was used. A search algorithm presented below gives 377 novel gene candidates without taking into account RNA orientation. It is about the same number as in the art. To use information about RNA orientation the search was performed twice (for reads mapped in the forward direction and then for reads mapped in the reverse direction), yielding in 549 novel gene candidates in forward orientation and 512 - in reverse. The whole genome was split into 50 bp windows (non-sliding). A "new gene" was defined as a joined group of more than 2 consecutive windows, with at least 2 sequence reads (of the same direction) mapped per window. The gap between "new genes" should be at least 50 bp, the gap between "new gene" and an annotated gene (with the same transcription direction as the "new gene") - at least 100 bp.

[0137] With the RNA-Seq procedure it is possible to determine the 3'-boundaries of genes using those sequencing reads, which overlap with 3'-borders of genes. These reads may be mapped to the reference genome only after removal of oligo(dT) tail. The inventive method has the advantage of reduced noise in comparison with the RNA-Seq protocol because only one orientation of the homopolymeric stretch is allowed.

Example 2: Transcriptome of the brain of *Mus musculus*

[0138] In this example a further application of the method according to the invention is presented. Strand-specific sequencing of the whole mouse brain transcriptome is performed. Subsequently, the sequence of each clone is determined and matched with already existing DNA sequence of the mouse genome. Methods are similar to those used in Example 1. Therefore only steps which differ from Example 1 are indicated.

RNA isolation:

[0139] Two eleven weeks old female mice (C57B1/6J) were dissected and the whole brain was taken for RNA preparation. Total RNA was extracted using the Trizol method.

IDT Exhibit 1020.9430

Data analysis:

[0140] The resulting sequencing reads were aligned to the *M. musculus* (UCSC mm9) genomes using modification of Eland software (Gerald module v.1.27, Illumina).

Example 3: Comparison of strand-specific and non-strand-specific amplification of polynucleotides

[0141] To demonstrate that second-strand synthesis with deoxyuridine does not disturb the transcriptional landscape both strand-specific (according to the inventive method) and non-strand-specific transcriptome (RNA-Seq) sequencing for the same RNA sample was performed (mouse liver transcriptome). The resulting scatter plot (Fig. 7A) shows that both RNA-Seq and the inventive method produce identical transcription patterns.

Example 4: Reproducibiliity of the inventive method and comparison with previously published results from another laboratory

[0142] In order to test the reproducibility of the inventive method independent analyses of whole brain mouse (Fig. 7B) transcriptomes were carried out under identical conditions. A high reproducibility can be seen in the respective scatter plots. A comparison of the data with previously published results also demonstrates good correlation (Fig. 7C).

Example 5: Stability of the inventive method

[0143] The inventive method results in a high degree of certainty in identifying transcript polarity. In theory, one uridine base in a molecule of a sequencing library is enough to prevent the appearance of a sequencing read from the false strand. Even if UNG occasionally does not remove the uridine base, the molecule still would not be amplified, since uridine-containing template strongly suppresses Phusion DNA polymerase used for library amplification. Apparently, most second cDNA strands in the library contain more than one uridine base, because a 75%/25% dUTP/dTTP mixture in the second strand synthesis reaction gives essentially the same results as 100% dUTP (Fig. 8A and 8B).

Example 6: Influence of actinomycin D

[0144] To prevent spurious second-strand cDNA synthesis, which was shown to be the major source of artifactual antisense transcripts. Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. Actinomycin D was included in the reverse transcription reaction. Actinomycin D specifically inhibits DNA-dependent, but not RNA-dependent, DNA synthesis. The presence of actinomycin D in the reaction practically did not influence the level of antisense transcription (Fig. 9A and 9B). Actinomycin D is included in the protocol to make the reaction more reproducible.

IDT Exhibit 1020.9431

SEQUENCE LISTING

<110> Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.

<120> Method for differentiation of polynucleotide strands

<130> GAR-P03013WO

<140> 000000
<141> 2010-07-06

<160> 39

<170> PatentIn version 3.3

<210> 1
<211> 113
<212> RNA
<213> Artificial

<220>
<223> first-strand synthesis: primers annealing

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (8)..(13)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (16)..(20)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (22)..(24)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (26)..(31)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (34)..(44)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (46)..(49)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (51)..(57)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (59)..(66)

IDT Exhibit 1020.9432

<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (75)..(76)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (78)..(79)
<223> n is a, c, g, or u

<400> 1
nnnnnnunnn nnnuunnnnn unnnunnnnn nuunnnnnnn nnnnunnnnu nnnnnnnunn 60

nnnnnnuuun nnnunnunna aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 113

<210> 2
<211> 6
<212> DNA
<213> Artificial

<220>
<223> Primer first-strand synthesis

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 2
nnnnnn 6

<210> 3
<211> 20
<212> DNA
<213> Artificial

<220>
<223> oligo (dT) first-strand synthesis

<400> 3
tttttttttt tttttttttt 20

<210> 4
<211> 42
<212> DNA
<213> Artificial

<220>
<223> first-strand synthesis: reverse transcription; cDNA starting from primer

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

IDT Exhibit 1020.9433

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(42)
<223> n is a, c, g, or t

<400> 4
nnnnnnnnnn naannnnnna nnnannnnna annnnnnann nn 42

<210> 5
<211> 50
<212> DNA
<213> Artificial

<220>
<223> first-strand synthesis: reverse transcription; cDNA starting from oligo (dT)

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (30)..(33)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (37)..(44)
<223> n is a, c, g, or t

<400> 5
ttttttttttt ttttttttttt tttnnannan nnnaaannnn nnnnaaaaaa 50

<210> 6
<211> 102
<212> DNA
<213> Artificial

IDT Exhibit 1020.9434

<220>
<223> Second-strand synthesis: Second Strand

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (8)..(13)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (16)..(20)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (22)..(24)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (26)..(31)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (34)..(44)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (46)..(49)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (51)..(57)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (59)..(66)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (75)..(76)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (78)..(79)
<223> n is a, c, g, t or u

<400> 6
nnnnnnunnn nnnuunnnnn unnnunnnnn nuunnnnnnn nnnnunnnnu nnnnnnnunn 60

nnnnnnuuun nnnunnunna aaaaaaaaaa aaaaaaaaaa aa 102

IDT Exhibit 1020.9435

<210> 7
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion): Fragment A

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(36)
<223> n is a, c, g, t or u

<400> 7
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnn 36

<210> 8
<211> 47
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion): Fragment B (second strand)

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

IDT Exhibit 1020.9436

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 8
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaaaaa 47

<210> 9
<211> 42
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion) Fragment B (first strand)

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<400> 9
ttttttttt ttttttttt nnannannnn aaannnnnnn na 42

<210> 10
<211> 39
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment A (second strand))

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>

IDT Exhibit 1020.9437

<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(39)
<223> n is a, c, g, t or u

<400> 10
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnnnnn 39

<210> 11
<211> 39
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment A first strand

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(39)
<223> n is a, c, g, or t

<400> 11
nnnnnnnnnn naannnnnna nnnannnnna annnnnnan 39

<210> 12
<211> 44

IDT Exhibit 1020.9438

<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment B second strand

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 12
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaa 44

<210> 13
<211> 44
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment B first strand

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (43)..(44)

IDT Exhibit 1020.9439

<223> n is a, c, g, or t

<400> 13
tttttttttt tttttttttt nnannannnn aaannnnnnn nann 44

<210> 14
<211> 40
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment A second strand

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(39)
<223> n is a, c, g, t or u

<400> 14
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnnnnna 40

<210> 15
<211> 40
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment A first strand

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

IDT Exhibit 1020.9440

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(39)
<223> n is a, c, g, or t

<400> 15
nnnnnnnnnn naannnnnna nnnannnnna annnnnnana 40

<210> 16
<211> 45
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment B second strand

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 16
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaaa 45

<210> 17
<211> 45
<212> DNA
<213> Artificial

<220>

IDT Exhibit 1020.9441

<223> A-tailing Fragment B first strand

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (43)..(44)
<223> n is a, c, g, or t

<400> 17
ttttttttt ttttttttt nnannannnn aaannnnnnn nanna 45

<210> 18
<211> 107
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment A second strand

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (37)..(42)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (45)..(49)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (51)..(53)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (55)..(60)
<223> n is a, c, g, t or u

<220>
<221> misc_feature

IDT Exhibit 1020.9442

<222> (63)..(68)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, t or u

<400> 18
tacactcttt ccctacacga cgctcttccg atctnunnnn nnuunnnnnu nnnunnnnnn 60

uunnnnnnun nnnagatcgg aagagctcgt atgccgtctt ctgcttg 107

<210> 19
<211> 106
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment A first strand

<220>
<221> misc_feature
<222> (35)..(45)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (48)..(53)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (55)..(56)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (58)..(62)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (65)..(70)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (72)..(72)
<223> n is a, c, g, or t

<400> 19
tacactcttt ccctacacga cgctcttccg atctnnnnnn nnnnnaannn nnnannannn 60

nnaannnnnn anagatcgga agagctcgta tgccgtcttc tgcttg 106

<210> 20
<211> 112
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment B second strand

IDT Exhibit 1020.9443

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (38)..(45)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (49)..(52)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (54)..(55)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (57)..(58)
<223> n is a, c, g, t or u

<400> 20
tacactcttt ccctacacga cgctcttccg atctnnunnn nnnnnuuunn nnunnunnaa 60

aaaaaaaaaa aaaaaaaaag atcggaagag ctcgtatgcc gtcttctgct tg 112

<210> 21
<211> 112
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment B first strand

<220>
<221> misc_feature
<222> (55)..(56)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (58)..(59)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (61)..(64)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (68)..(75)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (77)..(78)
<223> n is a, c, g, or t

<400> 21
tacactcttt ccctacacga cgctcttccg atctttttttt ttttttttttt ttttnnanna 60

IDT Exhibit 1020.9444

nnnnaaannn nnnnnannag atcggaagag ctcgtatgcc gtcttctgct tg 112

<210> 22
<211> 35
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment Fragment A second strand piece 1

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, or t

<400> 22
tacactcttt ccctacacga cgctcttccg atctn 35

<210> 23
<211> 6
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second sttrand piece 2

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 23
nnnnnn 6

<210> 24
<211> 5
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 3

<220>
<221> misc_feature
<222> (1)..(5)
<223> n is a, c, g, or t

<400> 24
nnnnn 5

<210> 25
<211> 3
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 4

<220>
<221> misc_feature

IDT Exhibit 1020.9445

<222> (1)..(3)
<223> n is a, c, g, or t

<400> 25
nnn 3

<210> 26
<211> 6
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 5

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 26
nnnnnn 6

<210> 27
<211> 44
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment Fragment A second strand piece 6

<220>
<221> misc_feature
<222> (1)..(10)
<223> n is a, c, g, or t

<400> 27
nnnnnnnnnn agatcggaag agctcgtatg ccgtcttctg cttg 44

<210> 28
<211> 36
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 1

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, or t

<400> 28
tacactcttt ccctacacga cgctcttccg atctnn 36

<210> 29
<211> 8
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 2

IDT Exhibit 1020.9446

<220>
<221> misc_feature
<222> (1)..(8)
<223> n is a, c, g, or t

<400> 29
nnnnnnnn 8

<210> 30
<211> 4
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 3

<220>
<221> misc_feature
<222> (1)..(4)
<223> n is a, c, g, or t

<400> 30
nnnn 4

<210> 31
<211> 2
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 4

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, or t

<400> 31
nn 2

<210> 32
<211> 56
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 5

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, or t

<400> 32
nnaaaaaaaa aaaaaaaaaa aaagatcgga agagctcgta tgccgtcttc tgcttg 56

<210> 33
<211> 34
<212> DNA

IDT Exhibit 1020.9447

<213> Artificial

<220>
<223> Preamplification fragment A second strand part 1

<400> 33
caagcagaag acggcatacg agctcttccg atct 34

<210> 34
<211> 58
<212> DNA
<213> Artificial

<220>
<223> Preamplification fragment A second strand part 2

<400> 34
tctagccttc tcgcagcaca tccctttctc acatctagag ccaccagcgg catagtaa 58

<210> 35
<211> 131
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment A superior

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (37)..(42)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (45)..(49)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (51)..(53)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (55)..(60)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (63)..(73)
<223> n is a, c, g, or t

<400> 35
caagcagaag acggcatacg agctcttccg atctntnnnn nnttnnnnnt nnntnnnnnn 60

ttnnnnnnnn nnnagatcgg aagagcgtcg tgtagggaaa gagtgtagat ctcggtggtc 120

gccgtatcat t 131

IDT Exhibit 1020.9448

<210> 36
<211> 131
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment A inferior

<220>
<221> misc_feature
<222> (59)..(69)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (72)..(77)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (79)..(81)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (83)..(87)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (90)..(95)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (97)..(97)
<223> n is a, c, g, or t

<400> 36
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctnn 60

nnnnnnnnna annnnnnann nannnnnaan nnnnnanaga tcggaagagc tcgtatgccg 120

tcttctgctt g 131

<210> 37
<211> 136
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment B superior

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (38)..(45)
<223> n is a, c, g, or t

<220>
<221> misc_feature

IDT Exhibit 1020.9449

<222> (49)..(52)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (54)..(55)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (57)..(58)
<223> n is a, c, g, or t

<400> 37
caagcagaag acggcatacg agctcttccg atctnntnnn nnnnntttnn nntnntnnaa 60

aaaaaaaaaa aaaaaaaaag atcggaagag cgtcgtgtag ggaaagagtg tagatctcgg 120

tggtcgccgt atcatt 136

<210> 38
<211> 136
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment B inferior

<220>
<221> misc_feature
<222> (79)..(80)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (82)..(83)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (85)..(88)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (92)..(99)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (101)..(102)
<223> n is a, c, g, or t

<400> 38
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatcttt 60

ttttttttt ttttttttnn annannnnaa annnnnnnna nnagatcgga agagctcgta 120

tgccgtcttc tgcttg 136

<210> 39
<211> 33
<212> DNA
<213> Artificial

IDT Exhibit 1020.9450

<220>
<223> Sequencing primer annealing: sequencing primer

<400> 39
acactctttc cctacacgac gctcttccga tct 33

IDT Exhibit 1020.9451

SEQUENCE LISTING

<110> Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.

<120> Method for differentiation of polynucleotide strands

<130> GAR-P03013EP

<140> EP09008808.9
<141> 2009-07-06

<160> 39

<170> PatentIn version 3.3

<210> 1
<211> 113
<212> RNA
<213> Artificial

<220>
<223> first-strand synthesis: primers annealing

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (8)..(13)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (16)..(20)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (22)..(24)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (26)..(31)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (34)..(44)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (46)..(49)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (51)..(57)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (59)..(66)

IDT Exhibit 1020.9452

<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (75)..(76)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (78)..(79)
<223> n is a, c, g, or u

<400> 1
nnnnnnnunnn nnnuunnnnn unnnunnnnn nuunnnnnnn nnnnunnnnu nnnnnnnunn 60

nnnnnnuuun nnnunnunna aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 113

<210> 2
<211> 6
<212> DNA
<213> Artificial

<220>
<223> Primer first-strand synthesis

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 2
nnnnnn 6

<210> 3
<211> 20
<212> DNA
<213> Artificial

<220>
<223> oligo (dT) first-strand synthesis

<400> 3
ttttttttt ttttttttt 20

<210> 4
<211> 42
<212> DNA
<213> Artificial

<220>
<223> first-strand synthesis: reverse transcription; cDNA starting from primer

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

IDT Exhibit 1020.9453

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(42)
<223> n is a, c, g, or t

<400> 4
nnnnnnnnnn naannnnnna nnnannnnna annnnnnann nn 42

<210> 5
<211> 50
<212> DNA
<213> Artificial

<220>
<223> first-strand synthesis: reverse transcription; cDNA starting from oligo (dT)

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (30)..(33)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (37)..(44)
<223> n is a, c, g, or t

<400> 5
ttttttttt ttttttttt tttnnannan nnnaaannnn nnnnaaaaaa 50

<210> 6
<211> 102
<212> DNA
<213> Artificial

IDT Exhibit 1020.9454

<220>
<223> Second-strand synthesis: Second Strand

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (8)..(13)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (16)..(20)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (22)..(24)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (26)..(31)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (34)..(44)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (46)..(49)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (51)..(57)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (59)..(66)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (75)..(76)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (78)..(79)
<223> n is a, c, g, t or u

<400> 6
nnnnnnunnn nnnuunnnnn unnnunnnnn nuunnnnnnn nnnnunnnnu nnnnnnnunn 60

nnnnnnuuun nnnunnunna aaaaaaaaaa aaaaaaaaaa aa 102

IDT Exhibit 1020.9455

<210> 7
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion): Fragment A

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(36)
<223> n is a, c, g, t or u

<400> 7
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnn 36

<210> 8
<211> 47
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion): Fragment B (second strand)

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

IDT Exhibit 1020.9456

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 8
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaaaaa 47

<210> 9
<211> 42
<212> DNA
<213> Artificial

<220>
<223> Fragmentation (ultrasound digestion) Fragment B (first strand)

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<400> 9
ttttttttttt ttttttttttt nnannannnn aaannnnnnn na 42

<210> 10
<211> 39
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment A (second strand))

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>

IDT Exhibit 1020.9457

<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(39)
<223> n is a, c, g, t or u

<400> 10
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnnnnn 39

<210> 11
<211> 39
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment A first strand

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(39)
<223> n is a, c, g, or t

<400> 11
nnnnnnnnnn naannnnnna nnnannnnna annnnnnan 39

<210> 12
<211> 44

IDT Exhibit 1020.9458

<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment B second strand

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 12
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaa 44

<210> 13
<211> 44
<212> DNA
<213> Artificial

<220>
<223> End-repair Fragment B first strand

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (43)..(44)

IDT Exhibit 1020.9459

<223> n is a, c, g, or t

<400> 13
ttttttttt ttttttttt nnannannnn aaannnnnnn nann 44

<210> 14
<211> 40
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment A second strand

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (3)..(8)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (11)..(15)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (17)..(19)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (21)..(26)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (29)..(39)
<223> n is a, c, g, t or u

<400> 14
nunnnnnnuu nnnnnunnnu nnnnnnuunn nnnnnnnnna 40

<210> 15
<211> 40
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment A first strand

<220>
<221> misc_feature
<222> (1)..(11)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (14)..(19)
<223> n is a, c, g, or t

IDT Exhibit 1020.9460

<220>
<221> misc_feature
<222> (21)..(23)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(29)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (32)..(37)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (39)..(39)
<223> n is a, c, g, or t

<400> 15
nnnnnnnnnn naannnnnna nnnannnnna annnnnnana 40

<210> 16
<211> 45
<212> DNA
<213> Artificial

<220>
<223> A-tailing Fragment B second strand

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (4)..(11)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (15)..(18)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, t or u

<400> 16
nnunnnnnnn nuuunnnnun nunnaaaaaa aaaaaaaaaa aaaaa 45

<210> 17
<211> 45
<212> DNA
<213> Artificial

<220>

IDT Exhibit 1020.9461

<223> A-tailing Fragment B first strand

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(30)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34)..(41)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (43)..(44)
<223> n is a, c, g, or t

<400> 17
ttttttttt ttttttttt nnannannnn aaannnnnnn nanna 45

<210> 18
<211> 107
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment A second strand

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (37)..(42)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (45)..(49)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (51)..(53)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (55)..(60)
<223> n is a, c, g, t or u

<220>
<221> misc_feature

IDT Exhibit 1020.9462

<222> (63)..(68)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (70)..(73)
<223> n is a, c, g, t or u

<400> 18
tacactcttt ccctacacga cgctcttccg atctnunnnn nnuunnnnnu nnnunnnnnn 60

uunnnnnnun nnnagatcgg aagagctcgt atgccgtctt ctgcttg 107

<210> 19
<211> 106
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment A first strand

<220>
<221> misc_feature
<222> (35)..(45)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (48)..(53)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (55)..(56)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (58)..(62)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (65)..(70)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (72)..(72)
<223> n is a, c, g, or t

<400> 19
tacactcttt ccctacacga cgctcttccg atctnnnnnn nnnnnaannn nnnannannn 60

nnaannnnnn anagatcgga agagctcgta tgccgtcttc tgcttg 106

<210> 20
<211> 112
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment B second strand

IDT Exhibit 1020.9463

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (38)..(45)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (49)..(52)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (54)..(55)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (57)..(58)
<223> n is a, c, g, t or u

<400> 20
tacactcttt ccctacacga cgctcttccg atctnnunnn nnnnnuuunn nnunnunnaa 60

aaaaaaaaaa aaaaaaaaag atcggaagag ctcgtatgcc gtcttctgct tg 112

<210> 21
<211> 112
<212> DNA
<213> Artificial

<220>
<223> Adaptor ligation Fragment B first strand

<220>
<221> misc_feature
<222> (55)..(56)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (58)..(59)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (61)..(64)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (68)..(75)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (77)..(78)
<223> n is a, c, g, or t

<400> 21
tacactcttt ccctacacga cgctcttccg atctttttttt ttttttttttt ttttnnanna 60

IDT Exhibit 1020.9464

nnnnaaannn nnnnnannag atcggaagag ctcgtatgcc gtcttctgct tg 112

<210> 22
<211> 35
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment Fragment A second strand piece 1

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, or t

<400> 22
tacactcttt ccctacacga cgctcttccg atctn 35

<210> 23
<211> 6
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second sttrand piece 2

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 23
nnnnnn 6

<210> 24
<211> 5
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 3

<220>
<221> misc_feature
<222> (1)..(5)
<223> n is a, c, g, or t

<400> 24
nnnnn 5

<210> 25
<211> 3
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 4

<220>
<221> misc_feature

IDT Exhibit 1020.9465

<222> (1)..(3)
<223> n is a, c, g, or t

<400> 25
nnn 3

<210> 26
<211> 6
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment A second strand piece 5

<220>
<221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, or t

<400> 26
nnnnnn 6

<210> 27
<211> 44
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment Fragment A second strand piece 6

<220>
<221> misc_feature
<222> (1)..(10)
<223> n is a, c, g, or t

<400> 27
nnnnnnnnnn agatcggaag agctcgtatg ccgtcttctg cttg 44

<210> 28
<211> 36
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 1

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, or t

<400> 28
tacactcttt ccctacacga cgctcttccg atctnn 36

<210> 29
<211> 8
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 2

IDT Exhibit 1020.9466

<220>
<221> misc_feature
<222> (1)..(8)
<223> n is a, c, g, or t

<400> 29
nnnnnnnn 8

<210> 30
<211> 4
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 3

<220>
<221> misc_feature
<222> (1)..(4)
<223> n is a, c, g, or t

<400> 30
nnnn 4

<210> 31
<211> 2
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 4

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, or t

<400> 31
nn 2

<210> 32
<211> 56
<212> DNA
<213> Artificial

<220>
<223> UDGase treatment fragment B second strand piece 5

<220>
<221> misc_feature
<222> (1)..(2)
<223> n is a, c, g, or t

<400> 32
nnaaaaaaaa aaaaaaaaaa aaagatcgga agagctcgta tgccgtcttc tgcttg 56

<210> 33
<211> 34
<212> DNA

IDT Exhibit 1020.9467

<213> Artificial

<220>
<223> Preamplification fragment A second strand part 1

<400> 33
caagcagaag acggcatacg agctcttccg atct 34

<210> 34
<211> 58
<212> DNA
<213> Artificial

<220>
<223> Preamplification fragment A second strand part 2

<400> 34
tctagccttc tcgcagcaca tccctttctc acatctagag ccaccagcgg catagtaa 58

<210> 35
<211> 131
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment A superior

<220>
<221> misc_feature
<222> (35)..(35)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (37)..(42)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (45)..(49)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (51)..(53)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (55)..(60)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (63)..(73)
<223> n is a, c, g, or t

<400> 35
caagcagaag acggcatacg agctcttccg atctntnnnn nnttnnnnnt nnntnnnnnn 60

ttnnnnnnnn nnnagatcgg aagagcgtcg tgtagggaaa gagtgtagat ctcggtggtc 120

gccgtatcat t 131

IDT Exhibit 1020.9468

<210> 36
<211> 131
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment A inferior

<220>
<221> misc_feature
<222> (59)..(69)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (72)..(77)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (79)..(81)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (83)..(87)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (90)..(95)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (97)..(97)
<223> n is a, c, g, or t

<400> 36
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctnn 60

nnnnnnnnna annnnnnann nannnnnaan nnnnnanaga tcggaagagc tcgtatgccg 120

tcttctgctt g 131

<210> 37
<211> 136
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment B superior

<220>
<221> misc_feature
<222> (35)..(36)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (38)..(45)
<223> n is a, c, g, or t

<220>
<221> misc_feature

IDT Exhibit 1020.9469

<222> (49)..(52)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (54)..(55)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (57)..(58)
<223> n is a, c, g, or t

<400> 37
caagcagaag acggcatacg agctcttccg atctnntnnn nnnnntttnn nntnntnnaa 60

aaaaaaaaaa aaaaaaaaag atcggaagag cgtcgtgtag ggaaagagtg tagatctcgg 120

tggtcgccgt atcatt 136

<210> 38
<211> 136
<212> DNA
<213> Artificial

<220>
<223> Amplification fragment B inferior

<220>
<221> misc_feature
<222> (79)..(80)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (82)..(83)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (85)..(88)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (92)..(99)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (101)..(102)
<223> n is a, c, g, or t

<400> 38
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatcttt 60

ttttttttt ttttttttnn annannnnaa annnnnnnna nnagatcgga agagctcgta 120

tgccgtcttc tgcttg 136

<210> 39
<211> 33
<212> DNA
<213> Artificial

IDT Exhibit 1020.9470

<220>
<223> Sequencing primer annealing: sequencing primer

<400> 39
acactctttc cctacacgac gctcttccga tct 33

Claims

1. A method for differentiation of polynucleotide strands comprising the following steps:

   I) Performing strand-specific synthesis on a polynucleotide template by providing polynucleotide template, primers, correspondent buffer, polymerase and sufficient amounts of the four dNTPs: dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, wherein at least one of the four dNTPs is a modified nucleotide used for strand specific differentiation;
   II) Optionally, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof, if said polynucleotide template on step I) is a single-stranded RNA;
   III) Optionally, performing molecular cloning operations with double-stranded DNA; and
   IV) Strand specific differentiation using different interaction of modified and unmodified nucleotides with some selection factor.

2. A method for differentiation of polynucleotide strands according to claim 1 comprising the following steps:

   I) Providing at least one polynucleotide template, primers, buffer solution, at least one enzyme with polymerase activity, sufficient amounts of the four dNTPs dATP, dCTP, dGTP and dTTP or analogues thereof, wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific differentiation;
   II) Performing specific strand synthesis on said at least one polynucleotide;
   and
   IIIa) Performing strand-specific differentiation by using a selection factor for differentiating strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide;
   or if said polynucleotide template is a single-stranded RNA,
   IIIb) Performing second-strand synthesis with all four unmodified dNTPs or analogues thereof after removing the at least one modified dNTP; and
   IVb) Performing strand-specific differentiation by using a selection factor for differentiating strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide;
   or
   IIIc) Performing one or more molecular cloning operations on the double-stranded DNA, wherein the molecular cloning operations are selected from shearing, restriction, digestion, end-polishing, A-tailing, G-tailing and adaptor ligation; and
   IVc) Performing strand-specific differentiation by using a selection factor for differentiating strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide;
   or if said polynucleotide template is a single-stranded RNA,
   IIId) Performing second-strand synthesis with all four unmodified dNTPs or analogues thereof after removing the at least one modified dNTP;
   IVd) Performing one or more molecular cloning operations on the double-stranded DNA, wherein the molecular cloning operations are selected from shearing, restriction, digestion, end-polishing, A-tailing, G-tailing and adaptor ligation; and
   Vd) Performing strand-specific differentiation by using a selection factor for differentiating strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide.

3. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said polynucleotide template is selected from single-stranded RNA, single-stranded DNA, double-stranded DNA and DNA/RNA hybrids.

4. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said selection factor for differentiating strands is at least one chemical agent modifying selectively the at least one modified nucleotide or one or more chemical groups of the at least one modified nucleotide.

IDT Exhibit 1020.9471

5. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said selection factor for differentiating strands is at least one factor displaying selective affinity to an affinity group of the at least one modified nucleotide.

6. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said selection factor for differentiating strands is one or more antibodies binding selectively to the at least one modified nucleotide.

7. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said selection factor for differentiating strands is a solid phase at which the strand containing the at least one modified nucleotide is immobilized.

8. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein said selection factor for differentiating strands is a different sensitivity to nucleases of the modified and the unmodified nucleotide.

9. The method for differentiation of polynucleotide strands according to claim 8, wherein said nuclease is Uracil-N-Glycosylase and said modified nucleotide is dUTP.

10. The method for differentiation of polynucleotide strands according to claim 5, wherein said factor displaying selective affinity to an affinity group of the at least one modified nucleotide is streptavidin or avidin and the at least one modified nucleotide is a biotinylated nucleotide.

11. The method for differentiation of polynucleotide strands according to claim 4, wherein the at least one modified nucleotide with said chemical group is Aminoallyl-d UTP.

12. The method for differentiation of polynucleotide strands according to claim 6, wherein the at least one modified nucleotide recognized selectively by an antibody is 5-bromodeoxyuridine.

13. The method for differentiation of polynucleotide strands according to claim 1 or 2, wherein the ratio of the at least one modified dNTP towards the corresponding unmodified dNTP or analogues thereof is at least 1:20.

14. The method for differentiation of polynucleotide strands according to claim 2, wherein if the polynucleotide strand is a single RNA strand the RNA is selected from mRNA, tRNA, rRNA, hnRNA, snRNA, snoRNA, scaRNA, asRNA, micro RNA, siRNA, short transcripts, synthetic RNA polynucleotides and semisynthetic RNA polynucleotides as well as physiological or externally induced modifications thereof, or if the polynucleotide strand is a DNA the DNA is selected from siDNA, single-stranded DNA copies from retroviral or hepadnaviral RNA and single-stranded DNA, single-stranded synthetic DNA polynucleotides and single-stranded semisynthetic DNA polynucleotides as well as physiological or externally induced modifications thereof.

15. A kit for performing differentiation of polynucleotide strands comprising:

A) All four dNTPs dATP, dCTP, dGTP and dTTP and corresponding modified deoxynucleotide;
B) Primers;
C) One or more DNA polymerases; and
D) One or more selection factors corresponding to the modified deoxynucleotide and/or one or more selective media for removing the strand containing the modified nucleotide;
or additionally comprising
E) One or more DNA polymerases with reverse transcriptase activity; or additionally comprising; and
F) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

IDT Exhibit 1020.9472

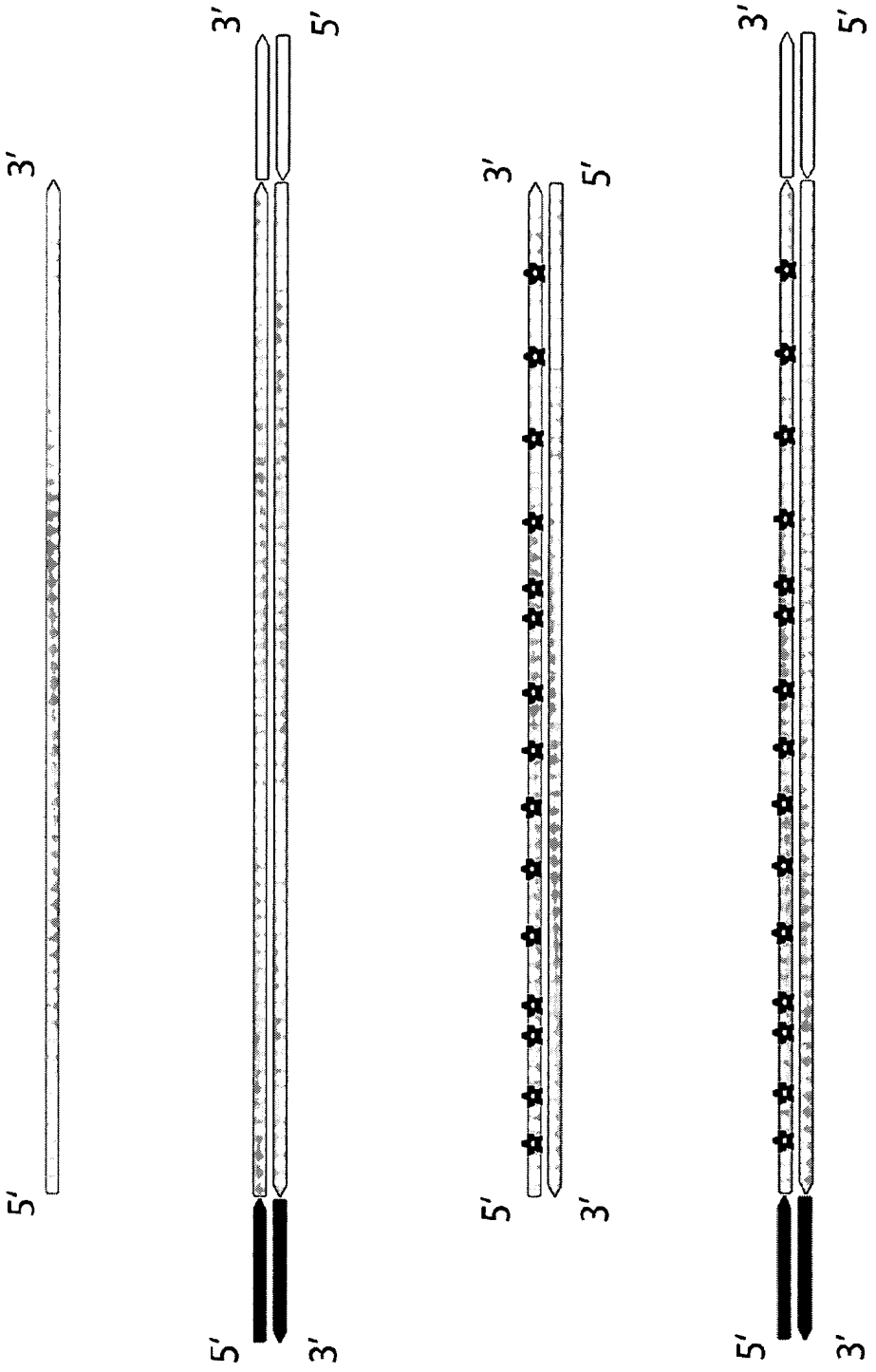


Figure 1

IDT Exhibit 1020.9473

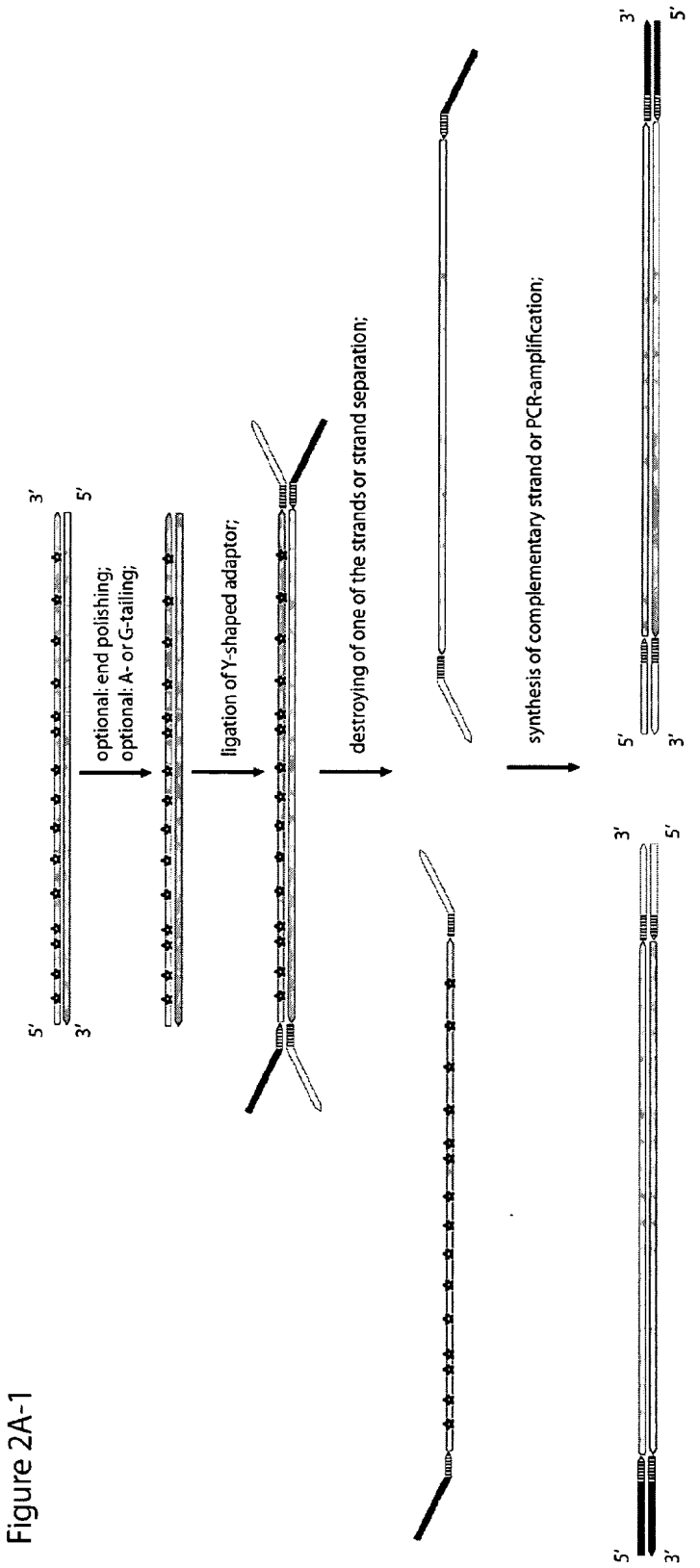


Figure 2A-1

IDT Exhibit 1020.9474

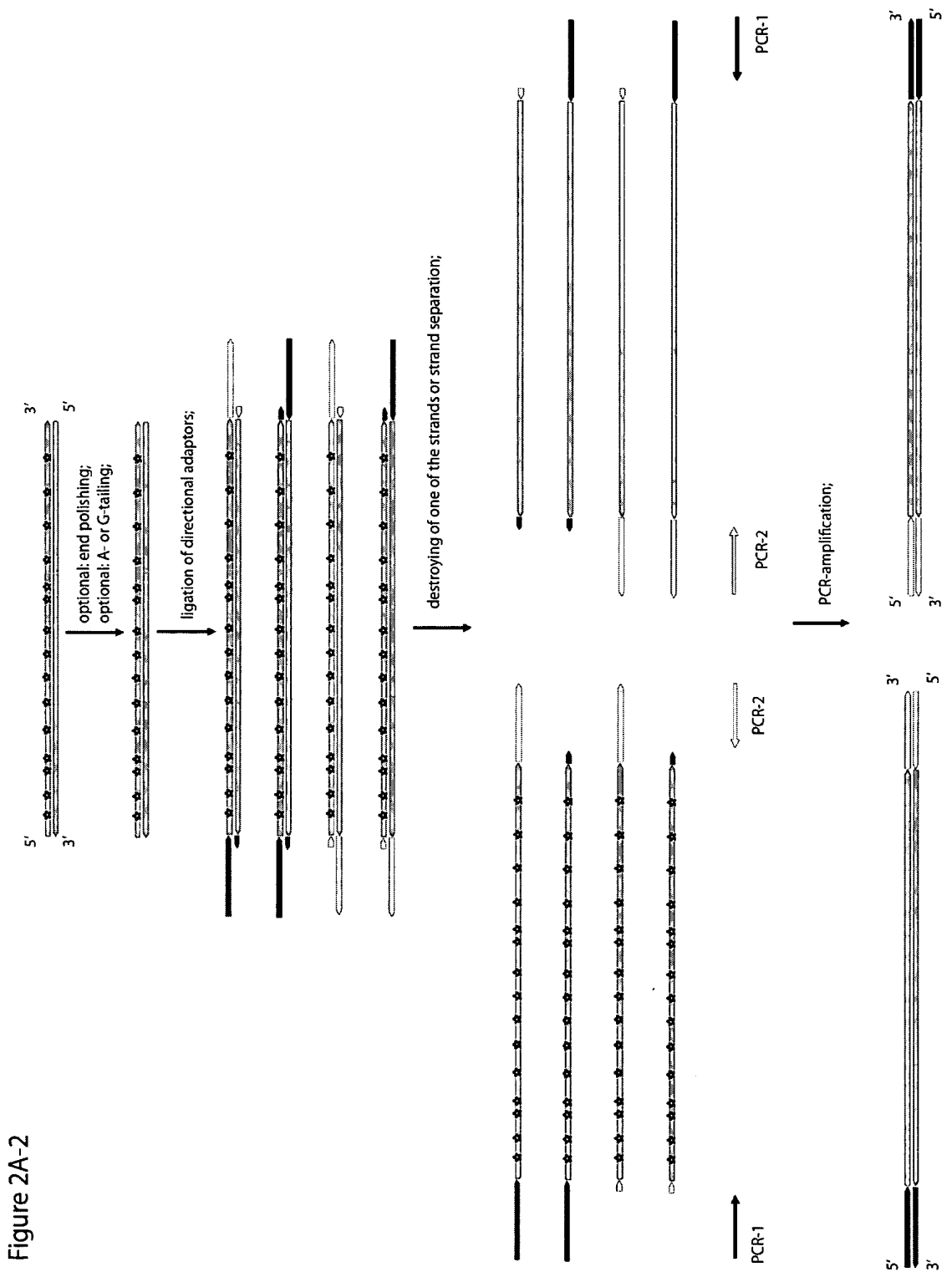


Figure 2A-2

IDT Exhibit 1020.9475

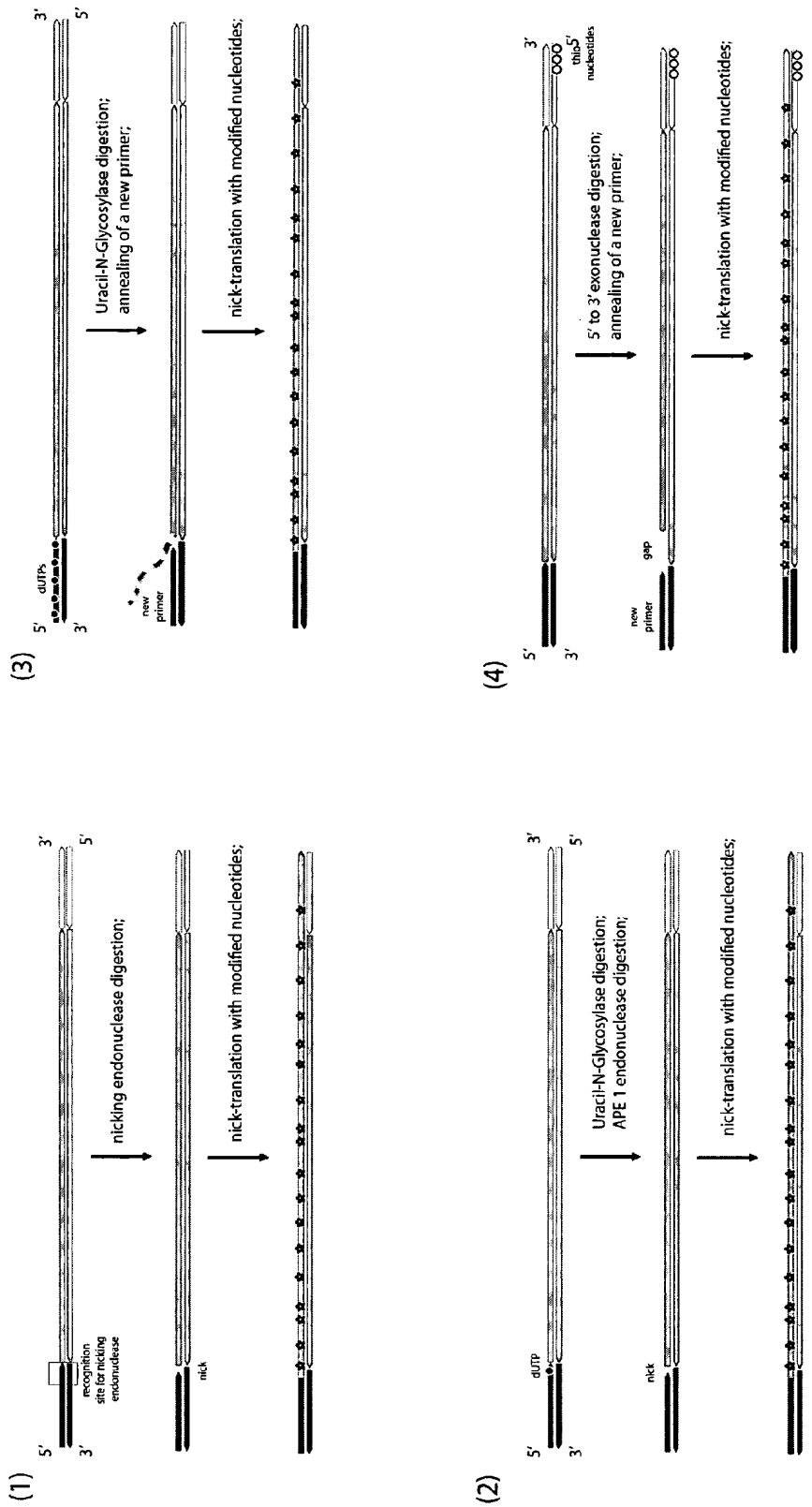


Figure 2B

IDT Exhibit 1020.9476

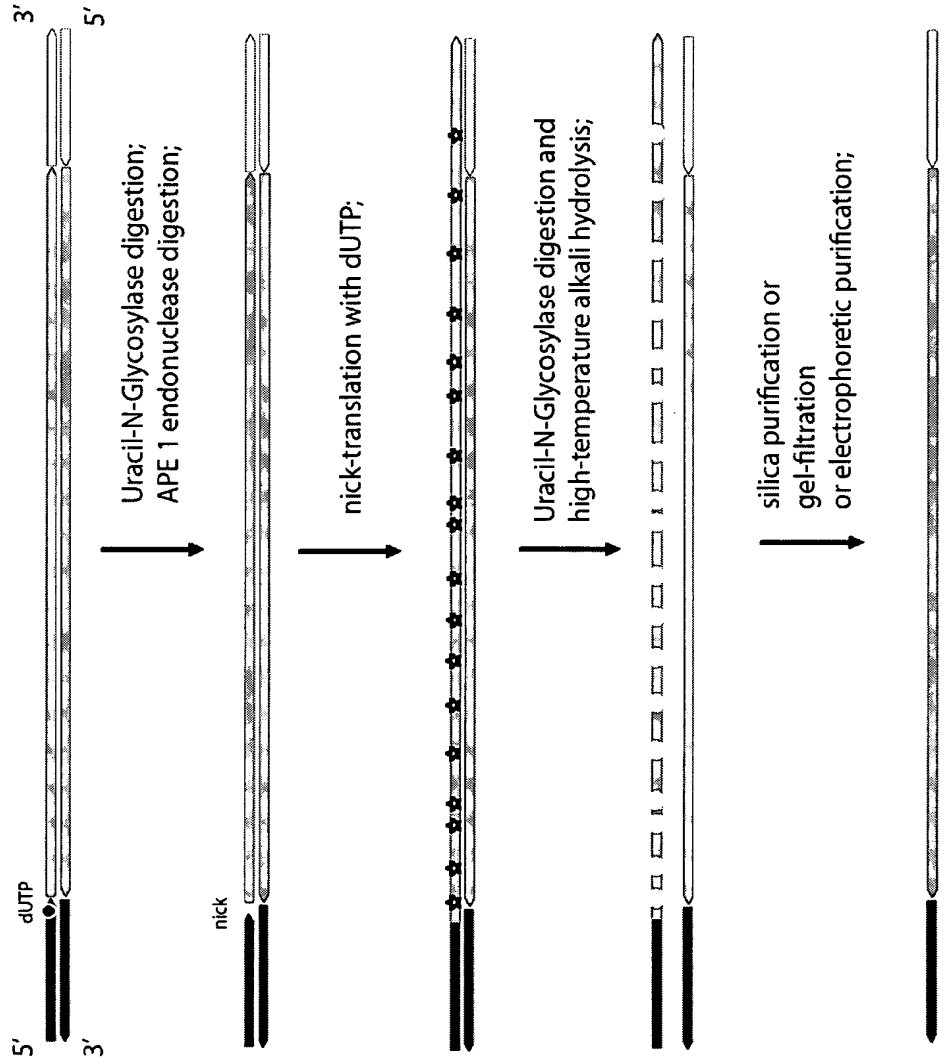


Figure 3A

IDT Exhibit 1020.9477

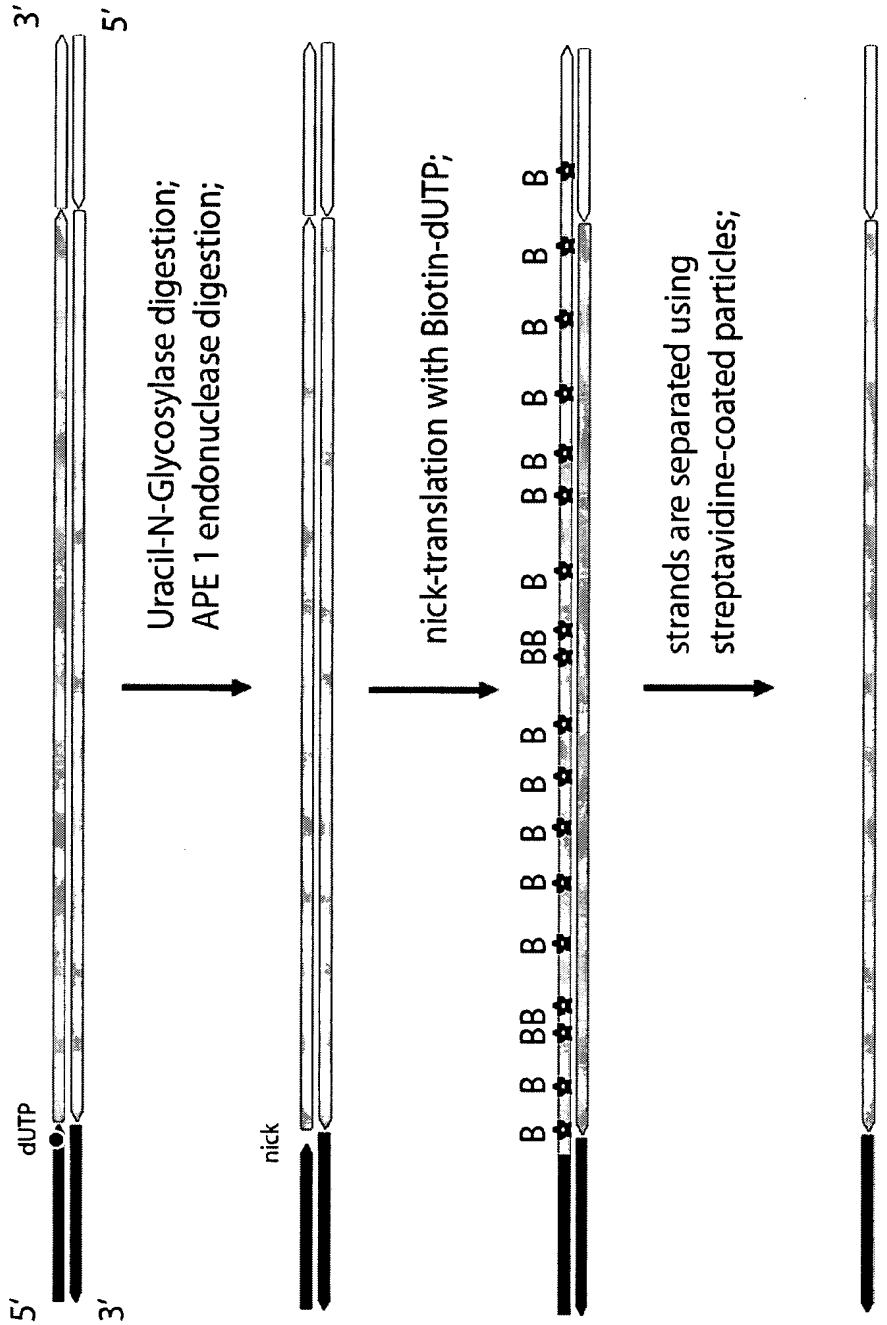


Figure 3B

IDT Exhibit 1020.9478

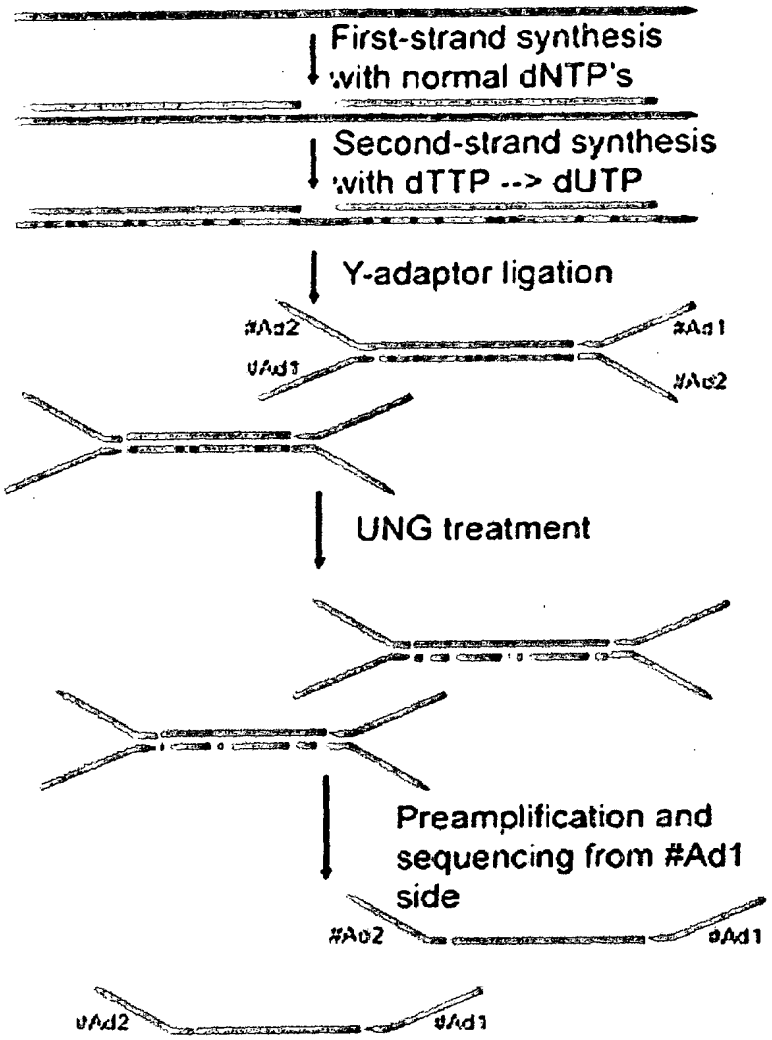
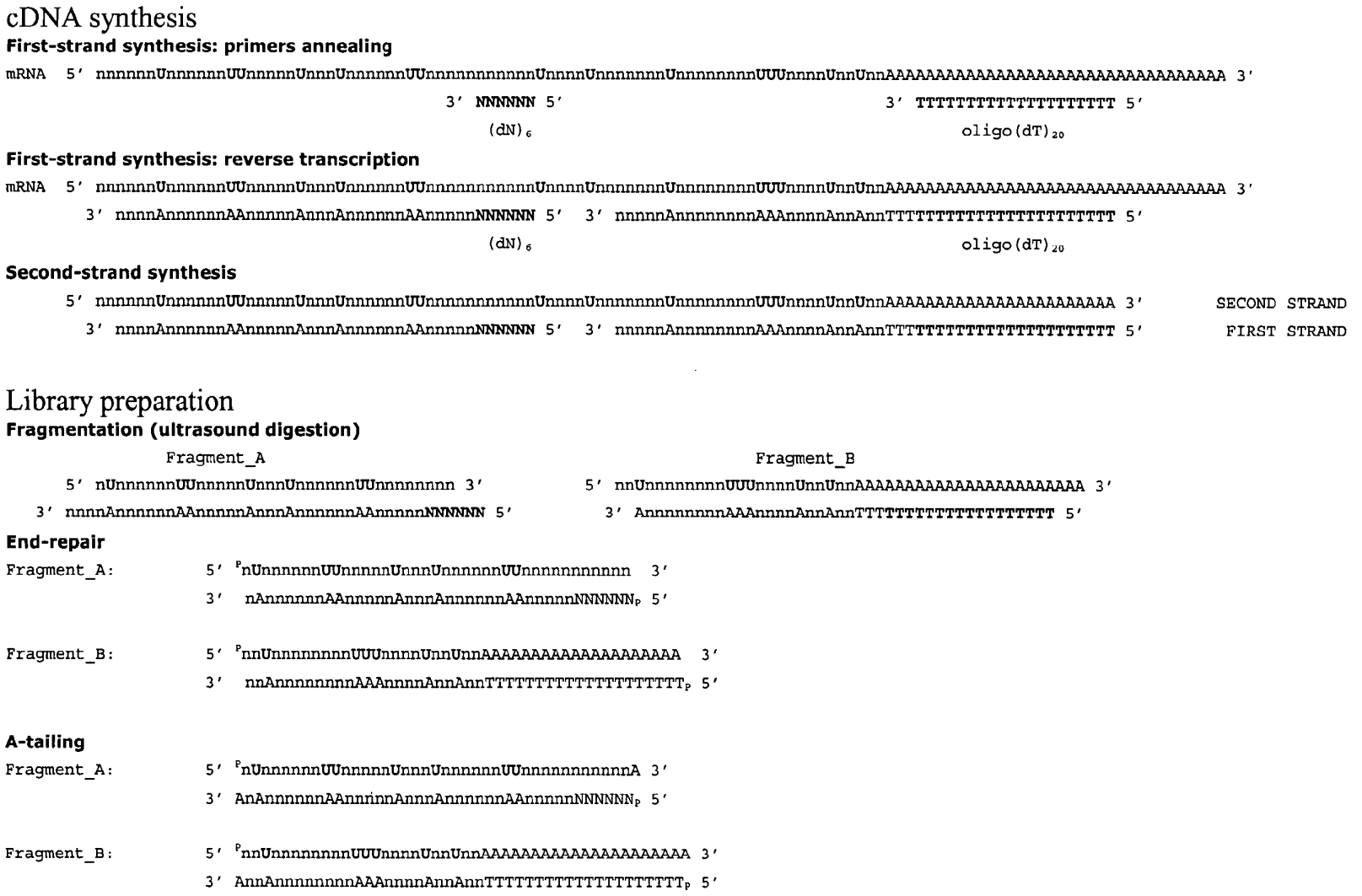


Figure 4

IDT Exhibit 1020.9479

Figure 5. Schematic view of the inventive method on the level of base pairs



EP 2 272 976 A1

IDT Exhibit 1020.9480

Adaptor ligation

```
Fragment_A:
                                                                #Ad_2
5' TACACTCTTTCCCTACACGAC                                      TCGTATGCCGTCTTCTGCTTG 3'
                  GCTCTTCCGATCT nUnnnnnnUUnnnnnnUnnnUnnnnnnUUnnnnnnUnnnnA GATCGGAAGAGC
                  CGAGAAGGCTAG AnAnnnnnnAAnnnnnnAnnnAnnnnnnAAnnnnnnNNNNNN TCTAGCCTTCTCG
3' GTTCGTCTTCTGCCGTATGCT                                      CAGCACATCCCTTTCTCACAT 5'
                        #Ad_1                                                 #Ad_1
Fragment_B:
                        #Ad_1                                                 #Ad_2
5' TACACTCTTTCCCTACACGAC                                      TCGTATGCCGTCTTCTGCTTG 3'
                  GCTCTTCCGATCT nnUnnnnnnnnnUUUnnnnnUnnUnnAAAAAAAAAAAAAAAAAA GATCGGAAGAGC
                  CGAGAAGGCTAG AnnAnnnnnnnnnAAAnnnnnAnnAnnTTTTTTTTTTTTTTTTT TCTAGCCTTCTCG
3' GTTCGTCTTCTGCCGTATGCT                                      CAGCACATCCCTTTCTCACAT 5'
                        #Ad_2                                                 #Ad_1
```

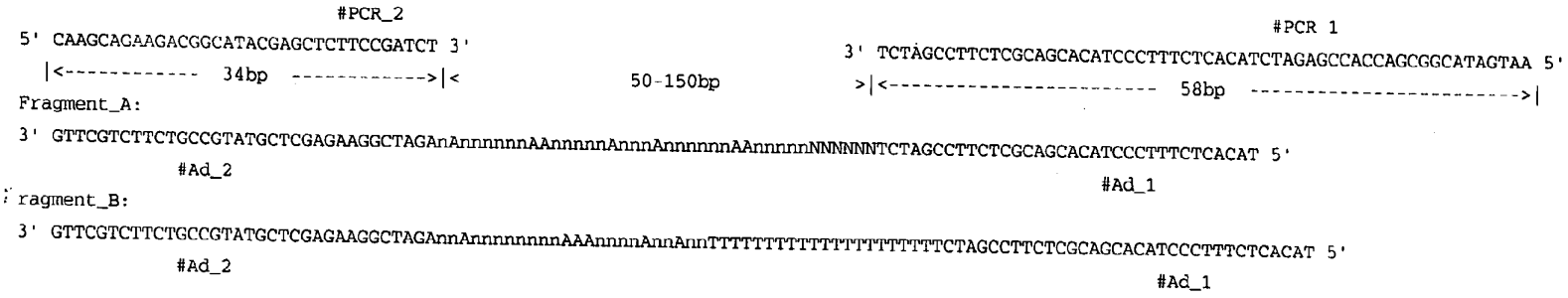
UDGase treatment

```
Fragment_A:
                        #Ad_1                                                 #Ad_2
5' TACACTCTTTCCCTACACGAC                                      TCGTATGCCGTCTTCTGCTTG 3'
                  GCTCTTCCGATCT n·nnnnnn··nnnnn·nnn·nnnnnn··nnnnnnnnnnnnA GATCGGAAGAGC
                  CGAGAAGGCTAG AnAnnnnnnAAnnnnnnAnnnAnnnnnnAAnnnnnnNNNNNN TCTAGCCTTCTCG
3' GTTCGTCTTCTGCCGTATGCT                                      CAGCACATCCCTTTCTCACAT 5'
                        #Ad_2                                                 #Ad_1
Fragment_B:
                        #Ad_1                                                 #Ad_2
5' TACACTCTTTCCCTACACGAC                                      TCGTATGCCGTCTTCTGCTTG 3'
                  GCTCTTCCGATCT nn·nnnnnnnn···nnnn·nn·nnAAAAAAAAAAAAAAAAAA GATCGGAAGAGC
                  CGAGAAGGCTAG AnnAnnnnnnnnnAAAnnnnnAnnAnnTTTTTTTTTTTTTTTTT TCTAGCCTTCTCG
3' GTTCGTCTTCTGCCGTATGCT                                      CAGCACATCCCTTTCTCACAT 5'
                        #Ad_2                                                 #Ad_1
```


IDT Exhibit 1020.9481

Preamplification

Fragment_A:



FlowCell preparation

Amplification

Fragment_A:

5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnTnnnnnnTTnnnnnTnnnTnnnnnnTTnnnnnnnnnnnnAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT 3'

3' GTTCGTCTTCTGCCGTATGCTCGAGAAGGCTAGAnAnnnnnnAAnnnnnAnnnAnnnnnnAAnnnnnNNNNNNTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA 5'

Fragment_B:

5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnnTnnnnnnnnnTTTnnnnTnnTnnAAAAAAAAAAAAAAAAAAAAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT 3'

3' GTTCGTCTTCTGCCGTATGCTCGAGAAGGCTAGAnnAnnnnnnnnnAAAnnnnAnnAnnTTTTTTTTTTTTTTTTTTTTTTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA 5'

Linearisation

Fragment_A:

5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnTnnnnnnTTnnnnnTnnnTnnnnnnTTnnnnnnnnnnnnAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT 3'

Fragment_B:

5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnnTnnnnnnnnnTTTnnnnTnnTnnAAAAAAAAAAAAAAAAAAAAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT 3'

EP 2 272 976 A1

IDT Exhibit 1020.9482

Sequencing primer annealing

Fragment_A:

```
                                    sequencing primer
                                <---- 3' TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA 5'
                                      |||||||||||||||||||||||||||||||||
5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnnTnnnnnnnnTTnnnnnnnTnnnTnnnnnnnnTTnnnnnnnnnnnnnnnnnnnnAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT 3'
Fragment_B:
                                                                <---- 3' TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA 5'
                                                                      |||||||||||||||||||||||||||||||||
5' CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTnnTnnnnnnnnnTTTnnnnnnnTnnTnnnATAAAAAAAAAAAAAAAAAAAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT
```


IDT Exhibit 1020.9483

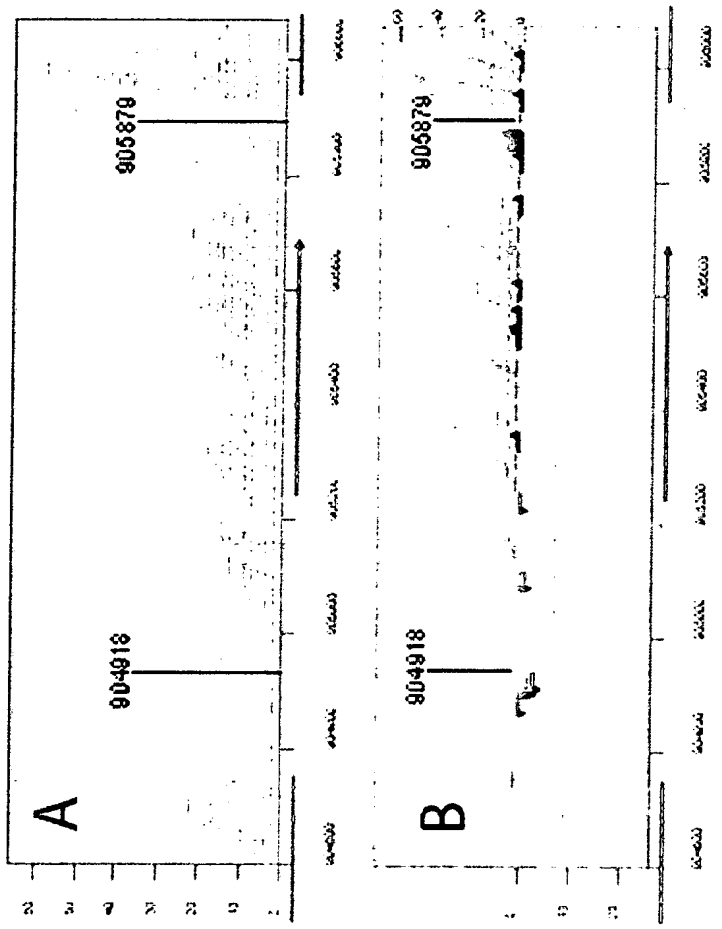


Figure 6

IDT Exhibit 1020.9484

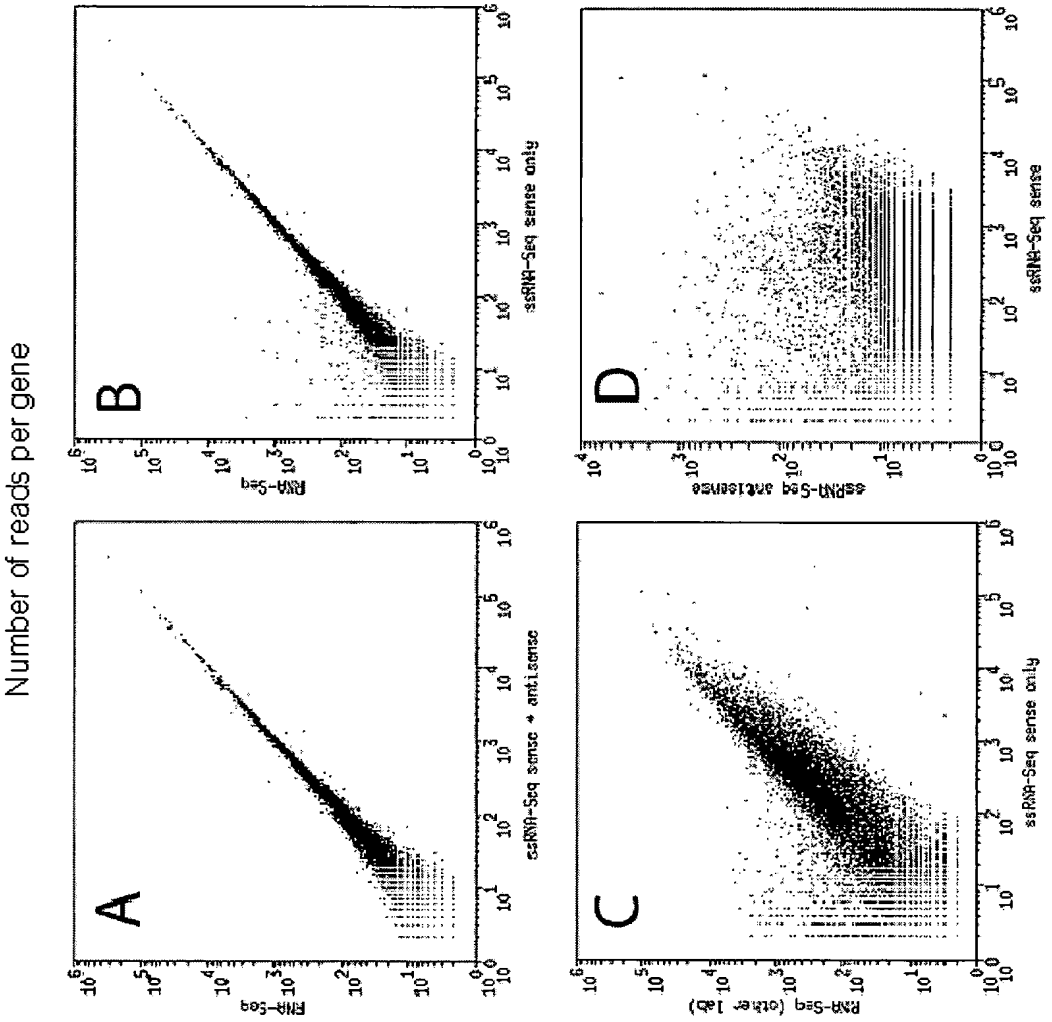
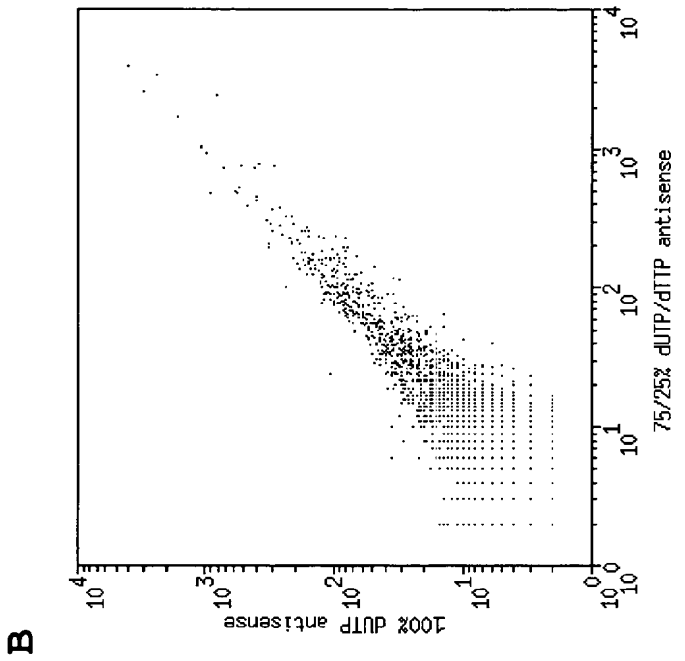


Figure 7

IDT Exhibit 1020.9485



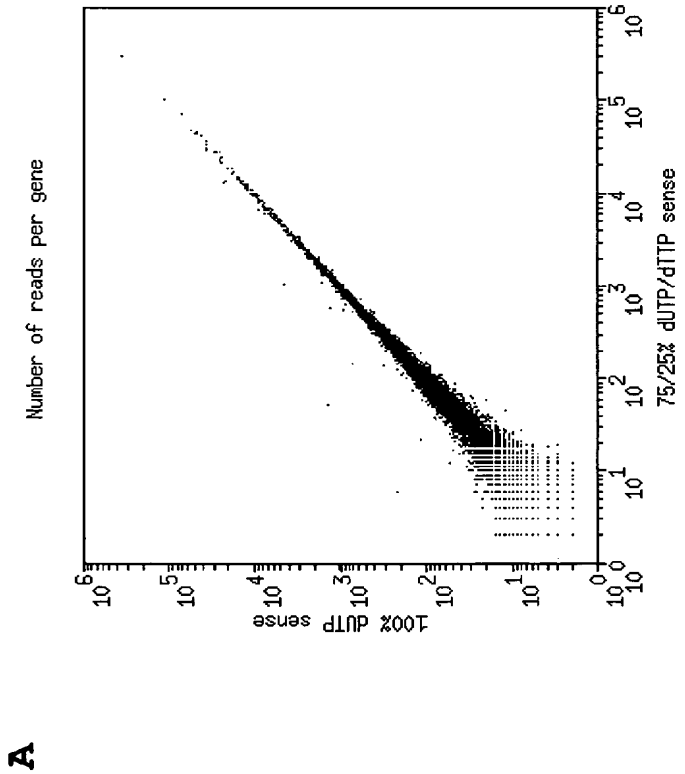


Figure 8

IDT Exhibit 1020.9486

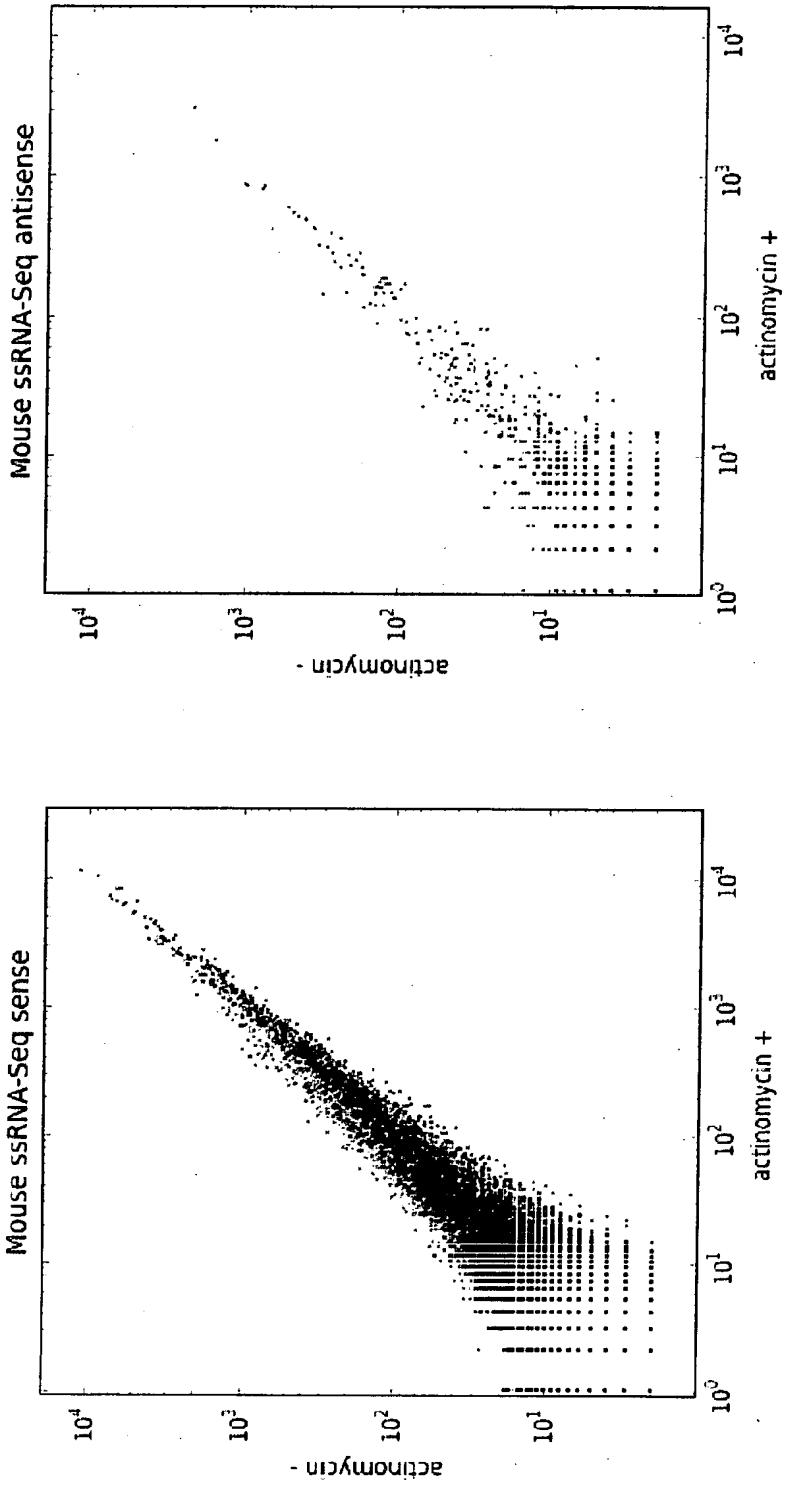


Figure 9

IDT Exhibit 1020.9487

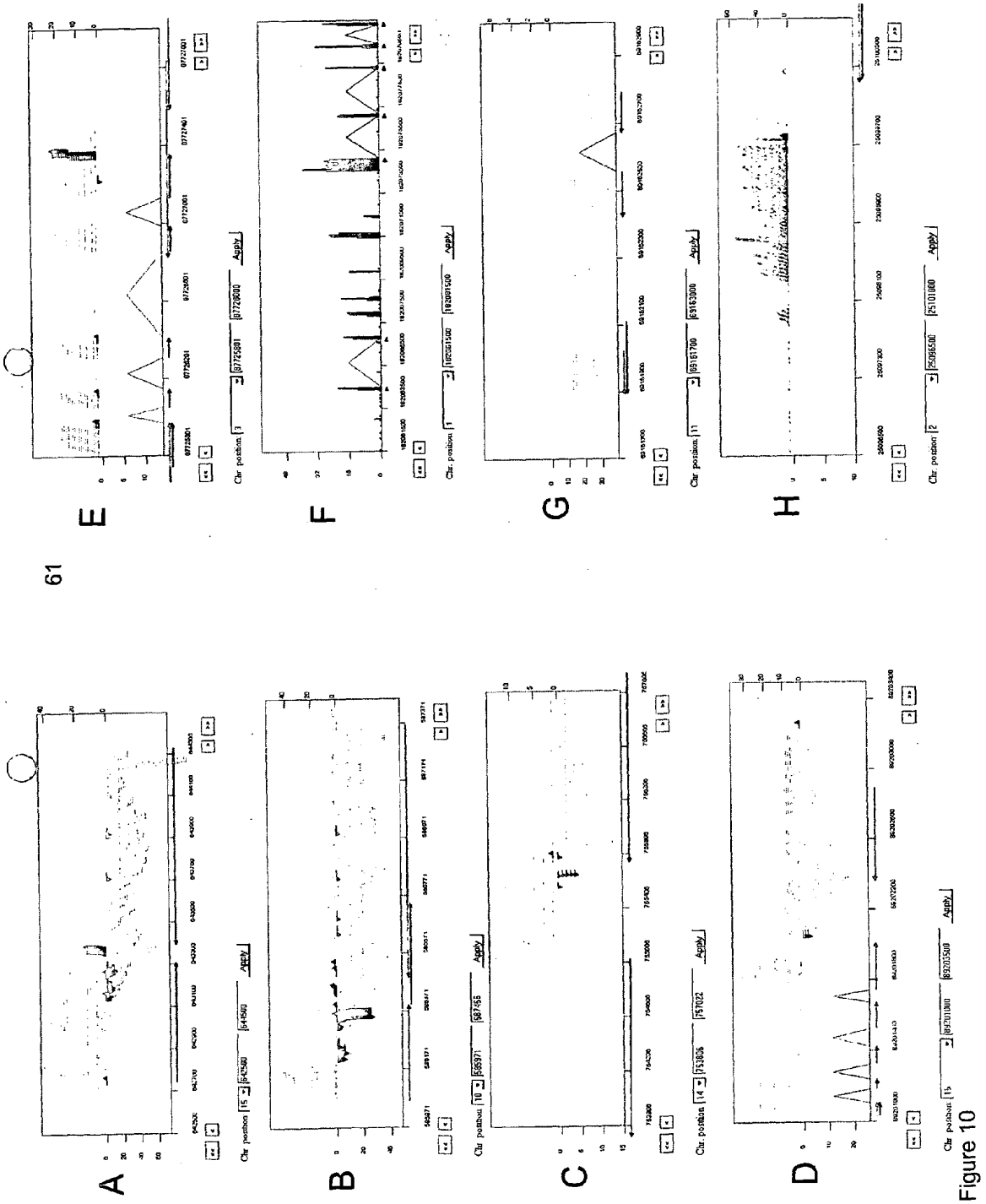


Figure 10

IDT Exhibit 1020.9488

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 09 00 8808

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2007/281863 A1 (RABBANI ELAZAR [US] ET AL) 6 December 2007 (2007-12-06) * abstract; examples 1,6 * * page 3, column 1, paragraph 4 - column 2, paragraph 1 * * page 6, column 2, paragraph 3; figure 2 *	1-14	INV. C12Q1/68
X	WO 97/12061 A1 (EPICENTRE TECHNOLOGIES CORP [US]) 3 April 1997 (1997-04-03) * claim 25 *	15	
A	WO 03/008538 A2 (MOLECULAR STAGING INC [US]) 30 January 2003 (2003-01-30) * abstract; claims 26-29; figure 5; example 9 *	6,12	
			TECHNICAL FIELDS SEARCHED (IPC) C12Q

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	22 December 2009	Tilkorn, A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

IDT Exhibit 1020.9489

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 09 00 8808

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

22-12-2009

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007281863 A1	06-12-2007	NONE	
WO 9712061 A1	03-04-1997	AU 704625 B2	29-04-1999
		AU 7118396 A	17-04-1997
		CA 2233079 A1	03-04-1997
		EP 0854936 A1	29-07-1998
		JP 11512607 T	02-11-1999
		JP 3561523 B2	02-09-2004
		US 6190865 B1	20-02-2001
WO 03008538 A2	30-01-2003	AU 2002259196 A1	03-03-2003
		US 2004091857 A1	13-05-2004

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

IDT Exhibit 1020.9490

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **J. Sambrook ; D. Russell.** Molecular Cloning: A Laboratory Manual **[0041]**
- **Hogrefe, H.H. ; Hansen, C.J. ; Scott, B.R. ; Nielsen K.B.** Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation. *Proc Natl Acad Sci USA,* 2002, vol. 99, 596-601 **[0065]**

IDT Exhibit 1020.9491

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68 | A1

(11) International Publication Number: **WO 00/18957**

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/GB99/03248

(22) International Filing Date: 30 September 1999 (30.09.99)

(30) Priority Data:
98307985.6 30 September 1998 (30.09.98) EP

(71) Applicant *(for all designated States except US)*: APPLIED RESEARCH SYSTEMS ARS HOLDING N.V. [NL/NL]; 14 John B. Gorsiraweg, Curacao (AN).

(71) Applicant *(for GB only)*: DZIEGLEWSKA, Hanna [GB/GB]; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ADESSI, Céline [FR/FR]; 8, rue du Jura, F–74100 Ambilly (FR). KAWASHIMA, Eric [US/CH]; 7, route de Lausanne, CH–1260 Nyon (CH). MAYER, Pascal [FR/FR]; Chemin du Ney, F–01200 Eloise (FR). MERMOD, Jean–Jacques [CH/CH]; 2, rue Maurice Barraud, CH–1206 Geneva (CH). TURCATTI, Gerardo [IT/CH]; 17, rue Dr. Alfred Vincent, CH–1201 Geneva (CH).

(74) Agents: DZIEGLEWSKA, Hanna et al.; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).

(81) Designated States: AE, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS OF NUCLEIC ACID AMPLIFICATION AND SEQUENCING

(57) Abstract

Methods for amplification and sequencing of at least one nucleic acid comprising the following steps: (1) forming at least one nucleic acid template comprising the nucleic acid(s) to be amplified or sequenced, wherein said nucleic acid(s) contains at the 5' end an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support; (2) mixing said nucleic acid template(s) with one or more colony primers X, which can hybridize to the oligonucleotide sequence Z and carries at the 5' end a means for attaching the colony primers to a solid support, in the presence of a solid support so that the 5' ends of both the nucleic acid template and the colony primers bind to the solid support; (3) performing one or more nucleic acid amplification reactions on the bound template(s), so that nucleic acid colonies are generated and optionally, performing at least one step of sequence determination of one or more of the nucleic acid colonies generated. Solid supports, kits and apparatus for use in these methods.

IDT Exhibit 1020.9492

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.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDT Exhibit 1020.9493

Methods of nucleic acid amplification and sequencing

This invention relates to the field of nucleic acid amplification and sequencing. More specifically, this invention relates to nucleic acid amplification and sequencing methods, and apparatus and kits useful for large scale high throughput amplification and sequencing of nucleic acids.

Nucleic acid sequence analysis has become a cornerstone in many activities in biology, biotechnology and medicine. The ability to determine nucleic acid sequences has become increasingly important as efforts have commenced to determine the sequences of the large genomes of humans and other higher organisms and also, for example, in single nucleotide polymorphism detection and screening and gene expression monitoring. The genetic information provided by nucleic acid sequencing has many applications in areas such as for example drug target discovery and validation, disease diagnosis and risk scoring and organism identification and characterization.

The first step in such applications is the determination of the actual chemical composition of the nucleic acids of interest, more precisely the determination of the sequence of occurrence of the four bases adenine (A), cytosine (C), guanine (G) and thymine (T) or uracil (U) which comprise nucleic acids. However, such applications require the sequencing of nucleic acids on a large scale, making high throughput methods of nucleic acid sequencing extremely desirable.

Methods of nucleic acid sequencing are documented in the art. The two most commonly used are the chemical cleavage technique by Maxam and Gilbert which relies on base-specific chemistry and the now more popular Sanger sequencing technique which relies on an enzymatic chain

IDT Exhibit 1020.9494

terminating principle and is now used on a routine basis for nucleic acid sequencing.

In Sanger sequencing, each nucleic acid to be sequenced is replicated in a reaction involving DNA polymerase, deoxynucleotide triphosphates (dNTPs) and dideoxynucleotide triphosphates (ddNTPs). The DNA polymerase can incorporate both dNTPs and ddNTPs into the growing DNA strand. However, once a ddNTP is incorporated, the 3' end of the growing DNA strand lacks a hydroxyl group and is no longer a substrate for chain elongation, thus terminating the nucleic acid chain. Hence, in a particular reaction including one type of ddNTP a mixture of nucleic acids of different lengths is produced, all terminating with the same ddNTP. Usually separate reactions are set up for each of the four types of ddNTP and the distribution of lengths of the nucleic acid fragments produced is analysed by denaturing gel electrophoresis (which resolves nucleic acid fragments according to their size), or more recently, by mass-spectroscopy. Usually, one or more of the deoxynucleotide triphosphates in the reaction mixture is labelled to enable detection of the fragments of different lengths.

The above described methods are disadvantageous because each nucleic acid to be sequenced has to be processed individually during the biochemical reaction. Gel electrophoresis is cumbersome, labour intensive and intrinsically slow even when capillary electrophoresis is used and is not well suited for large scale high throughput sequencing. In addition, the subsequent determination of the sequence is cumbersome. Mass-spectroscopy is still at the prototype level, requires very expensive apparatus and each sample has to be analysed individually.

One way to increase throughput is to process many samples in parallel. Methods using DNA hybridization of nucleic acid probes are in use and allow for some

multiplexing of the process during the biochemical and the electrophoretic processes, but at the cost of lengthy additional manipulations.

More recently methods based on DNA chips and DNA hybridization are becoming available (Thomas and Burke Exp. Opin. Ther. Patents 8: 503-508 (1998)). These methods are disadvantageous because for each application, a DNA chip has to be designed and manufactured first: this is a lengthy operation and the price of an individual chip drops only when very large numbers of the chip are required. Also, the chips are not reusable and for each chip only one sample of nucleic acids, e.g. one patient to be diagnosed, can be processed at each time. Finally, the extent of sequence which can be analysed by such a chip is limited to less than 100,000 bases, and is limited to some applications such as DNA genotyping and gene expression profiling.

In most known techniques for nucleic acid sequence analysis, amplification of the nucleic acids of interest is a prerequisite step in order to obtain the nucleic acid in a quantity sufficient for analysis.

Several methods of nucleic acid amplification are well known and documented in the art. For example, nucleic acids can be amplified by inserting the nucleic acid of interest into an expression vector construct. Such vectors can then be introduced into suitable biological host cells and the vector DNA, including the nucleic acid of interest, amplified by culturing the biological host using well established protocols.

Nucleic acids amplified by such methods can be isolated from the host cells by methods well known and documented in the art. However, such methods have the disadvantage of being generally time consuming, labour intensive and difficult to automate.

The technique of DNA amplification by the polymerase chain reaction (PCR) was disclosed in 1985 (Saiki et al. Science 230, 1350-1354) and is now a

method well known and documented in the art. A target nucleic acid fragment of interest can be amplified using two short oligonucleotide sequences (usually referred to as primers) which are specific to known sequences flanking the DNA sequence that is to be amplified. The primers hybridize to opposite strands of the double-stranded DNA fragment after it has been denatured, and are oriented so that DNA synthesis by the DNA polymerase proceeds through the region between the two primers, with the primer sequences being extended by the sequential incorporation of nucleotides by the polymerase. The extension reactions create two double-stranded target regions, each of which can again be denatured ready for a second cycle of hybridisation and extension. The third cycle produces two double-stranded molecules that comprise precisely the target region in double-stranded form. By repeated cycles of heat denaturation, primer hybridisation, and extension, there follows a rapid exponential accumulation of the specific target fragment of DNA. Traditionally, this method is performed in solution and the amplified target nucleic acid fragment purified from solution by methods well known in the art, for example by gel electrophoresis.

More recently, however, methods have been disclosed which use one primer grafted to a surface in conjunction with free primers in solution. These methods allow the simultaneous amplification and attachment of a PCR product onto the surface (Oroskar, A.A. et al., *Clinical Chemistry* 42:1547 (1996)).

WO96/04404 (Mosaic Technologies, Inc. et al) discloses a method of detection of a target nucleic acid in a sample which potentially contains the target nucleic acid. The method involves the induction of a PCR based amplification of the target nucleic acid only when the target nucleic acid is present in the sample being tested. For the amplification of the target sequence, both primers are attached to a solid support,

which results in the amplified target nucleic acid sequences also being attached to the solid support. The amplification technique disclosed in this document is sometimes referred to as the "bridge amplification" technique. In this technique the two primers are, as for conventional PCR, specifically designed so that they flank the particular target DNA sequence to be amplified. Thus, if the particular target nucleic acid is present in the sample, the target nucleic acid can hybridise to the primers and be amplified by PCR. The first step in this PCR amplification process is the hybridisation of the target nucleic acid to the first specific primer attached to the support ("primer 1"). A first amplification product, which is complementary to the target nucleic acid, is then formed by extension of the primer 1 sequence. On subjecting the support to denaturation conditions the target nucleic acid is released and can then participate in further hybridisation reactions with other primer 1 sequences which may be attached to the support. The first amplification product which is attached to the support, may then hybridise with the second specific primer ("primer 2") attached to the support and a second amplification product comprising a nucleic acid sequence complementary to the first amplification product can be formed by extension of the primer 2 sequence and is also attached to the support. Thus, the target nucleic acid and the first and second amplification products are capable of participating in a plurality of hybridisation and extension processes, limited only by the initial presence of the target nucleic acid and the number of primer 1 and primer 2 sequences initially present and the result is a number of copies of the target sequence attached to the surface.

Since, on carrying out this process, amplification products are only formed if the target nucleic acid is present, monitoring the support for the presence or

absence of one or more amplification products is indicative of the presence or absence of a specific target sequence.

The Mosaic technique can be used to achieve an amount of multiplexing in that several different target nucleic acid sequences can be amplified simultaneously by arraying different sets of first and second primers, specific for each different target nucleic acid sequence, on different regions of the solid support.

The disadvantage of the Mosaic process is that, as the first and second primer sequences have to be specific for each target nucleic acid to be amplified, it can only be used to amplify known sequences. In addition, the throughput is limited by the number of different sets of specific primers and subsequently amplified target nucleic acid molecules which can be arrayed in distinct regions of a given solid support and the time taken to array the nucleic acids in distinct regions. Also, the Mosaic process requires that 2 different primers are homogeneously attached by the 5' end to the support within the distinct region where the amplification product is formed. This cannot be achieved with presently available DNA chip manufacturing technology and has to be achieved by some means of sample dispensing. Thus, the density that can be achieved by this approach has the same limitation as other classical arraying technologies. A further limitation is the speed of monitoring the individual distinct regions of the support for the presence or absence of the amplified target nucleic acids.

Arraying of DNA samples is classically performed on membranes (e.g., nylon or nitro-cellulose membranes). The use of suitable robotics (e.g., Q-bot™, Genetix Ltd, Dorset BH23 3TG UK) means that it is possible to obtain a density of up to 10 samples/mm^2. In such methods, the DNA is covalently linked to a membrane by physicochemical means (e.g., UV irradiation) and the

arraying of large DNA molecules (e.g. molecules over 100 nucleotides long) as well as smaller DNA molecules such as oligonucleotide primers is possible.

Other techniques are known whereby higher density arrays of oligonucleotides can be obtained. For example, approaches based on pre-arrayed glass slides wherein arrays of reactive areas are obtained by ink-jet technology (Blanchard, A.P.and L. Hood, *Microbial and Comparative Genomics*, 1:225 (1996)) or arrays of reactive polyacrylamide gels (Yershov, G. et al., *Proceedings of the National Academy of Science, USA*, 93:4913-4918 (1996)) allow in theory the arraying of up to 100 samples/mm^2.

Higher sample densities still are achievable by the use of DNA chips (Fodor, S.P.A. et al., *Science* 251:767(1991)). Currently, chips with 625 oligonucleotide probes/mm^2 are used in molecular biology techniques (Lockhart, D.J. et al., *Nature Biotechnology* 14:1675 (1996)). Probe densities of up to 250 000 samples/cm^2 (2500/mm^2) are claimed to be achievable (Chee, M. et al., *Science* 274:610 (1996)). However, at present up to 132000 different oligonucleotides can be arrayed on a single chips of approximately 2.5 cm^2. Importantly, these chips are manufactured in such a way so that the 3'OH end of the oligonucleotide is attached to the solid surface. This means that oligonucleotides attached to chips in such a way cannot be used as primers in a PCR amplification reaction.

Importantly, when PCR products are linked to the vessel in which PCR amplification takes place, the density of the resultant array of PCR products is limited by the available vessel. Currently available vessels are only in 96 well microtiter plate format. These allow only around 0.02 samples of PCR products/mm^2 of surface to be obtained.

For example, using the commercially available Nucleolink™ system (Nunc A/S, Roskilde, Denmark) it is

possible to achieve simultaneous amplification and arraying of samples at a density of 0.02 samples/mm^2 in wells on the surface of which oligonucleotide primers have been grafted. However, technical problems mean that it is unlikely that a significant increase in this sample density will be achieved with this approach.

Thus, it can be seen that in order to increase throughput there is a need in the art for new methods of nucleic acid amplification which allow the simultaneous amplification and array of nucleic acid samples at a higher density, and furthermore, allows the monitoring of samples at a faster rate, preferably in parallel.

In addition, it is apparent that there is a need in the art for new methods of sequencing which allow large numbers of samples to be processed and sequenced in parallel, i.e. there is a need for methods of sequencing which allow significant multiplexing of the process. Significant multiplexing of the sequencing process would in turn lead to a higher throughput than that achievable with the methods of sequencing known in the art. Such new methods would be even more desirable if they could achieve such high throughput sequencing at a reasonable cost and with less labour intensiveness than conventional sequencing techniques.

The present invention describes new methods of solid-phase nucleic acid amplification which enable a large number of distinct nucleic acid sequences to be arrayed and amplified simultaneously and at a high density. The invention also describes methods by which a large number of distinct amplified nucleic acid sequences can be monitored at a fast rate and, if desired, in parallel. The invention also describes methods by which the sequences of a large number of distinct nucleic acids can be determined simultaneously and within a short period of time. The methods are particularly useful in, but not limited to, the sequencing of a whole genome, or situations where many

genes (e.g. 500) from many individuals (e.g. 500) have to be sequenced simultaneously, or the simultaneous scoring of large numbers (e.g. millions) of polymorphisms, or the monitoring of the expression of a large number of genes (e.g. 100,000) simultaneously.

The present invention therefore provides a method for amplification of at least one nucleic acid comprising the following steps:-

(1) forming at least one nucleic acid template comprising the nucleic acid(s) to be amplified, wherein said nucleic acid(s) contains at the 5' end an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support;

(2) mixing said nucleic acid template(s) with one or more colony primers X, which can hybridize to the oligonucleotide sequence Z and carries at the 5' end a means for attaching the colony primers to a solid support, in the presence of a solid support so that the 5' ends of both the nucleic acid template and the colony primers bind to the solid support;

(3) performing one or more nucleic acid amplification reactions on the bound template(s), so that nucleic acid colonies are generated.

In a further embodiment of the invention, two different colony primers X are mixed with the nucleic acid template(s) in step (2) of the method. Preferably the sequences of colony primers X are such that the oligonucleotide sequence Z can hybridise to one of the colony primers X and the oligonucleotide sequence Y is the same as one of the colony primers X.

In an alternative embodiment of the invention, the oligonucleotide sequence Z is complementary to oligonucleotide sequence Y, referred to as Y' and colony primer X is of the same sequence as oligonucleotide sequence Y.

In a yet further embodiment of the invention, the colony primer X may comprise a degenerate primer sequence and the nucleic acid template(s) comprise the nucleic acid(s) to be amplified and do not contain oligonucleotide sequences Y or Z at the 5' and 3' ends respectively.

In a further aspect of the invention, the method comprises the additional step of performing at least one step of sequence determination of one or more of the nucleic acid colonies generated in step (3).

Thus the invention also provides a method for sequencing of at least one nucleic acid comprising the following steps:-

(1) forming at least one nucleic acid template comprising the nucleic acid(s) to be sequenced, wherein said nucleic acid(s) contains at the 5' end an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support;

(2) mixing said nucleic acid template(s) with one or more colony primers X, which can hybridize to the oligonucleotide sequence Z and carries at the 5' end a means for attaching the colony primers to a solid support, in the presence of a solid support so that the 5' ends of both the nucleic acid template and the colony primers bind to the solid support;

(3) performing one or more nucleic acid amplification reactions on the bound template(s), so that nucleic acid colonies are generated; and

(4) performing at least one step of sequence determination of at least one of the nucleic acid colonies generated.

In a further embodiment of the invention the 5' ends of both the nucleic acid template(s) and the colony primers carry a means for attaching the nucleic acid sequences covalently to the solid support. Preferably

this means for covalent attachment is a chemically modifiable functional group, such as for example, a phosphate group, a carboxylic or aldehyde moiety, a thiol, a hydroxyl, a dimethoxyltrityl (DMT), or an amino group, preferably an amino group.

Nucleic acids which may be amplified according to the methods of the invention include DNA, for example, genomic DNA, cDNA, recombinant DNA or any form of synthetic or modified DNA, RNA, mRNA or any form of synthetic or modified RNA. Said nucleic acids may vary in length and may be fragments or smaller parts of larger nucleic acid molecules. Preferably the nucleic acid to be amplified is at least 50 base pairs in length and more preferably 150 to 4000 base pairs in length. The nucleic acid to be amplified may have a known or unknown sequence and may be in a single or double-stranded form. The nucleic acid to be amplified may be derived from any source.

"Nucleic acid template" as used herein refers to an entity which comprises the nucleic acid to be amplified or sequenced in a single-stranded form. As outlined below the nucleic acid to be amplified or sequenced can also be provided in a double stranded form. Thus, "nucleic acid templates" of the invention may be single or double stranded nucleic acids. The nucleic acid templates to be used in the method of the invention can be of variable lengths. Preferably they are at least 50 base pairs in length and more preferably 150 to 4000 base pairs in length. The nucleotides making up the nucleic acid templates may be naturally occurring or non-naturally occurring nucleotides. The nucleic acid templates of the invention not only comprise the nucleic acid to be amplified but may in addition contain at the 5' and 3' end short oligonucleotide sequences. The oligonucleotide sequence contained at the 5' end is referred to herein as Y. Oligonucleotide sequence Y is of a known sequence and can be of variable length.

Oligonucleotide sequence Y for use in the methods of the present invention is preferably at least five nucleotides in length, preferably between 5 and 100 nucleotides in length and more preferably of approximately 20 nucleotides in length. Naturally occurring or non-naturally occurring nucleotides may be present in the oligonucleotide sequence Y. As indicated above, preferably the sequence of oligonucleotide Y is the same as the sequence of colony primer X. The oligonucleotide sequence contained at the 3' end of the nucleic acid templates of the invention is referred to herein as Z. Oligonucleotide sequence Z is of a known sequence and can be of variable length. Oligonucleotide sequence Z for use in the methods of the present invention is preferably at least five nucleotides in length, preferably between 5 and 100 nucleotides in length and more preferably of approximately 20 nucleotides in length. Naturally occurring or non-naturally occurring nucleotides may be present in the oligonucleotide sequence Z. Oligonucleotide sequence Z is designed so that it hybridises with one of the colony primers X and preferably is designed so that it is complementary to oligonucleotide sequence Y, referred to herein as Y'. The oligonucleotide sequences Y and Z contained at the 5' and 3' ends respectively of a nucleic acid template need not be located at the extreme ends of the template. For example although the oligonucleotide sequences Y and Z are preferably located at or near the 5' and 3' ends (or termini) respectively of the nucleic acid templates (for example within 0 to 100 nucleotides of the 5' and 3' termini) they may be located further away (e.g. greater than 100 nucleotides) from the 5' or 3' termini of the nucleic acid template. The oligonucleotide sequences Y and Z may therefore be located at any position within the nucleic acid template providing the sequences Y and Z are on either side, i.e. flank, the nucleic acid sequence which is to be

amplified.

"Nucleic acid template" as used herein also includes an entity which comprises the nucleic acid to be amplified or sequenced in a double-stranded form. When the nucleic acid template is in a double-stranded form, the oligonucleotide sequences Y and Z are contained at the 5' and 3' ends respectively of one of the strands. The other strand, due to the base pairing rules of DNA, is complementary to the strand containing oligonucleotide sequences Y and Z and thus contains an oligonucleotide sequence Z' at the 5' end and an oligonucleotide sequence Y' at the 3' end.

"Colony primer" as used herein refers to an entity which comprises an oligonucleotide sequence which is capable of hybridizing to a complementary sequence and initiate a specific polymerase reaction. The sequence comprising the colony primer is chosen such that it has maximal hybridising activity with its complementary sequence and very low non-specific hybridising activity to any other sequence. The sequence to be used as a colony primer can include any sequence, but preferably includes 5'-AGAAGGAGAAGGAAAGGGAAAGGG or 5'-CACCAACCCAAACCAACCCAAACC. The colony primer can be 5 to 100 bases in length, but preferably 15 to 25 bases in length. Naturally occurring or non-naturally occurring nucleotides may be present in the primer. One or two different colony primers may be used to generate nucleic acid colonies in the methods of the present invention. The colony primers for use in the present invention may also include degenerate primer sequences.

"Degenerate primer sequences" as used herein refers to a short oligonucleotide sequence which is capable of hybridizing to any nucleic acid fragment independent of the sequence of said nucleic acid fragment. Such degenerate primers thus do not require the presence of oligonucleotide sequences Y or Z in the nucleic acid template(s) for hybridization to the template to occur,

although the use of degenerate primers to hybridise to a template containing the oligonucleotide sequences X or Y is not excluded. Clearly however, for use in the amplification methods of the present invention, the degenerate primers must hybridise to nucleic acid sequences in the template at sites either side, or flanking, the nucleic acid sequence which is to be amplified.

"Solid support" as used herein refers to any solid surface to which nucleic acids can be covalently attached, such as for example latex beads, dextran beads, polystyrene, polypropylene surface, polyacrylamide gel, gold surfaces, glass surfaces and silicon wafers. Preferably the solid support is a glass surface.

"Means for attaching nucleic acids to a solid support" as used herein refers to any chemical or non-chemical attachment method including chemically-modifiable functional groups. "Attachment" relates to immobilization of nucleic acid on solid supports by either a covalent attachment or via irreversible passive adsorption or via affinity between molecules (for example, immobilization on an avidin-coated surface by biotinylated molecules). The attachment must be of sufficient strength that it cannot be removed by washing with water or aqueous buffer under DNA-denaturing conditions.

"Chemically-modifiable functional group" as used herein refers to a group such as for example, a phosphate group, a carboxylic or aldehyde moiety, a thiol, or an amino group.

"Nucleic acid colony" as used herein refers to a discrete area comprising multiple copies of a nucleic acid strand. Multiple copies of the complementary strand to the nucleic acid strand may also be present in the same colony. The multiple copies of the nucleic acid strands making up the colonies are generally

immobilised on a solid support and may be in a single or double stranded form. The nucleic acid colonies of the invention can be generated in different sizes and densities depending on the conditions used. The size of colonies is preferably from 0.2 μm to 6 μm, more preferably from 0.3 μm to 4 μm. The density of nucleic acid colonies for use in the method of the invention is typically 10,000/mm^2 to 100,000/mm^2. It is believed that higher densities, for example, 100,000/mm^2 to 1,000,000/mm^2 and 1,000,000/mm^2 to 10,000,000/mm^2 may be achieved.

The methods of the invention can be used to generate nucleic acid colonies. Thus, a further aspect of the invention provides one or more nucleic acid colonies. A nucleic acid colony of the invention may be generated from a single immobilised nucleic acid template of the invention. The method of the invention allows the simultaneous production of a number of such nucleic acid colonies, each of which may contain different immobilised nucleic acid strands.

Thus, a yet further aspect of the invention provides a plurality of nucleic acid templates comprising the nucleic acids to be amplified, wherein said nucleic acids contain at their 5' ends an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support. Preferably this plurality of nucleic acid templates are mixed with a plurality of colony primers X which can hybridize to the oligonucleotide sequence Z and carry at the 5' end a means for attaching the colony primers to a solid support. Preferably said plurality of nucleic acid templates and colony primers are covalently bound to a solid support.

In a further embodiment of the invention, pluralities of two different colony primers X are mixed

with the plurality of nucleic acid templates. Preferably the sequences of colony primers X are such that the oligonucleotide sequence Z can hybridise to one of the colony primers X and the oligonucleotide sequence Y is the same as the sequence of one of the colony primers X.

In an alternative embodiment, the oligonucleotide sequence Z is complementary to oligonucleotide sequence Y,(Y') and the plurality of colony primers X are of the same sequence as oligonucleotide sequence Y.

In a yet further embodiment, the plurality of colony primers X may comprise a degenerate primer sequence and the plurality of nucleic acid templates comprise the nucleic acids to be amplified and do not contain oligonucleotide sequences Y or Z at the 5' and 3' ends respectively.

The nucleic acid templates of the invention may be prepared using techniques which are standard or conventional in the art. Generally these will be based on genetic engineering techniques.

The nucleic acids to be amplified can be obtained using methods well known and documented in the art. For example, by obtaining a nucleic acid sample such as, total DNA, genomic DNA, cDNA, total RNA, mRNA etc. by methods well known and documented in the art and generating fragments therefrom by, for example, limited restriction enzyme digestion or by mechanical means.

Typically, the nucleic acid to be amplified is first obtained in double stranded form. When the nucleic acid is provided in single stranded form, for example mRNA, it is first made into a double stranded form by means well known and documented in the art, for example, using oligo-dT primers and reverse transcriptase and DNA polymerase. Once the nucleic acid to be amplified is obtained in double stranded form of appropriate length, oligonucleotide sequences corresponding to the oligonucleotide sequences Y and Z

are joined to each end, i.e. to both the 5' and 3' ends of the nucleic acid sequence to form a nucleic acid template. This can be done using methods which are well known and documented in the art, for example by ligation, or by inserting the nucleic acid to be amplified into a biological vector at a site which is flanked by the appropriate oligonucleotide sequences. Alternatively, if at least part of the sequence of the nucleic acid to be amplified is known, the nucleic acid template containing oligonucleotide sequences Y and Z at the 5' and 3' ends respectively, may be generated by PCR using appropriate PCR primers which include sequences specific to the nucleic acid to be amplified. Before attaching the nucleic acid template to the solid support, it can be made into a single stranded form using methods which are well known and documented in the art, for example by heating to approximately 94°C and quickly cooling to 0°C on ice.

The oligonucleotide sequence contained at the 5' end of the nucleic acid can be of any sequence and any length and is denoted herein as sequence Y. Suitable lengths and sequences of oligonucleotide can be selected using methods well known and documented in the art. For example the oligonucleotide sequences attached to each end of the nucleic acid to be amplified are normally relatively short nucleotide sequences of between 5 and 100 nucleotides in length. The oligonucleotide sequence contained at the 3' end of the nucleic acid can be of any sequence and any length and is denoted herein as sequence Z. Suitable lengths and sequences of oligonucleotide can be selected using methods well known and documented in the art. For example the oligonucleotide sequences contained at each end of the nucleic acid to be amplified are normally relatively short nucleotide sequences of between 5 and 100 nucleotides in length.

The sequence of the oligonucleotide sequence Z is

such that it can hybridise to one of the colony primers X. Preferably, the sequence of the oligonucleotide sequence Y is such that it is the same as one of the colony primers X. More preferably, the oligonucleotide sequence Z is complementary to oligonucleotide sequence Y (Y') and the colony primers X are of the same sequence as oligonucleotide sequence Y.

The oligonucleotide sequences Y and Z of the invention may be prepared using techniques which are standard or conventional in the art, or may be purchased from commercial sources.

When producing the nucleic acid templates of the invention additional desirable sequences can be introduced by methods well known and documented in the art. Such additional sequences include, for example, restriction enzyme sites or certain nucleic acid tags to enable amplification products of a given nucleic acid template sequence to be identified. Other desirable sequences include fold-back DNA sequences (which form hairpin loops or other secondary structures when rendered single-stranded), 'control' DNA sequences which direct protein/DNA interactions, such as for example a promoter DNA sequence which is recognised by a nucleic acid polymerase or an operator DNA sequence which is recognised by a DNA-binding protein.

If there are a plurality of nucleic acid sequences to be amplified then the attachment of oligonucleotides Y and Z can be carried out in the same or different reaction.

Once a nucleic acid template has been prepared, it may be amplified before being used in the methods of the present invention. Such amplification may be carried out using methods well known and documented in the art, for example by inserting the template nucleic acid into an expression vector and amplifying it in a suitable biological host, or amplifying it by PCR. This amplification step is not however essential, as the

method of the invention allows multiple copies of the nucleic acid template to be produced in a nucleic acid colony generated from a single copy of the nucleic acid template.

Preferably the 5' end of the nucleic acid template prepared as described above is modified to carry a means for attaching the nucleic acid templates covalently to a solid support. Such a means can be, for example, a chemically modifiable functional group, such as, for example a phosphate group, a carboxylic or aldehyde moiety, a thiol, or an amino group. Most preferably, the thiol, phosphate or amino group is used for 5'-modification of the nucleic acid.

The colony primers of the invention may be prepared using techniques which are standard or conventional in the art. Generally, the colony primers of the invention will be synthetic oligonucleotides generated by methods well known and documented in the art or may be purchased from commercial sources.

According to the method of the invention one or two different colony primers X, can be used to amplify any nucleic acid sequence. This contrasts with and has an advantage over many of the amplification methods known in the art such as, for example, that disclosed in WO 96/04404, where different specific primers must be designed for each particular nucleic acid sequence to be amplified.

Preferably the 5' ends of colony primers X of the invention are modified to carry a means for attaching the colony primers covalently to the solid support. Preferably this means for covalent attachment is a chemically modifiable functional group as described above. If desired, the colony primers can be designed to include additional desired sequences such as, for example, restriction endonuclease sites or other types of cleavage sites each as ribozyme cleavage sites. Other desirable sequences include fold-back DNA

sequences (which form hairpin loops or other secondary structures when rendered single-stranded), 'control' DNA sequences which direct a protein/DNA interaction, such as for example a promoter DNA sequence which is recognised by a nucleic acid polymerase or an operator DNA sequence which is recognised by a DNA-binding protein.

Immobilisation of a colony primer X to a support by the 5' end leaves its 3' end remote from the support such that the colony primer is available for chain extension by a polymerase once hybridisation with a complementary oligonucleotide sequence contained at the 3' end of the nucleic acid template has taken place.

Once both the nucleic acid templates and colony primers of the invention have been synthesised they are mixed together in appropriate proportions so that when they are attached to the solid support an appropriate density of attached nucleic acid templates and colony primers is obtained. Preferably the proportion of colony primers in the mixture is higher than the proportion of nucleic acid templates. Preferably the ratio of colony primers to nucleic acid templates is such that when the colony primers and nucleic acid templates are immobilised to the solid support a "lawn" of colony primers is formed comprising a plurality of colony primers being located at an approximately uniform density over the whole or a defined area of the solid support, with one or more nucleic acid templates being immobilised individually at intervals within the lawn of colony primers.

The nucleic acid templates may be provided in single stranded form. However, they may also be provided totally or partly in double stranded form, either with one 5' end or both 5' ends modified so as to allow attachment to the support. In that case, after completion of the attachment process, one might want to separate strands by means known in the art, e.g. by

heating to 94°C, before washing the released strands away. It will be appreciated that in the case where both strands of the double stranded molecules have reacted with the surface and are both attached, the result will be the same as in the case when only one strand is attached and one amplification step has been performed. In other words, in the case where both strands of a double stranded template nucleic acid have been attached, both strands are necessarily attached close to each other and are indistinguishable from the result of attaching only one strand and performing one amplification step. Thus, single stranded and double stranded template nucleic acids might be used for providing template nucleic acids attached to the surface and suitable for colony generation.

The distance between the individual colony primers and the individual nucleic acid templates (and hence the density of the colony primers and nucleic acid templates) can be controlled by altering the concentration of colony primers and nucleic acid templates that are immobilised to the support. A preferred density of colony primers is at least 1 fmol/mm^2, preferably at least 10 fmol/mm^2, more preferably between 30 to 60 fmol/mm^2. The density of nucleic acid templates for use in the method of the invention is typically 10,000/mm^2 to 100,000/mm^2. It is believed that higher densities, for example, 100,000/mm^2 to 1,000,000/mm^2 and 1,000,000/mm^2 to 10,000,000/mm^2 may be achieved.

Controlling the density of attached nucleic acid templates and colony primers in turn allows the final density of nucleic acid colonies on the surface of the support to be controlled. This is due to the fact that according to the method of the invention, one nucleic acid colony can result from the attachment of one nucleic acid template, providing the colony primers of the invention are present in a suitable location on the

solid support (see in more detail below). The density of nucleic acid molecules within a single colony can also be controlled by controlling the density of attached colony primers.

Once the colony primers and nucleic acid templates of the invention have been immobilised on the solid support at the appropriate density, nucleic acid colonies of the invention can then be generated by carrying out an appropriate number of cycles of amplification on the covalently bound template nucleic acid so that each colony comprises multiple copies of the original immobilised nucleic acid template and its complementary sequence. One cycle of amplification consists of the steps of hybridisation, extension and denaturation and these steps are generally performed using reagents and conditions well known in the art for PCR.

A typical amplification reaction comprises subjecting the solid support and attached nucleic acid template and colony primers to conditions which induce primer hybridisation, for example subjecting them to a temperature of around 65°C. Under these conditions the oligonucleotide sequence Z at the 3' end of the nucleic acid template will hybridise to the immobilised colony primer X and in the presence of conditions and reagents to support primer extension, for example a temperature of around 72°C, the presence of a nucleic acid polymerase, for example, a DNA dependent DNA polymerase or a reverse transcriptase molecule (i.e. an RNA dependent DNA polymerase), or an RNA polymerase, plus a supply of nucleoside triphosphate molecules or any other nucleotide precursors, for example modified nucleoside triphosphate molecules, the colony primer will be extended by the addition of nucleotides complementary to the template nucleic acid sequence.

Examples of nucleic acid polymerases which can be used in the present invention are DNA polymerase (Klenow

fragment, T4 DNA polymerase), heat-stable DNA polymerases from a variety of thermostable bacteria (such as Taq, VENT, Pfu, Tfl DNA polymerases) as well as their genetically modified derivatives (TaqGold, VENTexo, Pfu exo). A combination of RNA polymerase and reverse transcriptase can also be used to generate the amplification of a DNA colony. Preferably the nucleic acid polymerase used for colony primer extension is stable under PCR reaction conditions, i.e. repeated cycles of heating and cooling, and is stable at the denaturation temperature used, usually approximately 94°C. Preferably the DNA polymerase used is Taq DNA polymerase.

Preferably the nucleoside triphosphate molecules used are deoxyribonucleotide triphosphates, for example dATP, dTTP, dCTP, dGTP, or are ribonucleoside triphosphates for example dATP, dUTP, dCTP, dGTP. The nucleoside triphosphate molecules may be naturally or non-naturally occurring.

After the hybridisation and extension steps, on subjecting the support and attached nucleic acids to denaturation conditions two immobilised nucleic acids will be present, the first being the initial immobilised nucleic acid template and the second being a nucleic acid complementary thereto, extending from one of the immobilised colony primers X. Both the original immobilised nucleic acid template and the immobilised extended colony primer formed are then able to initiate further rounds of amplification on subjecting the support to further cycles of hybridisation, extension and denaturation. Such further rounds of amplification will result in a nucleic acid colony comprising multiple immobilised copies of the template nucleic acid and its complementary sequence.

The initial immobilisation of the template nucleic acid means that the template nucleic acid can only hybridise with colony primers located at a distance

within the total length of the template nucleic acid. Thus the boundary of the nucleic acid colony formed is limited to a relatively local area to the area in which the initial template nucleic acid was immobilised. Clearly, once more copies of the template molecule and its complement have been synthesised by carrying out further rounds of amplification, ie. further rounds of hybridisation, extension and denaturation, then the boundary of the nucleic acid colony being generated will be able to be extended further, although the boundary of the colony formed is still limited to a relatively local area to the area in which the initial nucleic acid template was immobilised.

A schematic representation of a method of nucleic acid colony generation according to an embodiment of the present invention is shown in Figure 1. Figure 1(a) shows a colony primer X of the invention (shown here as having the sequence ATT), and a nucleic acid template of the invention containing at the 5' end an oligonucleotide sequence Y, here shown as ATT and at the 3' end an oligonucleotide sequence Z, here shown as AAT, which can hybridise to the colony primer sequence X. In the schematic representation the colony primer X and the oligonucleotide sequences Y and Z are shown as being of only three nucleotides in length. In practice however, it will be appreciated that longer sequences would normally be used. The 5' ends of both the colony primer and the nucleic acid template carry a means for attaching the nucleic acid to a solid support. This means is denoted in Figure 1 as a black square. This means of attaching may result in a covalent or a non-covalent attachment.

Only one colony primer X and one template nucleic acid are shown in Figure 1(a) for simplicity. However, in practice a plurality of colony primers X will be present with a plurality of nucleic acid templates. The plurality of colony primers X may comprise two different

colony primers X. However, for simplicity the schematic representation shown in Figure 1 shows only one type of colony primer X, with the sequence ATT. The plurality of nucleic acid templates may comprise different nucleic acid sequences in the central portion between the oligonucleotides Y and Z, but contain the same oligonucleotide sequences Y and Z at the 5' and 3' ends respectively. Only one species of nucleic acid template is shown for simplicity in Figure 1, in which a portion of the sequence in the central portion is shown as CGG.

In the presence of a solid support, the 5' ends of both the nucleic acid template and colony primer bind to the support. This is depicted in Figure 1(b). The support and the attached nucleic acid template and colony primers are then subjected to conditions which induce primer hybridisation. Figure 1(c) shows a nucleic acid template that has hybridised to a colony primer. Such hybridisation is enabled by virtue of the fact that the oligonucleotide sequence Z at the 3' end of the nucleic acid template can hybridise to the colony primer. In the schematic representation oligonucleotide sequence Z is shown to be complementary to the colony primer, although in practice an exact complementary sequence is not essential, providing hybridisation can occur under the conditions the nucleic acid templates and colony primers are subjected to.

Figure 1(d) shows the stage of primer extension. Here, under appropriate conditions of temperature and in the presence of a DNA polymerase and a supply of nucleotide precursors, for example dATP, dTTP, dCTP and dGTP, the DNA polymerase extends the colony primer from its 3' end using the nucleic acid template as a template. When primer extension is complete, see Figure 1(e), it can be seen that a second immobilised nucleic acid strand has been generated which is complementary to the initial nucleic acid template. On separating the two nucleic acid strands by, for example heating, two

immobilised nucleic acids will be present, the first being the initial immobilised nucleic acid template and the second being a nucleic acid complementary thereto, extending from one of the immobilised colony primers X, see Figure 1(f).

Both the original immobilised nucleic acid template and the immobilised extended colony primer formed are then able to hybridise to other colony primers present (depicted as colony primers 2 and 3 in Figure 1(g)) and after a further round of primer extension (Figure 1(h)) and strand separation (Figure 1(i)), four single stranded immobilised strands are provided. Two of these contain sequences corresponding to the original nucleic acid template and two contain sequences complementary thereto.

Further rounds of amplification beyond those shown in Figure 1 can be carried out to result in a nucleic acid colony comprising multiple immobilised copies of the template nucleic acid and its complementary sequence.

It can thus be seen that the method of the present invention allows the generation of a nucleic acid colony from a single immobilised nucleic acid template and that the size of these colonies can be controlled by altering the number of rounds of amplification that the nucleic acid template is subjected to. Thus the number of nucleic acid colonies formed on the surface of the solid support is dependent upon the number of nucleic acid templates which are initially immobilised to the support, providing there is a sufficient number of immobilised colony primers within the locality of each immobilised nucleic acid template. It is for this reason that preferably the solid support to which the colony primers and nucleic acid templates have been immobilised comprises a lawn of immobilised colony primers at an appropriate density with nucleic acid templates immobilised at intervals within the lawn of

primers.

Such so called "autopatterning" of nucleic acid colonies has an advantage over many methods of the prior art in that a higher density of nucleic acid colonies can be obtained due to the fact that the density can be controlled by regulating the density at which the nucleic acid templates are originally immobilised. Such a method is thus not limited by, for example, having specifically to array specific primers on particular local areas of the support and then initiate colony formation by spotting a particular sample containing nucleic acid template on the same local area of primer. The numbers of colonies that can be arrayed using prior art methods, for example those disclosed in WO96/04404 (Mosaic Technologies, Inc.) is thus limited by the density/spacing at which the specific primer areas can be arrayed in the initial step.

By being able to control the initial density of the nucleic acid templates and hence the density of the nucleic acid colonies resulting from the nucleic acid templates, together with being able to control the size of the nucleic acid colonies formed and in addition the density of the nucleic acid templates within individual colonies, an optimum situation can be reached wherein a high density of individual nucleic acid colonies can be produced on a solid support of a large enough size and containing a large enough number of amplified sequences to enable subsequent analysis or monitoring to be performed on the nucleic acid colonies.

Once nucleic acid colonies have been generated it may be desirable to carry out an additional step such as for example colony visualisation or sequence determination (see later). Colony visualisation might for example be required if it was necessary to screen the colonies generated for the presence or absence of for example the whole or part of a particular nucleic acid fragment. In this case the colony or colonies

which contain the particular nucleic acid fragment could be detected by designing a nucleic acid probe which specifically hybridises to the nucleic acid fragment of interest.

Such a nucleic acid probe is preferably labelled with a detectable entity such as a fluorescent group, a biotin containing entity (which can be detected by for example an incubation with streptavidin labelled with a fluorescent group), a radiolabel (which can be incorporated into a nucleic acid probe by methods well known and documented in the art and detected by detecting radioactivity for example by incubation with scintillation fluid), or a dye or other staining agent.

Alternatively, such a nucleic acid probe may be unlabelled and designed to act as a primer for the incorporation of a number of labelled nucleotides with a nucleic acid polymerase. Detection of the incorporated label and thus the nucleic acid colonies can then be carried out.

The nucleic acid colonies of the invention are then prepared for hybridisation. Such preparation involves the treatment of the colonies so that all or part of the nucleic acid templates making up the colonies is present in a single stranded form. This can be achieved for example by heat denaturation of any double stranded DNA in the colonies. Alternatively the colonies may be treated with a restriction endonuclease specific for a double stranded form of a sequence in the template nucleic acid. Thus the endonuclease may be specific for either a sequence contained in the oligonucleotide sequences Y or Z or another sequence present in the template nucleic acid. After digestion the colonies are heated so that double stranded DNA molecules are separated and the colonies are washed to remove the non-immobilised strands thus leaving attached single stranded DNA in the colonies.

After preparation of the colonies for

hybridisation, the labelled or unlabelled probe is then added to the colonies under conditions appropriate for the hybridisation of the probe with its specific DNA sequence. Such conditions may be determined by a person skilled in the art using known methods and will depend on for example the sequence of the probe.

The probe may then be removed by heat denaturation and, if desired, a probe specific for a second nucleic acid may be hybridised and detected. These steps may be repeated as many times as is desired.

Labelled probes which are hybridised to nucleic acid colonies can then be detected using apparatus including an appropriate detection device. A preferred detection system for fluorescent labels is a charge-coupled device (CCD) camera, which can optionally be coupled to a magnifying device, for example a microscope. Using such technology it is possible to simultaneously monitor many colonies in parallel. For example, using a microscope with a CCD camera and a 10x or 20x objective it is possible to observe colonies over a surface of between 1 mm^2 and 4 mm^2, which corresponds to monitoring between 10 000 and 200 000 colonies in parallel. Moreover, it is anticipated that this number will increase with improved optics and larger chips.

An alternative method of monitoring the colonies generated is to scan the surface covered with colonies. For example systems in which up to 100 000 000 colonies could be arrayed simultaneously and monitored by taking pictures with the CCD camera over the whole surface can be used. In this way, it can be seen that up to 100 000 000 colonies could be monitored in a short time.

Any other devices allowing detection and preferably quantitation of fluorescence on a surface may be used to monitor the nucleic acid colonies of the invention. For example fluorescent imagers or confocal microscopes could be used.

If the labels are radioactive then a radioactivity

detection system would be required.

In methods of the present invention wherein the additional step of performing at least one step of sequence determination of at least one of the nucleic acid colonies generated is performed, said sequence determination may be carried out using any appropriate solid phase sequencing technique. For example, one technique of sequence determination that may be used in the present invention involves hybridising an appropriate primer, sometimes referred to herein as a "sequencing primer", with the nucleic acid template to be sequenced, extending the primer and detecting the nucleotides used to extend the primer. Preferably the nucleic acid used to extend the primer is detected before a further nucleotide is added to the growing nucleic acid chain, thus allowing base by base in situ nucleic acid sequencing.

The detection of incorporated nucleotides is facilitated by including one or more labelled nucleotides in the primer extension reaction. Any appropriate detectable label may be used, for example a fluorophore, radiolabel etc. Preferably a fluorescent label is used. The same or different labels may be used for each different type of nucleotide. Where the label is a fluorophore and the same labels are used for each different type of nucleotide, each nucleotide incorporation can provide a cumulative increase in signal detected at a particular wavelength. If different labels are used then these signals may be detected at different appropriate wavelengths. If desired a mixture of labelled and unlabelled nucleotides are provided.

In order to allow the hybridisation of an appropriate sequencing primer to the nucleic acid template to be sequenced the nucleic acid template should normally be in a single stranded form. If the nucleic acid templates making up the nucleic acid

colonies are present in a double stranded form these can be processed to provide single stranded nucleic acid templates using methods well known in the art, for example by denturation, cleavage etc.

The sequencing primers which are hybridised to the nucleic acid template and used for primer extension are preferably short oligonucleotides, for example of 15 to 25 nucleotides in length. The sequence of the primers is designed so that they hybridise to part of the nucleic acid template to be sequenced, preferably under stringent conditions. The sequence of the primers used for sequencing may have the same or similar sequences to that of the colony primers used to generate the nucleic acid colonies of the invention. The sequencing primers may be provided in solution or in an immobilised form.

Once the sequencing primer has been annealed to the nucleic acid template to be sequenced by subjecting the nucleic acid template and sequencing primer to appropriate conditions, determined by methods well known in the art, primer extension is carried out, for example using a nucleic acid polymerase and a supply of nucleotides, at least some of which are provided in a labelled form, and conditions suitable for primer extension if a suitable nucleotide is provided. Examples of nucleic acid polymerases and nucleotides which may be used are described above.

Preferably after each primer extension step a washing step is included in order to remove unincorporated nucleotides which may interfere with subsequent steps. Once the primer extension step has been carried out the nucleic acid colony is monitored in order to determine whether a labelled nucleotide has been incorporated into an extended primer. The primer extension step may then be repeated in order to determine the next and subsequent nucleotides incorporated into an extended primer.

Any device allowing detection and preferably

quantitation of the appropriate label, for example fluorescence or radioactivity, may be used for sequence determination. If the label is fluorescent a CCD camera optionally attached to a magnifying device (as described above), may be used. In fact the devices used for the sequence determining aspects of the present invention may be the same as those described above for monitoring the amplified nucleic acid colonies.

The detection system is preferably used in combination with an analysis system in order to determine the number and nature of the nucleotides incorporated at each colony after each step of primer extension. This analysis, which may be carried out immediately after each primer extension step, or later using recorded data, allows the sequence of the nucleic acid template within a given colony to be determined.

If the sequence being determined is unknown, the nucleotides applied to a given colony are usually applied in a chosen order which is then repeated throughout the analysis, for example dATP, dTTP, dCTP, dGTP. If however, the sequence being determined is known and is being resequenced, for example to analyse whether or not small differences in sequence from the known sequence are present, the sequencing determination process may be made quicker by adding the nucleotides at each step in the appropriate order, chosen according to the known sequence. Differences from the given sequence are thus detected by the lack of incorporation of certain nucleotides at particular stages of primer extension.

Thus it can be seen that full or partial sequences of the amplified nucleic acid templates making up particular nucleic acid colonies may be determined using the methods of the present invention.

In a further embodiment of the present invention, the full or partial sequence of more than one nucleic acid can be determined by determining the full or

partial sequence of the amplified nucleic acid templates present in more than one nucleic acid colony. Preferably a plurality of sequences are determined simultaneously.

Carrying out sequence determination of nucleic acids using the method of the present invention has the advantage that it is likely to be very reliable due to the fact that large numbers of each nucleic acid to be sequenced are provided within each nucleic acid colony of the invention. If desired, further improvements in reliability can be obtained by providing a plurality of nucleic acid colonies comprising the same nucleic acid template to be sequenced, then determining the sequence for each of the plurality of colonies and comparing the sequences thus determined.

Preferably the attachment of the colony primer and nucleic acid template to the solid support is thermostable at the temperature to which the support may be subjected to during the nucleic acid amplification reaction, for example temperatures of up to approximately 100°C, for example approximately 94°C. Preferably the attachment is covalent in nature.

In a yet further embodiment of the invention the covalent binding of the colony primers and nucleic acid template(s) to the solid support is induced by a crosslinking agent such as for example 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC), succinic anhydride, phenyldiisothiocyanate or maleic anhydride, or a hetero-bifunctional crosslinker such as for example m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), N-succinimidyl[4-iodoacethyl] aminobenzoate (SIAB), Succinimidyl 4-[N-maleimidomethyl]cyclohexane-1-carboxylate (SMCC), N-y-maleimidobutyryloxy-succinimide ester (GMBS), Succinimidyl-4-[p-maleimidophenyl] butyrate (SMPB) and the sulfo (water-soluble) corresponding compounds. The preferred crosslinking reagents for use in the present invention, are s-SIAB,

s-MBS and EDC.

In a yet further embodiment of the invention the solid support has a derivatised surface. In a yet further embodiment the derivatised surface of the solid support is subsequently modified with bifunctional crosslinking groups to provide a functionalised surface, preferably with reactive crosslinking groups.

"Derivatised surface" as used herein refers to a surface which has been modified with chemically reactive groups, for example amino, thiol or acrylate groups.

"Functionalised surface" as used herein refers to a derivatised surface which has been modified with specific functional groups, for example the maleic or succinic functional moieties.

In the method of the present invention, to be useful for certain applications, the attachment of colony primers and nucleic acid templates to a solid support has to fulfill several requirements. The ideal attachment should not be affected by either the exposure to high temperatures and the repeated heating/cooling cycles employed during the nucleic acid amplification procedure. Moreover the support should allow the obtaining of a density of attached colony primers of at least 1 fmol/mm^2, preferably at least 10 fmol/mm^2, more preferably between 30 to 60 fmol/mm^2. The ideal support should have a uniformly flat surface with low fluorescence background and should also be thermally stable (non-deformable). Solid supports, which allow the passive adsorption of DNA, as in certain types of plastic and synthetic nitrocellulose membranes, are not suitable. Finally, the solid support should be disposable, thus should not be of a high cost.

For these reasons, although the solid support may be any solid surface to which nucleic acids can be attached, such as for example latex beads, dextran beads, polystyrene, polypropylene surface, polyacrylamide gel, gold surfaces, glass surfaces and

silicon wafers, preferably the solid support is a glass surface and the attachment of nucleic acids thereto is a covalent attachment.

The covalent binding of the colony primers and nucleic acid templates to the solid support can be carried out using techniques which are known and documented in the art. For example, epoxysilane-amino covalent linkage of oligonucleotides on solid supports such as porous glass beads has been widely used for solid phase in situ synthesis of oligonucleotides (via a 3'end attachment) and has also been adapted for 5' end oligonucleotide attachment. Oligonucleotides modified at the 5' end with carboxylic or aldehyde moieties have been covalently attached on hydrazine-derivatized latex beads (Kremsky et al 1987).

Other approaches for the attachment of oligonucleotides to solid surfaces use crosslinkers, such as succinic anhydride, phenyldiisothiocyanate (Guo et al 1994), or maleic anhydride (Yang et al 1998). Another widely used crosslinker is 1-ethyl-3-(3-dimethylamonipropyl)-carbodiimide hydrochloride (EDC). EDC chemistry was first described by Gilham et al (1968) who attached DNA templates to paper (cellulose) via the 5' end terminal phosphate group. Using EDC chemistry, other supports have been used such as, latex beads (Wolf et al 1987, Lund et al 1988), polystyrene microwells (Rasmussen et al 1991), controlled-pore glass (Ghosh et al 1987) and dextran molecules (Gingeras et al 1987). The condensation of 5' amino-modified oligonucleotides with carbodiimide mediated reagent have been described by Chu et al(1983), and by Egan et al (1982) for 5' terminal phosphate modification group.

The yield of oligonucleotide attachment via the 5' termini using carbodiimides can reach 60%, but non-specific attachment via the internal nucleotides of the oligonucleotide is a major drawback. Rasmussen et al (1991) have enhanced to 85% the specific attachment via

the 5' end by derivatizing the surface using secondary amino groups.

More recent papers have reported the advantages of the hetero-bifunctional cross-linkers. Hetero- or mono-bifunctional cross-linkers have been widely used to prepare peptide carrier conjugate molecules (peptide-protein) in order to enhance immunogenicity in animals (Peeters et al 1989). Most of these grafting reagents have been described to form stable covalent links in aqueous solution. These crosslinking reagents have been used to bind DNA onto a solid surface at only one point of the molecule.

Chrisey et al (1996) have studied the efficiency and stability of DNA solid phase attachment using 6 different hetero-bifunctional cross-linkers. In this example, the attachment occurs only at the 5' end of DNA oligomers modified by a thiol group. This type of attachment has also been described by O'Donnell-Maloney et al (1996) for the attachment of DNA targets in a MALDI-TOF sequence analysis and by Hamamatsu Photonics F.K. company (EP-A-665293) for determining base sequence of nucleic acid on a solid surface.

Very few studies concerning the thermal stability of the attachment of the oligonucleotides to the solid support have been done. Chrisey et al (1996) reported that with the Succinimidyl-4-[p-maleimidophenyl] butyrate (SMPB) cross-linker, almost 60% of molecules are released from the glass surface during heat treatment. But the thermal stability of the other reagents have not been described.

In order to generate nucleic acid colonies via the solid phase amplification reaction as described in the present application, colony primers and nucleic acid templates need to be specifically attached at their 5' ends to the solid surface, preferably glass. Briefly, the glass surface can be derivatized with reactive amino groups by silanization using amino-alkoxy silanes.

Suitable silane reagents include aminopropyltrimethoxysilane, aminopropyltriethoxysilane and 4-aminobutyltriethoxysilane. Glass surfaces can also be derivatized with other reactive groups, such as acrylate or epoxy using epoxysilane, acrylatesilane and acrylamidesilane. Following the derivatization step, nucleic acid molecules (colony primers or nucleic acid templates) having a chemically modifiable functional group at their 5' end, for example phosphate, thiol or amino groups are covalently attached to the derivatized surface by a crosslinking reagent such as those described above.

Alternatively, the derivatization step can be followed by attaching a bifunctional cross-linking reagent to the surface amino groups thereby providing a modified functionalized surface. Nucleic acid molecules (colony primers or nucleic acid templates) having 5'-phosphate, thiol or amino groups are then reacted with the functionalized surface forming a covalent linkage between the nucleic acid and the glass.

Potential cross-linking and grafting reagents that can be used for covalent DNA/oligonucleotide grafting on the solid support include succinic anhydride, (1-ethyl-3[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC), m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), N-succinimidyl[4-iodoacethyl] aminobenzoate (SIAB), Succinimidyl 4-[N-maleimidomethyl]cyclohexane-1-carboxylate (SMCC), N-y-maleimidobutyryloxy-succinimide ester (GMBS), Succinimidyl-4-[p-maleimidophenyl] butyrate (SMPB) and the sulfo (water-soluble) corresponding compounds. The preferred crosslinking reagents, according to the present invention, are s-SIAB, s-MBS and EDC. s-MBS is a maleimide-succinimide hetero-bifunctional cross-linker and s-SIAB is an iodoacethyl-succinimide hetero-bifunctional cross-linker. Both are capable of forming a covalent bond respectively with SH groups and primary amino groups. EDC is a carbodiimide-reagent that mediates covalent

attachment of phosphate and amino groups.

The colony primers and nucleic acid templates are generally modified at the 5' end by a phosphate group or by a primary amino group (for EDC grafting reagent) or a thiol group (for s-SIAB or s-MBS linkers).

Thus, a further aspect of the invention provides a solid support, to which there is attached a plurality of colony primers X as described above and at least one nucleic acid template as described above, wherein said nucleic acid templates contain at their 5' ends an oligonucleotide sequence Y as described above, and at their 3' ends an oligonucleotide sequence Z as described above. Preferably a plurality of nucleic acid templates are attached to said solid support, which is preferably glass. Preferably the attachment of the nucleic acid templates and colony primers to the solid support is covalent. By performing one or more rounds of nucleic acid amplification on the immobilised nucleic acid template(s) using methods as described above, nucleic acid colonies of the invention may be formed. A yet further aspect of the invention is, therefore, a support comprising one or more nucleic acid colonies of the invention. A yet further aspect of the invention provides the use of the solid supports of the invention in methods of nucleic acid amplification or sequencing. Such methods of nucleic acid amplification or sequencing include the methods of the present invention.

A yet further aspect of the invention provides the use of a derivatized or functionalized support, prepared as described above in methods of nucleic acid amplification or sequencing. Such methods of nucleic acid amplification or sequencing include the methods of the present invention.

A yet further aspect of the invention provides an apparatus for carrying out the methods of the invention or an apparatus for producing a solid support comprising nucleic acid colonies of the invention. Such apparatus

might comprise for example a plurality of nucleic acid templates and colony primers of the invention bound, preferably covalently, to a solid support as outlined above, together with a nucleic acid polymerase, a plurality of nucleotide precursors such as those described above, a proportion of which may be labelled, and a means for controlling temperature. Alternatively, the apparatus might comprise for example a support comprising one or more nucleic acid colonies of the invention. Preferably the apparatus also comprises a detecting means for detecting and distinguishing signals from individual nucleic acid colonies arrayed on the solid support according to the methods of the present invention. For example such a detecting means might comprise a charge-coupled device operatively connected to a magnifying device such as a microscope as described above.

Preferably any apparatuses of the invention are provided in an automated form.

The present application provides a solution to current and emerging needs that scientists and the biotechnology industry are trying to address in the fields of genomics, pharmacogenomics, drug discovery, food characterization and genotyping. Thus the method of the present invention has potential application in for example: nucleic acid sequencing and re-sequencing, diagnostics and screening, gene expression monitoring, genetic diversity profiling, whole genome polymorphism discovery and scoring, the creation of genome slides (whole genome of a patient on a microscope slide) and whole genome sequencing.

Thus the present invention may be used to carry out nucleic acid sequencing and re-sequencing, where for example a selected number of genes are specifically amplified into colonies for complete DNA sequencing. Gene re-sequencing allows the identification of all known or novel genetic polymorphisms of the investigated

genes. Applications are in medical diagnosis and genetic identification of living organisms.

For use of the present invention in diagnostics and screening, whole genomes or fractions of genomes may be amplified into colonies for DNA sequencing of known single nucleotide polymorphisms (SNP). SNP identification has application in medical genetic research to identify genetic risk factors associated with diseases. SNP genotyping will also have diagnostic applications in pharmaco-genomics for the identification and treatment of patients with specific medications.

For use of the present invention in genetic diversity profiling, populations of for example organisms or cells or tissues can be identified by the amplification of the sample DNA into colonies, followed by the DNA sequencing of the specific "tags" for each individual genetic entity. In this way, the genetic diversity of the sample can be defined by counting the number of tags from each individual entity.

For use of the present invention in gene expression monitoring, the expressed mRNA molecules of a tissue or organism under investigation are converted into cDNA molecules which are amplified into sets of colonies for DNA sequencing. The frequency of colonies coding for a given mRNA is proportional to the frequency of the mRNA molecules present in the starting tissue. Applications of gene expression monitoring are in bio-medical research.

A whole genome slide, where the entire genome of a living organism is represented in a number of DNA colonies numerous enough to comprise all the sequences of that genome may be prepared using the methods of the invention. The genome slide is the genetic card of any living organism. Genetic cards have applications in medical research and genetic identification of living organisms of industrial value.

The present invention may also be used to carry out

whole genome sequencing where the entire genome of a living organism is amplified as sets of colonies for extensive DNA sequencing. Whole genome sequencing allows for example, 1) a precise identification of the genetic strain of any living organism; 2) to discover novel genes encoded within the genome and 3) to discover novel genetic polymorphisms.

The applications of the present invention are not limited to an analysis of nucleic acid samples from a single organism/patient. For example, nucleic acid tags can be incorporated into the nucleic acid templates and amplified, and different nucleic acid tags can be used for each organism/patient. Thus, when the sequence of the amplified nucleic acid is determined, the sequence of the tag may also be determined and the origin of the sample identified.

Thus, a further aspect of the invention provides the use of the methods of the invention, or the nucleic acid colonies of the invention, or the plurality of nucleic acid templates of the invention, or the solid supports of the invention, for providing nucleic acid molecules for sequencing and re-sequencing, gene expression monitoring, genetic diversity profiling, diagnosis, screening, whole genome sequencing, whole genome polymorphism discovery and scoring and the preparation of whole genome slides (ie. the whole genome of an individual on one support), or any other applications involving the amplification of nucleic acids or the sequencing thereof.

A yet further aspect of the invention provides a kit for use in sequencing, re-sequencing, gene expression monitoring, genetic diversity profiling, diagnosis, screening, whole genome sequencing, whole genome polymorphism discovery and scoring, or any other applications involving the amplification of nucleic acids or the sequencing thereof. This kit comprises a plurality of nucleic acid templates and colony primers

of the invention bound to a solid support, as outlined above.

The invention will now be described in more detail in the following non-limiting Examples with reference to the following drawings in which:

Figure 1: shows a schematic representation of a method of nucleic acid colony generation according to an embodiment of the invention.

Figure 2: Schematic representation of template preparation and subsequent attachment to the solid surface. In Figure 2a the preparation of Templates A, B and B' containing colony primer sequences is shown. The 3.2 Kb template is generated from genomic DNA using PCR primers TP1 and TP2. Templates A (854 bp) and B (927 bp) are generated using PCR primers TPA1/TPA2 or TPB1/TPB2, respectively. The TPA1 and TPB1 oligonucleotides are modified at their 5'-termini with either a phosphate or thiol group for subsequent chemical attachment (★). Note that the templates obtained contain sequences corresponding to colony primers CP1 and/or CP2. The 11 exons of the gene are reported as "E1 to E11". In Figure 2b the chemical attachment of colony primers and templates to glass surface is shown. Derivatization by ATS (aminopropyl-triethoxysilane) is exemplified.

Figure 3: DNA colonies generated from a colony primer. It shows the number of colonies observed per 20X field as a function of the concentration of template bound to the well. The lowest concentration of detectable template corresponds to 10^{-13} M.

Figure 4: Representation of discrimination between colonies originated from two different templates. Figure 4a shows the images of colonies made from both templates and negative controls. Figure 4b shows the colonies from both templates at the same position in the same well visualised with two different colours and negative controls. Figure 4c shows the coordinates of

both colony types in a sub-section of a microscopy field. Figure 4c demonstrates that colonies from different templates do not coincide.

Figure 5: Reaction schemes of the template or oligonucleotide attachment on glass. Step A is the derivatization of the surface: glass slide are treated with acidic solution to enhance free hydroxyl group on the surface. The pretreated slides are immersed into a solution of aminosilane. ATS: Aminopropyl triethoxysilane. Step B: B1 or B2 is the functionalization of glass surface with cross-linkers followed by oligonucleotide attachment. Amino group reacts with a cross linking agent via an amide bond: step B1; s-MBS (sulfo m-maleimidobenzoyl-N-hydroxy-succinimide ester) step B2; s-SIAB (sulfo N-succinimidyl[4-iodoacethyl] aminobenzoate). The oligonucleotides (5'end thiol modified oligonucleotide) are attached to the surface via formation of a covalent bound between the thiol and the double bond of the maleimide. Phosphate buffered saline: (PBS, 0.1 M NaH_2PO_4, pH:6.5, 0.15 M NaCl). B3: Attachment of oligonucleotides using EDC and Imidazole. 5'end phosphate of the modified oligonucleotides reacts with imidazole in the presence of EDC to give 5'-phosphor-imidazolide derivatives (not shown). The derivatives form a phosphoramidate bond with amino groups of the derivatized glass surface. EDC: 1-ethyl-3-(3-dimethyl-amonipropyl)-carbodiimide hydrochloride.

Figure 6: it shows the number of colonies observed per 20X field as a function of the concentration of template bound to the well. DNA template were bound at different concentration either via the mediated coupling reagent (EDC) on amino derivatized glass surface (A) or on s-MBS functionalized glass surface (B). Double strand DNA colonies were submitted to restriction enzyme and the recovered single strands hybridized with a complementary oligonucleotide, cy5 fluorescently

labeled.

Figure 7: shows an example of in situ sequencing from DNA colonies generated on glass. Figure 7A shows the result after incubation with Cy5™-dCTP on a sample that has not been incubated with primer p181. One will appreciate only 5 blurry spots can be observed, indicating that no dramatic spurious effect is taking place (such as Cy5™-dCTP aggregate precipitation, adsorption or simply non specific incorporation to the DNA in the colonies or on the surface). Figure 7B shows the result after incubation with Cy5™-dUTP on a sample that has been incubated with primer p181. One will appreciate that no fluorescent spot can be observed, indicating that the incorporation of a fluorescent base cannot take place in detectable amounts when the nucleotide proposed for incorporation does not correspond to the sequence of the template following the hybridized primer. Figure 7C shows the result after incubation with Cy5™-dCTP on a sample that has been incubated with primer p181. One will appreciate that many fluorescent spots can be observed, indicating that the incorporation of a fluorescent base can indeed take place in detectable amounts when the nucleotide proposed for incorporation does correspond to the sequence of the template following the hybridized primer.

Figure 8: shows hybridization of probes to oligonucleotides attached to Nucleolink, before and after PCR cycling. The figure shows R58 hybridization to CP2 (5'-(phosphate)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) closed circles, CP8 (5'(amino-hexamethylene)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) closed triangles, CP9 (5'(hydroxyl)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) diamonds, CP10 (5'(dimethoxytrityl)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) open circles and

CP11 (5'(biotin)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) open triangles.

EXAMPLES

EXAMPLE 1: Preparation of DNA templates suitable for the generation of DNA colonies

DNA colonies have been generated from DNA templates and colony primers. The term "colony primer sequence" as used herein refers to a sequence corresponding to the sequence of a colony primer and is elsewhere sometimes referred to as "oligonucleotide sequence Y" or "oligonucleotide sequence Z'".

The properties of the colony primers have been chosen based on a selection for oligonucleotide primers that show little non-specific nucleotide incorporation in the presence of heat-stable DNA polymerases. The colony primers, CPα (5'-p CACCAACCCAAACCAACCCAAACC) and CPβ (5'-p AGAAGGAGAAGGAAAGGGAAAGGG) have been selected due to their low incorporation of radiolabeled [α^{32}P-dCTP] in the presence of a stable DNA polymerase (AmpliTaq, Perkin Elmer, Foster City, CA) in the standard buffer and under thermocycling conditions (94°C for 30 seconds, 65°C for 1 minute, 72°C for 2 minutes, 50 cycles).

A 3.2 Kb DNA fragment was taken as a model system to demonstrate the feasibility of colony generation using colony primers and DNA templates. The chosen template comprises the human gene for the receptor for advanced glycosylation end-products (HUMOXRAGE, GenBank Acc. No. D28769). The RAGE-specific primers are depicted in Table 1. The 3.2 Kb template was generated by PCR amplification from 0.1 μg human genomic DNA with 1 μM primers TP1 and TP2 with 1 unit of DNA polymerase (AmpliTaq, Perkin Elmer, Foster City, CA) in the

standard buffer and under thermocycling conditions (94°C for 30 seconds, 65°C for 1 minute, 72°C for 5 minutes, 40 cycles). This 3.2 Kb DNA fragment was used as a template for secondary PCR to generate two shorter templates for colony generation (Templates A and B). The primers used to generate the shorter templates contain both sequences specific to the template and sequences of colony primers CP1 and CP2 to amplify the DNA on the solid surface. In general, the PCR primer used to generate a DNA template is modified at the 5'-terminus with either a phosphate or thiol moiety. Thus after the PCR amplification, DNA fragments are generated which contain the colony primer sequences at one or both termini adjoining the RAGE DNA fragment of interest (see Figure 2a).

Template A (double stranded template containing the colony primer sequence, CPβ at both termini) was generated with 0.1 ng of the 3.2 Kb template with 1 μM primers TPA1 and 1 μM TPA2 with 1 unit of DNA polymerase (AmpliTaq, Perkin Elmer, Foster City, CA) in the standard buffer and under thermocycling conditions (94°C for 30 seconds, 65°C for 1 minute, 72°C for 1 minutes, 30 cycles). The products were then purified over Qiagen Qia-quick columns (Qiagen GmbH, Hilden, Germany).

Template B (double stranded template which contains colony primer sequences corresponding to CPβ) was generated with 0.1 ng of the 3.2 Kb template with 1 μM primers TPB1 and 1 μM TPB2 with 1 unit of DNA polymerase (AmpliTaq, Perkin Elmer, Foster City, CA) in the standard buffer and under thermocycling conditions (94°C for 30 seconds, 65°C for 1 minute, 72°C for 1 minutes, 30 cycles). The products were then purified over Qiagen Qia-quick columns (Qiagen GmbH, Hilden, Germany).

Template B' (double stranded template containing colony

primer sequences corresponding to CPα and CPβ at either end) was generated with 0.1 ng of the 3.2 Kb template with 1 μM primers TPB3 and 1 μM TPB4 with 1 unit of (AmpliTaq, Perkin Elmer, Foster City, CA) in the standard buffer and under thermocycling conditions (94°C for 30 seconds, 65°C for 1 minute, 72°C for 1 minutes, 30 cycles). The products were then purified over Qiagen Qia-quick columns (Qiagen GmbH, Hilden, Germany).

All the specific oligonucleotides employed for the DNA templates preparation and for the DNA colony generation have been reported in the Table 1 together with any chemical modification.

A general scheme showing the chemical attachment of colony primers and templates to the glass surface is reported in Figure 2b, where the derivatization by ATS (aminopropyltriethoxysilane) is reported, as a non-limitative example.

TABLE 1

List of oligonucleotides used for templates preparation and colonies generation:

Name	DNA sequence	Coordinates (orientation)	Oligonucleotide Modification	Use
TP1	GAGGCCAGAACAGT TCAAGG	9810 (R)		Template 3.2 Kb
TP2	CCTGTGACAAGACG ACTGAA	6550 (F)		Template 3.2 Kb
CP1	TTTTTTTTTTCACC AACCCAAACCAACC CAAACC	None	5'P	Generate colonies
CP2	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'P	Generate colonies
CP3	TTTTTTTTTTCACC AACCCAAACCAACC CAAACC	None	5'SH	Generate colonies
CP4	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'SH	Generate colonies
CP5	AGAAGGAGAAGGAA AGGGAAAGGGTTTT TTTTTTTTTTTNN	None	5'P	Generate colonies
CP6	AGAAGGAGAAGGAA AGGGAAAGGGGG	None	5'P	Generate colonies
CP7	TTTTTTTTTTCACC AACCCAAACCAACC CAAACC	None	5'(NH_2)	Generate colonies
CP8	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'(NH_2)	Generate colonies
CP9	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'(OH)	Control oligo
CP10	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'(DMT)	Control oligo
CP11	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGG	None	5'(biotin)	Control oligo

TPA1	AGAAGGAGAAGGAA AGGGAAAGGGCCTG TGACAAGACGACTG AA	6550 (F)	5'P	Template A
TPA2	TTTTTTTTTTAGAA GGAGAAGGAAAGGG AAAGGGGCGGCCGC TGAGGCCAGTGGAA GTCAGA	7403 (R)	5'P	Template A
TPB3	TTTTTTTTTTCACC AACCCAAACCAACC CAAACCGAGCTCAG GCTGAGGCAGGAGA ATTG	9049 (F)	None	Template B'
TPB1	AGAAGGAGAAGGAA AGGGAAAGGGGAGC TGAGGAGGAAGAGA GG	9265 (F)	None	Template B
TPB2	AGAAGGAGAAGGAA AGGGAAAGGGGCGG CCGCTCGCCTGGTT CTGGAAGACA	8411 (R)	5'P	Template B
TPB4	AGAAGGAGAAGGAA AGGGAAAGGGGCGG CCGCTCGCCTGGTT CTGGAAGACA	9265 (R)	5'SH	Template B'

Coordinate from HUMOXRAGE gene Accession number D28769
(R) means "reverse" and (F) means "forward"

EXAMPLE 1a: Preparation of a random DNA template flanked by a degenerate primer

A 3.2 Kb DNA fragment was taken as a model system to demonstrate the feasibility of colony generation from random primer PCR amplification. This strategy can be applied to sequencing of DNA fragments of approximately 100 Kb in length and, by combination of fragments to whole genomes. A fragment of DNA of 3.2 Kb was generated by PCR from human genomic DNA using PCR primers; TP1 5'-pGAGGCCAGAACAGTTCAAGG and TP2 5'-pCCTGTGACAAGACGACTGAA as described in example 1. The 3.2 Kb fragment was cut in smaller fragments by a combination of restriction enzymes (EcoR1 and HhaI

yielding 4 fragments of roughly 800 bp). The cut or uncut fragment DNAs were then mixed with the degenerate primer, p252 (5'-P TTTTTTTTTTISISISISISIS, where I stands for inosine (which pairs with A, T and C) and S stands for G or C) and covalently coupled to the Nucleolink wells (Nunc, Denmark). The tubes were then subjected to random solid phase PCR amplification and visualized by hybridisation with labeled DNA probes, as will be described in Example 2a.

EXAMPLE 2: Covalent binding of DNA templates and colony primers on solid support (plastic) and colony formation with a colony primer

Covalent binding of template and colony primer to the solid support (plastic)

A colony primer (CP2, 5'-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), phosphorylated at its 5' terminus (Microsynth GmBH, Switzerland), was attached onto Nucleolink plastic microtitre wells (Nunc, Denmark) in the presence of varying doses of Template A (prepared as described in example 1). 8 wells were set up in duplicate with seven 1/10 dilutions of template with CP2, starting with the highest concentration of 1 nM.

Microtitre wells, to which CP2 colony primer and the template are covalently bound were prepared as follows. In each Nucleolink well, 30 μl of a 1 μM solution of the colony primer with varying concentrations of template diluted down from 1 nM in 10 mM 1-methyl-imidazole (pH 7.0) (Sigma Chemicals) was added. To each well, 10 μl of 40 mM 1-ethyl-3-{3-dimethylaminopropyl)-carbodiimide (pH 7.0) (Sigma Chemicals) in 10 mM 1-methyl-imidazole, was added to the solution of colony primer and template. The wells were then sealed and incubated at 50°C overnight. After the incubation, wells were rinsed

twice with 200 µl of RS (0.4 N NaOH, 0.25% Tween 20), incubated 15 minutes with 200 µl RS, washed twice with 200 µl RS and twice with 200 µl TNT (100 mM TrisHCl pH 7.5, 150 mM NaCl, 0.1% Tween 20). Tubes were dried at 50°C and were stored in a sealed plastic bag at 4°C.

Colony generation

Colony growing was initiated in each well with 20 µl of PCR mix; the four dNTPs (0.2 mM), 0.1% BSA (bovine serum albumin), 0.1% Tween 20, 8% DMSO (dimethylsulfoxide, Fluka, Switzerland), 1X PCR buffer and 0.025 units/µl of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA). The wells were then placed in the thermocycler and growing was performed by incubating the sealed wells 5 minutes at 94°C and cycling for 50 repetitions the following conditions: 94°C for 30 seconds, 65°C for 2 minutes, 72°C for 2 minutes. After completion of this program, the wells were kept at 8°C until further use. Prior to hybridization wells are filled with 50 µL TE (10 mM Tris, 1 mM EDTA, pH 7.4) heated at 94°C for 5 minutes and chilled on ice before probe addition at 45°C.

Colonies Visualization

Probe: The probe was a DNA fragment of 1405 base pairs comprising the sequence of the template at their 3' end (nucleotide positions 8405 to 9259). The DNA probe was synthesized by PCR using two primers: p47 (5'-GGCTAGGAGCTGAGGAGGAA), amplifying from base 8405, and TP2, biotinylated at 5' end, amplifying from base 9876 of the antisense strand.

Hybridization and detection: The probe was diluted to 1 nM in "easyhyb" (Boehringer-Mannheim, Germany) and 20 µL added to each well. The probe and the colonies were denatured at 94°C for 5 min and then incubated 6 hours at 50°C. Excess probes was washed at 50°C in 2xSSC with

0.1% Tween. The DNA probes were bound to avidin coated green fluorescence fluorospheres of a diameter of 0.04 μ (Molecular Probes) in TNT for 1 hour at room temperature. Excess beads were washed with TNT. Colonies were visualized by microscopy and image analysis as described in example 2a. Figure 3 shows the number of colonies observed per 20X field as a function of the concentration of template bound to the well. The lowest concentration of detectable template corresponds to 10^{-13} M.

Example 2a: Covalent binding of DNA templates and colony primers on solid support (plastic) and colony formation with a degenerate primer

Covalent binding of template and colony primer to the solid support (plastic)

Microtitre wells with p252 and template DNA fragments were prepared as follows:

In each Nucleolink well, 30 μl of a 1 μM solution of the colony primer p252 with varying concentrations of template diluted down from 0.5 nM in 10 mM 1-methyl-imidazole (pH 7.0) (Sigma Chemicals) was added. To each well, 10 μl of 40 mM 1-ethyl-3-(3-dimethylamino-propyl)-carbodiimide (pH 7.0) (Sigma Chemicals) in 10 mM 1-methyl-imidazole, was added to the solution of colony primer and template. The wells were then sealed and incubated at 50°C overnight. After the incubation, wells were rinsed twice with 200 μl of RS (0.4N NaOH, 0.25% Tween 20), incubated 15 minutes with 200 μl RS, washed twice with 200 μl RS and twice with 200μl TNT (100mM TrisHCl pH7.5, 150mM NaCl, 0.1% Tween 20). Tubes were dried at 50°C and were stored in a sealed plastic bag at 4°C.

Colony generation

DNA colony generation was performed with a modified protocol to allow random priming in each well with 20 µl of PCR mix; the four dNTPs (0.2 mM each), 0.1% BSA, 0.1% Tween 20, 8% DMSO (dimethylsulfoxide, Fluka, Switzerland), 1X PCR buffer and 0.025 units/µl of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA). The wells were then placed in the thermocycler and amplification was performed by incubating the sealed wells 5 minutes at 94°C and cycling for 50 repetitions the following conditions: 94°C for 30 seconds, 65°C for 2 minutes, 72°C for 2 minutes. After completion of this program, the wells were kept at 8°C until further use. Prior to hybridization wells are filled with 50 µL TE (10 mM Tris 1 mM EDTA pH 7.4) heated at 94°C for 5 minutes and chilled on ice before probe addition at 45°C.

Colonies Visualization

Probes: Two DNA fragments of 546 and 1405 base pairs comprising the sequences of either extremities of the original template were amplified by PCR. The antisense strand of the probe was labeled with biotin, through the use of a 5'-biotinylated PCR primer. The base pair coordinates of the probes were 6550 to 7113 and 6734 to 9805.

Hybridization and detection: The probes were diluted to 1 nM in "easyhyb" (Boehringer-Mannheim, Germany) and 20 µL added to each well. The probe and the colonies were denatured at 94°C for 5 min and then incubated 6 hours at 50°C. Excess probes was washed at 50°C in 2xSSC with 0.1% tween. The DNA probes were bound to avidin coated green fluorescence fluorospheres of a diameter of 40 nanometers (Molecular Probes, Portland OR) in TNT for 1 hour at room temperature. Excess beads were washed off with TNT. Fluorescence was detected using an inverted

microscope (using the 20x/0.400 LD Achroplan objective, on the Axiovert S100TV, with an arc mercury lamp HBO 100W/2, Carl Zeiss, Oberkochen, Germany) coupled to a 768(H)x512(V)pixel-CCD camera (Princeton Instruments Inc. Trenton, NJ, USA). Exposure were 20 seconds through filter sets XF22 (Ex: 485DF22, Dichroic: 505DRLPO2 Em: 530DF30) and XF47 (Ex: 640DF20, Dichroic: 670DRLPO2 Em: 682DF22) from Omega Optical (Brattleboro VT) for FITC and Cy5 respectively. Data were analyzed using Winwiew software (Princeton Instruments Inc., Trenton NJ, USA). The numbers of colonies per field were counted in duplicate wells with image analysis software developed in house.

EXAMPLE 3: Sequence discrimination in different colonies originated from varying ratios of 2 different covalently bound templates and a colony primer

Covalent binding of templates and colony primer to the solid support (plastic)

A colony primer (CP2: 5'pTTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), phosphorylated at its 5' termini (Microsynth GmbH, Switzerland), was grafted onto Nucleolink plastic microtitre wells (Nunc, Denmark) in the presence of varying doses of the two templates A and B (prepared as described in example 1). Series of 8 wells were set up in triplicate with seven 1/10 dilutions of both templates starting with the highest concentration of 1 nM. Template dilutions are set up in opposite directions such that the highest concentration of one template coincides with the lowest of the other.

Microtitre wells, to which CP2 primer and both templates are covalently bound were prepared as follows. In each Nucleolink well, 30 μl of a 1 μM solution of the CP2 primer with varying concentrations of both templates

diluted down from 1 nM in 10 mM 1 methyl-imidazole (pH 7.0) (Sigma Chemicals) were added. To each well, 10 µl of 40 mM 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (pH 7.0) (Sigma Chemicals) in 10 mM 1-methyl-i imidazole (pH 7.0), was added to the solution of colony primer and templates. The wells were then sealed and incubated at 50°C for 4 hours. After the incubation, the wells were rinsed three times with 50 µl of RS (0.4 N NaOH, 0.25% Tween 20), incubated 15 minutes with 50 µl RS, washed three times with 50 µl RS and three times with 50 µl TNT (100 mM TrisHCl pH 7.5, 150 mM NaCl, 0.1% Tween 20). Tubes were stored in TNT at 4°C.

Colonies generation

Colony growing was initiated in each well with 20 µl of PCR mix; the four dNTPs (0.2 mM), 0.1% BSA, 0.1% Tween 20, 8% DMSO (dimethylsulfoxide, Fluka, Switzerland), 1X PCR buffer and 0.025 units/µl of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA).

The wells were then placed in the thermocycler and growing was performed by incubating the sealed wells 5 minutes at 94°C and cycling for 50 repetitions the following conditions: 94°C for 30 seconds, 65°C for 5 minutes, 72°C for 5 minutes. After completion of this program, the wells were kept at 8°C until further use. Prior to hybridization wells are filled with 50 µl TE (10 mM Tris, 1 mM EDTA, pH 7.4) heated at 94°C for 5 minutes and chilled on ice before probe addition at 50°C.

Colonies visualization

Probe: Two DNA fragments of 546 and 1405 base pairs corresponding to the sequences of the 3.2 Kb DNA fragment at the 5'- and 3'- termini were amplified by PCR using one biotinylated primer (see example 2). The two probes were denatured by heating at 94°C for 5

minutes, quick-chilled into 1 M NaCl, 10 mM Tris pH 7.4 and allowed to bind to Strepatividin coated fluorospheres of diameter 0.04 μm labeled with different colors for 2 hours at 4°C. The probes bound to bead were diluted 20 fold in "easyhyb" solution prewarmed to 50°C. 20 μl of probes was added to each well containing denatured colonies.

Hybridization and detection: The hybridization was carried out at 50°C for 5 hours. Excess probes was washed at 50°C in 2xSSC with 0.1% SDS. Colonies were visualized by microscopy with a 20X objective, 20 second exposure and image analysis as described in example 2a. Figure 4a shows the images of colonies made from both templates and negative controls. Figure 4b shows the colonies from both templates at the same position in the same well visualised with two different colours and negative controls. Figure 4c shows the coordinates of both colony types in a sub-section of a microscopy field. Figure 4c demonstrates that colonies from different templates do not coincide.

EXAMPLE 4: Covalent binding of DNA templates and oligonucleotides on glass solid supports

Aminosilane-derivatized glass slides have been used as solid support to covalently attach thiol-modified oligonucleotides probes using hetero-bifunctional cross-linkers. The reagents selected have thiol-reactive (maleimide) and amino-reactive groups (succinimidyl ester). Oligonucleotide attachment yields and stability of the immobilized molecules will be strongly dependent on the cross-linker stability towards the conditions of the different treatments performed. The reaction schemes of the DNA templates or oligonucleotides attachment on glass are described in Figure 5.

The storage stability of glass slides prepared with the cross-linkers s-MBS and s-SIAB and its thermal stability have been evaluated. An important factor affecting the extent of hybridization of immobilized oligonucleotide probes is the density of attached probes (Beattie et al., 1995; Joss et al., 1997). We have studied this effect by varying the concentration of oligonucleotides during the immobilization and assaying the density of attached oligos by hybridization.

Materials and methods

Microscope glass slides acid pre-treatment - Microscope glass slides (Knittel, Merck ABS) were soaked in basic Helmanex solution during 1 hour (HelmanexII^R 0.25%, 1N NaOH). The slides were rinsed with water, immersed overnight in 1N HCl, rinsed again in water and treated 1 hour in sulfuric acid solution (H_2SO_4/H_2O, 1/1, v/v, with a small amount of fresh ammonium persulfate added). The slides were rinsed in water, in ethanol and finally with pure acetone. Glass slides are dried and stored under vacuum for further use.

Silanization of the surface - The pre-treated slides were immersed into a 5% solution of ATS (aminopropyltriethoxysilane, Aldrich) in acetone. Silanization was carried out at room temperature for 2 hours. After three washes in acetone (5 min/wash) the slides were rinsed once with ethanol, dried and stored under vacuum.

Cross-linker attachment - Cross-linkers, s-MBS and s-SIAB (respectively sulfo m-maleimidobenzoyl-N-hydroxysuccinimide ester, sulfo N-succinimidyl[4-iodoacethyl] aminobenzoate, Pierce, Rockford IL), are prepared as 20 mM solutions in PBS (phosphate-buffered saline, 0.1 M NaH_2PO_4, pH 7.2, 0.15 M NaCl). Silanized

glass slides, on which 80 μL of cross-linker solution was applied, were covered by a cleaned micro cover glass and reacted for 5 hours at 20°C. The glass slides were rinsed in PBS, briefly immersed in water and rinsed in ethanol. Slides were then dried and stored under vacuum in the dark for further use.

Oligonucleotide Attachment - Oligonucleotides were synthesized with 5' modifications of a thiol (CP3 and CP4 Eurogentec, Brussels) or a phosphate moiety (CP1 and CP2, Eurogentec, Brussels) using standard phosphoramidite chemistry.

- 5'- thiol oligonucleotide primers (CP3 and CP4) were prepared as 100 μM solutions in a saline phosphate buffer (NaPi : 0.1M NaH_2PO_4 pH: 6.6, 0.15M NaCl) and drops of 1 μl applied on the functionalized glass slide (functionalized with cross-linker) for 5 hours at room temperature. Glass slides were kept under a saturated wet atmosphere to avoid evaporation. Glass slides were washed on a shaker in NaPi buffer. For thermal stability study glass slides were immersed 2 times in Tris buffer (10 mM, pH 8) for 5 min at 100°C and directly immersed in 5xSSC (0.75 M NaCl, 0.075 M NaCitrate pH 7) at 4°C for 5 min. Slides were stored in 5XSSC at 4°C for further use.

- 5'- phosphate oligonucleotides primers (CP1 and CP2) were applied (1 μl drops) for 5 hours at room temperature to amino-derivatized glass as 1 μM solution in 10 mM 1-methyl-imidazole (pH 7.0) (Sigma Chemicals) containing 40 mM of 1-ethyl-3-(3-dimethylamino-propyl) carbodiimide (EDC, Pierce, Rockford IL). The slides were washed 2 times at 100°C in Tris buffer (10 mM, pH 8) and directly immersed in 5XSSC at 4°C for 5 min. Slides were stored in 5XSSC at 4°C for further use.

Oligonucleotide and DNA template attachment

The 5'-thiol oligonucleotide primers (CP3 and CP4), and 5'-thiol template B' were mixed in a saline phosphate buffer (NaPi : 0.1M NaH_2PO_4 pH: 6.6, 0.15M NaCl). Concentration of DNA template varied from 0.001 to 1 μM and from 0.1 to 100 μM for primers but were optimized at 1 μM and 100 μM respectively for template and primers. The procedure described above for CP3 and CP4 attachment on functionalized glass surface was then followed.

The 5'-phosphate oligonucleotide primers (CP1 and CP2), and 5'-phosphate template B were mixed in a 10 mM 1-methyl-imidazole (pH 7.0)(Sigma Chemicals) solution containing 40 mM of 1-ethyl-3-(3-dimethylamino-propyl) carbodiimide (EDC, Pierce, Rockford IL). Concentration of DNA template varied from 0.001 to 10 nM and from 0.1 to 1 μM for primers, but were eventually optimized at 10 nM and 1 μM respectively for template and primers. The procedure described above for CP1 and CP2 attachment on amino-derivatized glass surface was followed.

Hybridization with fluorescent probes- Oligonucleotide probes, fluorescently labeled with Cy5 or FITC at their 5'end, were synthesized by Eurogentec (Brussels). To prevent non-specific hybridization, glass slides were incubated with a blocking solution (5xSSC, Tween 0.1%, BSA 0.1%) for 1 hour and washed on a shaker in 5xSSC (2 times, 5 min). Oligonucleotide probes were diluted at 0.5 μM in 5xSSC, Tween 0.1% and applied on the glass surface for 2 hours at room temperature. Glass slides were rinsed on a shaker at 37°C, once in 5xSSC for 5 min, and twice in 2xSSC containing 0.1% SDS for 5 minutes.

Hybridization with radiolabeled probes- Radiolabeled oligonucleotides complementary to covalently linked oligonucleotides were used as hybridization probes in

order to quantify hybridization yields. Oligonucleotides were enzymatically labeled at their 5' end terminus with [γ-^{32}P]dATP (Amersham, UK) using the bacteriophage T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Excess [γ-^{32}P]dATP was removed with a Chroma Spin column TE-10 (Clontech, Palo Alto CA). Radiolabeled oligonucleotides (0.5 μM in 5xSSC, Tween 0.1%) were then applied onto derivatized slides for 2 hours at room temperature. Glass slides were rinsed on a shaker once in 5xSSC for 5 min and twice in 2xSSC, SDS 0.1% for 5 minutes at 37°C. After hybridization the specific activity was determined by scintillation counting.

Microscope observation - Glass slides were overlaid with 5xSSC solution and a micro cover glass. Fluorescence was detected using an inverted microscope model Axiovert S100TV, with an arc mercury lamp HBO 100W/2 (Carl Zeiss, Oberkochen, Germany) coupled to a CCD camera equipped with a CCD array Kodak with a format 768(H)x512(V) pixels; 6.91x4.6 mm overall, pixel size 9x9 μm2 (Princeton Instruments Inc. Trenton, NJ, USA). Exposition times were between 1 and 50 seconds using the objective LD Achroplan 20x/0.400 (Carl Zeiss, Oberkochen, Germany) and filter sets XF22 (Ex: 485DF22, Dichroic: 505DRLPO2 Em: 530DF30) and XF47 (Ex: 640DF20, Dichroic: 670DRLPO2 Em: 682DF22) from Omega Optical (Brattleboro VT) for FITC and Cy5 fluorophores respectively. Data were analyzed using Winwiew software (Princeton Instruments Inc., Trenton NJ, USA).

Results

Evaluation of storage stability attachment and thermal stability

We evaluated the storage stability of glass plates prepared with s-MBS and s-SIAB. Since these reagents

are sensitive towards hydrolysis, oligonucleotide attachment yields will be dependent on their stability. Amino-derivatized glass plates were functionalized with freshly prepared crosslinking reagents, s-MBS and s-SIAB. The functionalized slides were stored after cross-linking attachment for 10 days in a dessicator under vacuum in the dark at room temperature. After this time, stored slides (t=10days) and freshly reacted slides with the cross-linker reagents (t = 0) were assayed. The results obtained after reaction of a thiol-oligonucleotide and hybridization of a complementary fluorescent probe were compared for both chemistries at t=0 and time= 10 days.

Once immobilized, the s-SIAB-functionalized slides are fully stable after 10 days storage as evidenced by the same yields of hybridization obtained at t=0 and t=10 days. In contrast, coupled s-MBS to glass was found to be less stable with a 30% loss in yield of oligonucleotide attachment and hybridization after 10 days storage. In conclusion, s-SIAB functionalized slides are preferred as they can be prepared in advance and stored dry under vacuum in the dark for at least ten days without any reduction in probe attachment yield.

To evaluate the thermal stability of oligonucleotides attached to glass, the slides were subjected to two 5-min treatments at 100°C in 10 mM Tris-HCl, pH 8. The remaining oligonucleotide still immobilized after washes was assayed by hybridization with a fluorescently labeled complementary oligonucleotide. About 14% of the molecules attached are released for s-SIAB glass slides and 17% for S-MBS glass slides after the first 5 minutes wash, but no further release was detected in the second wash for both chemistries (TABLE 1A). These results are encouraging compared to those obtained by Chrisey et al. 1996, where a release of more than 62% of

oligonucleotides attached on fused silica slides via the crosslinker SMPB (Succinimidyl 4-[p-maleimidophenyl] butyrate) was measured after a 10 min treatment in PBS at 80°C.

TABLE 1A

	Hybridisation results (arbitrary units, normalised to 100%)		
	Freshly attached	After 5 min wash at 100°C	After 2x5 min wash at 100°C
s-MBS	80±6	69±4	73±4
s-SIAB	100±9	84±8	87±3

Table 1A : Thermal stability study

Oligonucleotides were attached to glass slides functionalized with either s-MBS or s-SIAB. Attached oligonucleotides were assayed by hybridization with a fluorescently-labeled complementary oligonucleotide. Fluorescence signal is normalized at 100 for the highest signal obtained. Averaged values of triplicate analyses are reported.

Hybridization as a function of probe attachment

We have studied the extent of hybridization of covalently bound oligonucleotide probes as a function of the surface coverage of attached oligonucleotides using the s-MBS, s-SIAB cross-linkers and EDC-mediated reactions. The concentration of oligonucleotides applied for immobilization was 1 μM for EDC and has been varied between 1 and 800 μM for crosslinkers, the surface density was assayed by hybridization with ^{32}P-labeled probes. The optimal concentration for primer attachment using the heterobifunctional cross-linkers was 500 μM which equates with a surface density of hybridized molecules of 60 fmol/mm^2 for s-MBS and 270

fmol/mm^2 for s-SIAB. Similar coverage density as s-MBS was obtained using EDC/Imidazole-mediated attachment of 5'-phosphate oligonucleotides to aminosilanised glass. However, only 1 μM solutions of oligonucleotide were necessary to attain the same attachment yield, this represents a 500-fold excess of oligonucleotide to be attached for the s-MBS chemistry compared to the EDC/imidazole coupling strategy (Table 1B).

TABLE 1B

	Oligo hybridized (fmol/mm^2)		
Conc. of oligonucleotide used for attachment (μM)	s-MBS	s-SIAB	EDC
1	NT	NT	50
100	10	100	NT
500	60	270	NT

Table 1B : Hybridization as a function of probe attachment

Oligonucleotides were attached to glass slides functionalized with either s-MBS or s-SIAB or via mediated activating reagent EDC. Attached oligonucleotides were assayed by hybridization with a radiolabeled complementary oligonucleotide. The specific activity and therefore the density of hybridized molecules were determined by scintillation liquid. NT : not tested

The 60 fmol/cm^2 surface density corresponds to an average molecular spacing between bound oligonucleotides of 8 nm. According to our results, a coverage density of 30 fmol/mm^2 (spacing of 20 nm) is sufficient to obtain DNA colonies. This yield can be obtained by immobilizing

primers at 100 μM using the heterobifunctional cross-linker s-SIAB or 1μM probes using the EDC-mediated approach. The hybridization densities we have obtained are in the range of the highest densities obtained on glass slides of other grafting protocols previously reported (Guo et al-1994, Joss et al-1997, Beattie et a1-1995).

DNA colony generation on glass: colonies formation is dependent on the length, the concentration of template and the concentration of primers

Theoretically, DNA colony formation requires an appropriate density of primers attached on the surface corresponding to an appropriate length of the DNA template. For optimal DNA colony generation, it is important to define the range of densities of the bound primers and templates, as well as the stoichiometric ratio between template and primer.

MATERIALS AND METHODS

Glass slide preparation

Glass slides were derivatized and functionalized as described above (Materials and methods). DNA colony primers were CP1 and CP2. The colony templates were prepared as described in example 1 for template B', but using primers TPB3 and TPB2. The modified colony primers and templates were applied on glass surface at varying concentrations of both colony primer and colony template.

Generation of colonies

Glass slides stored in 5XSSC were washed in micro-filtered water to removed salts. Colony growing was initiated on glass surface with a PCR mix; the four dNTP (0.2 mM), 0.1% BSA, 0.1% Tween 20, 1X PCR buffer

and 0.05 U/µl of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA). The PCR mix is placed in a frame seal chamber (MJ Research, Watertown, MA). The slides were placed in the thermocycler (The DNA Engine, MJ Research Watertown, MA) and thermocycling was as carried out as follows : step 1 at 94°C for 1 min, step 2 at 65°C for 3 minutes, step 3 at 74°C for 6 min and this program is repeated 50 times. After completion of this program the slides are kept at 6°C until further use.

Digestion of double strand DNA colonies

The glass surface containing the DNA was cut with a restriction nuclease by overlaying with the restriction enzyme in a digestion 1X buffer. The reaction was run twice for 1h30 at 37°C. Double strand DNA colonies were denatured by immersing slides 2 times in tris buffer (10 mM, pH 8) at 100°C for 5 min, followed by a rinse in 5XSSC at 4°C. Slides were stored in 5XSSC for further use.

Hybridization of one strand DNA colonies

To prevent non-specific hybridiztion, glass slides were incubated with a blocking solution (5XSSC, 0.1% Tween, 0.1% BSA) for 1 hour and the slides rinsed in 5XSSC (2 times, 5min). Fluorescently Cy5 5' end labeled oligonucleotide (Eurogentec, Brussels) were diluted at 0.5 µM in SSC 5X, Tween 0.1% and applied to the glass surface for at least 2 hours. Glass slides are rinsed on a shaker once in SSC 5X for 5 min and twice in SSC 5X, SDS 0.1% 5 minutes at 37°C.

The glass slides were visualized as previously described.

We have previously observed that the extent of hybridization is a function of the density of oligonucleotide attachment. A similar study with

bound DNA templates has shown that colony formation is also a function of the concentration of template attached on glass slide. Depending on the chemistry used for oligonucleotide and template attachment, the optimal concentration of template is 1 μM for the bi-functional crosslinkers, s-MBS (Figure 6B), and 1 nM for EDC carbodiimide (Figure 6A). Interestingly, a higher concentration of template does not enhance number of colonies for EDC chemistry and a plateau corresponding to a maximal number of colonies seems to be reached.

We have studied colony formation (number) as a function of the concentration of primers, concentration of the DNA template applied on the surface and the length of the DNA template.

We have also evaluated the number of copy of template in each colony. The quantification was based on fluorescence detection with Cy5-, Cy3- or fluorescein-labeled fluorophores supplemented with an anti-bleaching reagent (Prolong, Molecular Probes, Portland OR). The calibration has been done by hybridization experiments on primers attached to the surface as the exact density corresponding has been determined by radioactivity

EXAMPLE 5: <u>colony in-situ DNA sequencing</u>

Glass slides (5mm diameter Verrerie de Carouge, Switzerland) were first placed into a Helmanex 0.2% (in H_2O), NaOH 1N bath for 1 h at room temperature, rinsed with distilled water, rinsed in pure Hexane, rinsed again two times with distilled water and treated with HCl 1M over night at room temperature. Then, they were rinsed two times in distilled water, and treated with H_2SO_4 (50%) + $K_2S_2O_8$ for 1 h at room temperature. They

were rinsed in distilled water, then two times in Ethanol. Glass slides were derivatized with ATS (as described in example 4).

Colony primers CP1 (5'-pTTTTTTTTTTCACCAACCCAAACCAACCCAAACC) and CP2 (5'-pTTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) which are 5'phosphorylated (Microsynth GmbH, Switzerland) and DNA template B (prepared as described in example 1) were 5' covalently attached onto 5 mm diameter glass slides (Verrerie de Carouge, Switzerland) to a final concentrations of 1 μM and 10 nM respectively, as follows: 2 nmoles of each primer were added to 0.2 nmoles of template in 1 ml of solution A (41 μl of Methylimidazole (Sigma, #M-8878) in 50 ml H_2O, pH adjusted to 7 with HCl) and then mixed 1:1 with solution D (0.2 mM EDC in 10 ml of solution A). On both glass slides sides, 3.5 μl of the mixture were loaded, and incubated over night at room temperature. The glass slides were then briefly rinsed with 5xSSC buffer and placed at 100°C in 10mM Tris buffer pH 8.0 for 2x5'.

Non specific sites on glass were blocked with Bovine Serum Albumin (BSA, Boehringer Mannheim GmbH, Germany, #238040) at 1 mg/ml in 5xSSC buffer for 1 h at room temperature and then rinsed with distilled water.

Glass slides were then individually placed onto a Microamp™ reaction tube (Perkin Elmer) containing 170 μl of PCR mix, and DNA colonies were then generated using Taq polymerase (AmpliTaq, PE-Applied Biosystems Inc., Foster City CA) with 50 cycles (94C/60", 60C/3', 72C/6') in a MTC 200 thermo-cycler (MJ Research, Watertown, MA). Each slide was digested twice using 1.3 units of Pvu II (Stratagene) in NEB 2 buffer (New England Biolabs) for 45 minutes at 37°C. After digestion, the tubes were placed at 100°C in 10mM Tris buffer pH 8.0 for 2x5',

then blocked with filtered (Millex GV4, Millipore) 1 mg/ml BSA in 2xSSC buffer for 30' at room temperature and rinsed first in 2xSSC 0.1% SDS buffer then in 5xSSC buffer. Each slide was incubated over night at room temperature with a 5xSSC/0.1% Tween 20 buffer containing 1 μM of the sequencing primer p181 (CGACAGCCGGAAGGAAGAGGGAGC) overnight at room temperature. Controls without primer were kept in 5xSSC 0.1% Tween 20 buffer. Glass slides were washed 2 times in 5xSSC 0.1% SDS at 37C for 5' and rinsed in 5xSSC. Primer p181 can hybridize to template B' and the sequence following p181 is CAGCT.... In order to facilitate focusing, green fluorescent beads have been adsorbed to the bottom of the well by incubating each well with 20 μl of a 1/2000 dilution of 200 nm yellow/green fluorescent, streptavidin coated FluoSpheres(R) (Molecular Probes, Eugene, OR) in 5X SSC for 20" at room temperature.

After hybridization with the primer, 2 μl of a solution containing 0.1% BSA, 6 mM dithiotreitol (Sigma Chemicals), 5 μM Cy5™-dCTP or 5 μM Cy5™-dUTP (Amersham, UK) and 1X Sequenase reaction buffer is added to each slide. The addition of the Cy5™-nucleotide is initiated with the addition of 1.3 unit of T7 Sequenase™ DNA polymerase (Amersham, UK) for two minutes at room temperature. The wells are washed 2 times in 5X SSC/0.1% SDS bath for 15' and rinsed with 5xSSC buffer.

The samples are observed using an inverted microscope (Axiovert S100TV, Carl Zeiss AG, Oberkochen, Germany) equipped with a Micromax 512x768 CCD camera and Winview software (Princeton Instruments, Trenton, NJ). For focusing, a 20X objective and a XF 22 filter set (Omega Optical, Brattleboro, VT) were used, and for observing Cy5TM incorporation on the samples, a 20X objective and a XF47 filter set (Omega Optical) with a 50 second

exposure using a 2x2 pixel binning. The yellow/green FluoSpheres(R) (approximately 100/field of view) do not give a detectable signal using the XF47 filter set and 50 second exposure (data not shown). The photos are generated by the program, Winview (Princeton Instruments).

Figure 7A shows the result after incubation with Cy5™-dCTP on a sample that has not been incubated with primer p181. One will appreciate only 5 blurry spots can be observed, indicating that no dramatic spurious effect is taking place (such as Cy5™-dCTP aggregate precipitation, adsorption or simply non specific incorporation to the DNA in the colonies or on the surface). Figure 7B shows the result after incubation with Cy5™-dUTP on a sample that has been incubated with primer p181. One will appreciate that no fluorescent spot can be observed, indicating that the incorporation of a fluorescent base cannot take place in detectable amounts when the nucleotide proposed for incorporation does not correspond to the sequence of the template following the hybridized primer. Figure 7C shows the result after incubation with Cy5™-dCTP on a sample that has been incubated with primer p181. One will appreciate that many fluorescent spots can be observed, indicating that the incorporation of a fluorescent base can indeed take place in detectable amounts when the nucleotide proposed for incorporation does correspond to the sequence of the template following the hybridized primer. To summarize, we showed that it is possible to incorporate on a sequence specific manner fluorescent nucleotides into the DNA contained in the colonies and to monitor this incorporation with the apparatus and method described. However, this is only a example. One will appreciate that if desired the incorporation of a fluorescent base could be repeated several times. As this is done on a sequence specific manner, it is thus

possible to deduce part of the sequence of the DNA contained in the colonies.

EXAMPLE 6: 5' mRNA sequence tag analysis

The most accurate way to monitor gene expression in cells or tissues is to reduce the number of steps between the collection of the sample and the scoring of the mRNA. New methods for rapidly isolating mRNA are commercially available. The most efficient methods involve the rapid isolation of the sample and immediate disruption of cells into a solution of guanidinium hydrochloride, which completely disrupts proteins and inactivates RNAses. This is followed by the purification of the mRNA from the supernatant of the disrupted cells by oligo-dT affinity chromatography. Finally, 5'-capped mRNA can be specifically targeted and transformed into cDNA using a simple strategy (SMART cDNA synthesis, Clontech, Palo Alto).

This method allows the synthesis of cDNA copies of only the translationally active, 5'-capped mRNA. By combining the above rapid methods of mRNA isolation and cDNA preparation with the grafted-template method of DNA colony generation described in the present application, we have an approach for the high-throughput identification of a large number of 5' mRNA sequence tags. The advantage of our invention is the possibility to sequence a large number of cDNA by directly grafting the product of the cDNA synthesis reaction, amplifying the cDNA into thousands of copies, followed by the simultaneous in situ sequencing of the cDNAs.

Materials and methods:

Synthetic oligonucleotides and plasmids -

Oligonucleotides were synthesized with 5'-phosphates by Eurogentec or Microsynth. Plasmids containing partial coding and 3'-untranslated sequences of the murine

potassium channel gene, mSlo, following the T3 RNA polymerase promoter were generated by standard methods.

mRNA synthesis - mSlo plasmids were linearized at a single SalI or SacI restriction nuclease site following the poly A+ sequence in the plasmid. After treatment of the cut plasmid with proteinase K, linear plasmid DNA was extracted once with phenol/CH_3Cl/isoamyl alcohol and precipitated with ethanol. The DNA precipitate was re-dissolved in H_20 at a concentration of 10 $\mu g/\mu l$. Synthetic mRNA capped with the 5'-methylguanosine were synthesized by the mMessage mMachine in vitro mRNA synthesis kit as per manufacturer instructions (Ambion, Austin TX). Synthetic mRNA was stored at 80°C.

Enzymes - Restriction enzymes were obtained from New England Biolabs (Beverly, MA).

cDNA synthesis - Synthetic mRNA was mixed with mouse liver poly A+ mRNA at different molar ratios (1:1, 1:10, 1:100) and cDNA synthesis on the mixture of synthetic and mouse liver mRNA was performed using the "SMART PCR cDNA synthesis kit" (Clontech, Palo Alto CA) with some minor modifications. In a cDNA reaction, approximately 1 μg of the mRNA mixture was mixed with the -primer CP5, having at the 5' -end the sequence of CPβ, (5'p-AGAAGGAGAAGGAAAGGGAAAGGGTTTTTTTTTTTTTTTNN). This primer has been used to make the 1st strand cDNA synthesis. For the 2nd strand synthesis, the "SMART" technique has been used. The basis of the SMART synthesis is the property of the Moloney murine viral reverse transcriptase to add three to five deoxycytosine residues at the 3'-termini of first strand cDNA, when the mRNA contains a 5'-methylguanosine-cap (SMART user manual, Clontech, Palo Alto CA). A CP6 primer, which contains the sequence of CPβ plus AAAGGGGG at the 3' end, (5'p-AGAAGGAGAAGGAAAGGGAAAGGGGG) has been used for

the 2nd strand cDNA synthesis. Buffer and SUPERSCRIPT™ II RNase H- reverse transcriptase from Moloney murine leukemia virus (Life Technologies, Ltd.) were used as described in the instructions and the reaction was carried out at 42°C for 1 hr. The cDNA was assayed by PCR using the primer p251, which contains a fragment of the CPβ sequence, (5'-GAGAAGGAAAGGGAAAGG) with Taq DNA polymerase (Platinum Taq, Life Technologies, Ltd.).

Preparation of DNA colonies - The 5'p-cDNA was mixed with different concentrations of the solid phase colony primer, CP2 (5'p-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) and chemically bound to Nucleolink PCR tubes (NUNC) following manufacturer instructions. DNA colonies were then generated using Taq polymerase (AmpliTaq Gold, PE-Applied Biosystems Inc., Foster City CA) with 30 cycles (94C/30", 65C/1', 72C/ 1.5') in a MTC 200 thermo-cycler (MJ Research, Watertown, MA).

DNA probes and hybridization - ^{32}Biotinylated and ^{32}P-radiolabelled DNA probes specific for the mSlo DNA sequence were synthesized with a 5'-biotinylated primer and a normal downstream primer by PCR on the template (mSlo plasmid DNA). The probe incorporated α[^{32}P]-dCTP (Amersham, Amersham UK) at a ratio of 300:1 (α[^{32}P]-dCTP to dCTP) in the PCR reaction, with a final concentration of the four deoxynucleoside triphosphates of 50 μM. The resulting biotinylated and radiolabelled DNA probe was desalted over a Chromaspin-1000 column (Clontech, Palo Alto CA). The DNA probes were hybridized to the samples in "easyhyb" buffer (Boehringer-Mannheim, Germany), using the following temperature scheme (in the MTC200 thermocycler): 94°C for 5 minutes, followed by 68 steps of 0.5°C decrease in temperature every 30 seconds (in other words, the temperature is decreased down to 60°C in 34 minutes), using sealed wells. The samples are then washed 3 times with 200 μl of TNT at room

temperature. The wells are then incubated for 30 minutes with 50 µl TNT containing 0.1 mg/ml BSA (New England Biolabs, Beverly, MA). Then the wells are incubated 5 minutes with 15 µl of solution of red fluorescent, steptavidin-coated, 40 nanometer microspheres (Molecular Probes, Portland, OR). The solution of microspheres is made of 2 µl of the stock solution of microspheres, which have been sonicated for 5 minutes in a 50 W ultra-sound water-bath (Elgasonic, Bienne, Switzerland), diluted in 1 ml of TNT solution containing 0.1 mg/ml BSA and filtered with Millex GV4 0.22 µm pore size filter (Millipore, Bedford, MA).
DNA colony visualization - The stained samples are observed using an inverted Axiovert 10 microscope using a 20X objective (Carl Zeiss AG, Oberkochen, Germany) equipped with a Micromax 512x768 CCD camera (Princeton instruments, Trenton, NJ), using a PB546/FT580/LP590 filter set, and 10 seconds of light collection. The files are converted to TIFF format and processed in the suitable software (PhotoPaint, Corel Corp. Ltd, Dublin, Ireland). The processing consisted in inversion and linear contrast enhancement, in order to provide a picture suitable for black and white printout on a laser printer.

RESULTS

Synthetic mRNA and cDNA synthesis - Following cDNA synthesis, the cDNA was checked in a PCR using the p251 primer (generated at each end of the first strand cDNA) for the correct lengths of products as assayed by agarose gel electrophoresis. The synthetic mSlo mRNA was diluted into the liver mRNA, which was evidenced by the decreasing intensity of the mSlo-specific band and the increase of a non-specific smear of liver cDNA.

Detection and quantification of DNA colonies - DNA colonies were assayed using fluorescent imaging CCD microscopy or scintillation counting. The numbers of fluorescently detectable colonies increased as a function of the amount of grafted template, as shown in Figure 6. This increase was mirrored by the amount of ^{32}P-radiolabel detected.

With radiolabelled probes it is possible to detect mRNA copies at about 1:100. But with fluorescent microscopic CCD imaging technology, one can detect mRNA to a dilution of 1:10000.

EXAMPLE 7 : Covalent binding of primer to the solid support (plastic)

Oligonucleotide primers were attached onto Nucleolink plastic microtitre wells (Nunc, Denmark) in order to determine optimal coupling times and chemistries. Oligonucleotides; CP2 (5'-(phosphate)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), CP8(5'-(amino-hexamethylene)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), CP9 (5'(hydroxyl)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), CP10 (5'-(dimethoxytrityl)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG) and CP11 (5'(biotin)-TTTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG), (Microsynth GmbH, Switzerland), were attached to Nucleolink microtitre wells as follows (8 wells each); to each well 20 μl of a solution containing 0.1 μM oligonucleotide, 10mM 1-methyl-imidazole (pH 7.0) (Sigma Chemicals) and 10mM 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (pH 7.0) (Sigma Chemicals) in 10mM 1-methyl-imidazole. The wells were then sealed and incubated 50°C for varying amounts of time. The coupling reaction was terminated at specific times by rinsing twice with 200 μl of RS (0.4 N NaOH, 0.25% Tween 20) and twice with 200 μl TNT (100mM TrisHCl pH 7.5, 150 mM NaCl, 0.1% Tween 20). Tubes were dried at 50°C for

30' and were stored in a sealed plastic bag at 4°C.

Stability of bound oligonucleotides under PCR Colony generation conditions

Stability was tested under colony growing conditions by adding a PCR mix (20 µl of four dNTPs (0.2 mM), 0.1% BSA, 0.1% Tween 20, 8% DMSO (dimethylsulfoxide, Fluka, Switzerland), IX PCR buffer). The wells were then placed in the thermocycler and for 33 repetitions under the following conditions: 94°C for 45 seconds, 60°C for 4 minutes, 72°C for 4 minutes. After completion of this program, the wells were rinsed with 5xSSC, 0.1% Tween 20 and kept at 8°C until further use. Prior to hybridization wells are filled with 50 µl 5xSSC, 0.1% Tween 20 heated at 94°C for 5 minutes and stored at RT. Probe: Oligonucleotide probes, R57 (5'(phosphate)-GTTTGGGTTGGTTTGGGTTGGTG, control probe) and R58 (5'-(phosphate)-CCCTTTCCCTTTCCTTCTCCTTCT, which is complementary to CP2, CP8, CP9, CP10 and CP11) were enzymatically labeled at their 5' end terminus with [γ-^{32}P]dATP (Amersham, UK) using the bacteriophage T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Excess ^{32}P dATP was removed with a Chroma Spin column TE-10 (Clontech, Palo Alto CA). Radiolabeled oligonucleotides (0.5 µM in 5xSSC, 0.1% Tween 20) were then hybridized to the oligonucleotide derivatized Nucleolink wells at 37°C for two hours. The wells were washed 4 times with 5xSSC, 0.1% Tween 20 at room temperature, followed by a wash with 0.5xSSC, 0.1% Tween 20 for 15' at 37°C. Wells were then assayed for bound probe by scintillation counting.

Results

There is a marked difference in the rate and specificity of oligonucleotide coupling depending on the nature of

5'-functional group on the oligonucleotide. Oligonucleotides carrying the 5'-amino group coupled approximately twice as fast as oligonucleotides functionalized with a 5'-phosphate group (see Table 2 and Figure 8). In addition, the control oligonucleotides functionalized with 5'hydroxyl, 5'-DMT or 5'-biotin all coupled at rates similar to that of the 5'-phosphate, which questions the 5'specific nature of the chemical attachment using the 5'-phosphate group.

Table 2

	5'-phosphate	5'-amino	5'-hydroxyl	5'-DMT	5'-biotin
Ka(min-1)	0.0068	0.0135	0.0063	0.0070	0.0068
Attached oligo-nucleotide (fmol/well)	608	1344	542	602	650
PCR stability (% remaining)	56	69	66	66	62

IDT Exhibit 1020.9569

Claims:

1. A method for amplification of at least one nucleic acid comprising the following steps:-

(1) forming at least one nucleic acid template comprising the nucleic acid(s) to be amplified, wherein said nucleic acid(s) contains at the 5' end an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support;

(2) mixing said nucleic acid template(s) with one or more colony primers X, which can hybridize to the oligonucleotide sequence Z and carries at the 5' end a means for attaching the colony primers to a solid support, in the presence of a solid support so that the 5' ends of both the nucleic acid template and the colony primers bind to the solid support;

(3) performing one or more nucleic acid amplification reactions on the bound template(s), so that nucleic acid colonies are generated.

2. A method as claimed in claim 1, wherein the oligonucleotide sequence Z is complementary to oligonucleotide sequence Y and colony primer X is of the same sequence as oligonucleotide sequence Y.

3. A method as claimed in claim 1, wherein two different colony primers X are mixed with said template(s) in step (2), and wherein the sequences of colony primers X are such that the oligonucleotide sequence Z can hybridise to one of the colony primers X and the oligonucleotide sequence Y is the same as one of the colony primers X.

4. A method as claimed in claim 1, wherein the colony primers X comprise a degenerate primer sequence and the

nucleic acid template(s) do not contain oligonucleotide sequences Y or Z.

5. A method as claimed in any one of claims 1 to 4 comprising the additional step of performing at least one step of sequence determination of one or more of the nucleic acid colonies generated.

6. A method as claimed in claim 5 wherein the sequence determination step(5) involves the incorporation and detection of labelled oligonucleotides.

7. A method as claimed in claim 5 or 6 wherein the full or partial sequences of the amplified nucleic acid templates present in more than one nucleic acid colonies are determined simultaneously.

8. A method as claimed in any one of claims 1 to 7 comprising the additional step of visualising the colonies generated.

9. A method as claimed in claim 8 wherein said visualisation step involves the use of a labelled or unlabelled nucleic acid probe.

10. A method as claimed in any one of claims 1 to 9, wherein the means for attaching the nucleic acid template(s) and the colony primers to the solid support comprises a means for attaching the nucleic acid sequences covalently to the said support.

11. A method as claimed in claim 10, wherein said means for attaching the nucleic acid sequences covalently to the solid support is a chemically modifiable functional group.

12. A method as claimed in claim 11, wherein said chemically modifiable functional group is a phosphate group, a carboxylic or aldehyde moiety, a thiol, a hydroxyl, a dimethoxytrityl (DMT), or an amino group.

13. A method as claimed in claim 12, wherein said chemically modifiable functional group is an amino group.

14. A method as claimed in any one of claims 1 to 13 wherein said solid support is selected from the group comprising latex beads, dextran beads, polystyrene, polypropylene surface, polyacrylamide gel, gold surfaces, glass surfaces and silicon wafers.

15. A method as claimed in claim 14, wherein said solid support is glass.

16. A method as claimed in any one of claims 1 to 15 wherein the density of nucleic acid colonies generated is 10,000/mm^2 to 100,000/mm^2.

17. A method as claimed in any one of claims 1 to 16 wherein the density of colony primers X attached to said solid support is at least 1 fmol/mm^2.

18. A method as claimed in any one of claims 1 to 17 wherein the density of nucleic acid templates is 10,000/mm^2 to 100,000/mm^2 .

19. A plurality of nucleic acid templates comprising the nucleic acids to be amplified, wherein said nucleic acids contain at their 5' ends an oligonucleotide sequence Y and at the 3' end an oligonucleotide sequence Z and, in addition, the nucleic acid(s) carry at the 5' end a means for attaching the nucleic acid(s) to a solid support.

20. A plurality of nucleic acid templates as claimed in claim 19, wherein said templates are mixed with a plurality of colony primers X.

21. A plurality of nucleic acid templates as claimed in claim 19 or claim 20, wherein the oligonucleotide sequence Z is complementary to oligonucleotide sequence Y and colony primer X is of the same sequence as oligonucleotide sequence Y.

22. A plurality of nucleic acid templates as claimed in claim 19 or claim 20 wherein two different colony primers X are mixed with said templates, and wherein the sequences of colony primers X are such that the oligonucleotide sequence Z can hybridise to one of the colony primers X and the oligonucleotide sequence Y is the same as one of the colony primers X.

23. A plurality of nucleic acid templates as claimed in claim 19 or claim 20 wherein the colony primers X comprise a degenerate primer sequence and the nucleic acid templates do not contain oligonucleotide sequences Y or Z.

24. A solid support, to which there is attached a plurality of colony primers X and at least one nucleic acid template, wherein said nucleic acid templates and colony primers are as defined in any one of the previous claims.

25. A solid support as claimed in claim 24 wherein a plurality of nucleic acid templates are attached to said solid support.

26. A solid support as claimed in claim 24 or 25 wherein the solid support is as defined in claims 14 and 15.

27. A solid support as claimed in any one of claims 24 to 26 wherein the attachment of nucleic acid templates and colony primers to the solid support is covalent.

28. A solid support comprising one or more nucleic acid colonies generated by a method as defined in any one of claims 1 to 18.

29. Use of the solid support of any one of claims 24 to 28 in methods of nucleic acid amplification or sequencing.

30. Use as claimed in claim 29 wherein said method is a method as claimed in any one of claims 1 to 18.

31. Use of a method of any one of claims 1 to 18, for nucleic acid amplification or sequencing.

32. Use of a method as defined in any one of claims 1 to 18 or the nucleic acid colonies generated by said methods, or the plurality of nucleic acid templates of claims 19 to 23, or the solid supports of claims 24 to 28, for providing nucleic acid molecules for sequencing and re-sequencing, gene expression monitoring, genetic diversity profiling, diagnosis, screening, whole genome sequencing, whole genome polymorphism discovery and scoring and the preparation of whole genome slides, or any other applications involving the amplification of nucleic acids or the sequencing thereof.

33. A kit for use in nucleic acid amplification or sequencing comprising a plurality of nucleic acid templates and colony primers as defined in any of the preceding claims bound to a solid support.

34. A kit as claimed in claim 33 for use in sequencing, re-sequencing, gene expression monitoring, genetic

diversity profiling, diagnosis, screening, whole genome sequencing, whole genome polymorphism discovery and scoring, or any other applications involving the amplification of nucleic acids or the sequencing thereof.

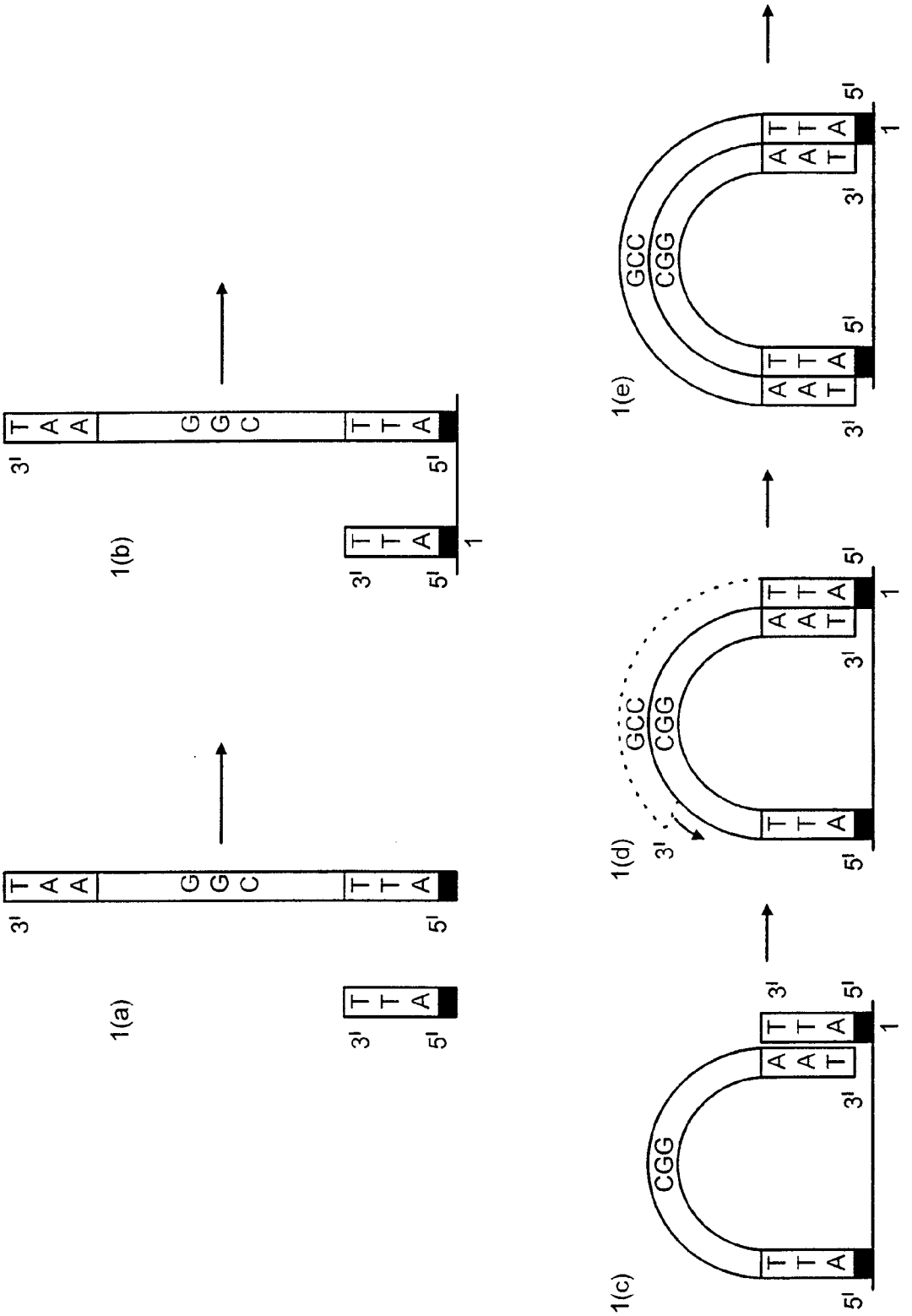
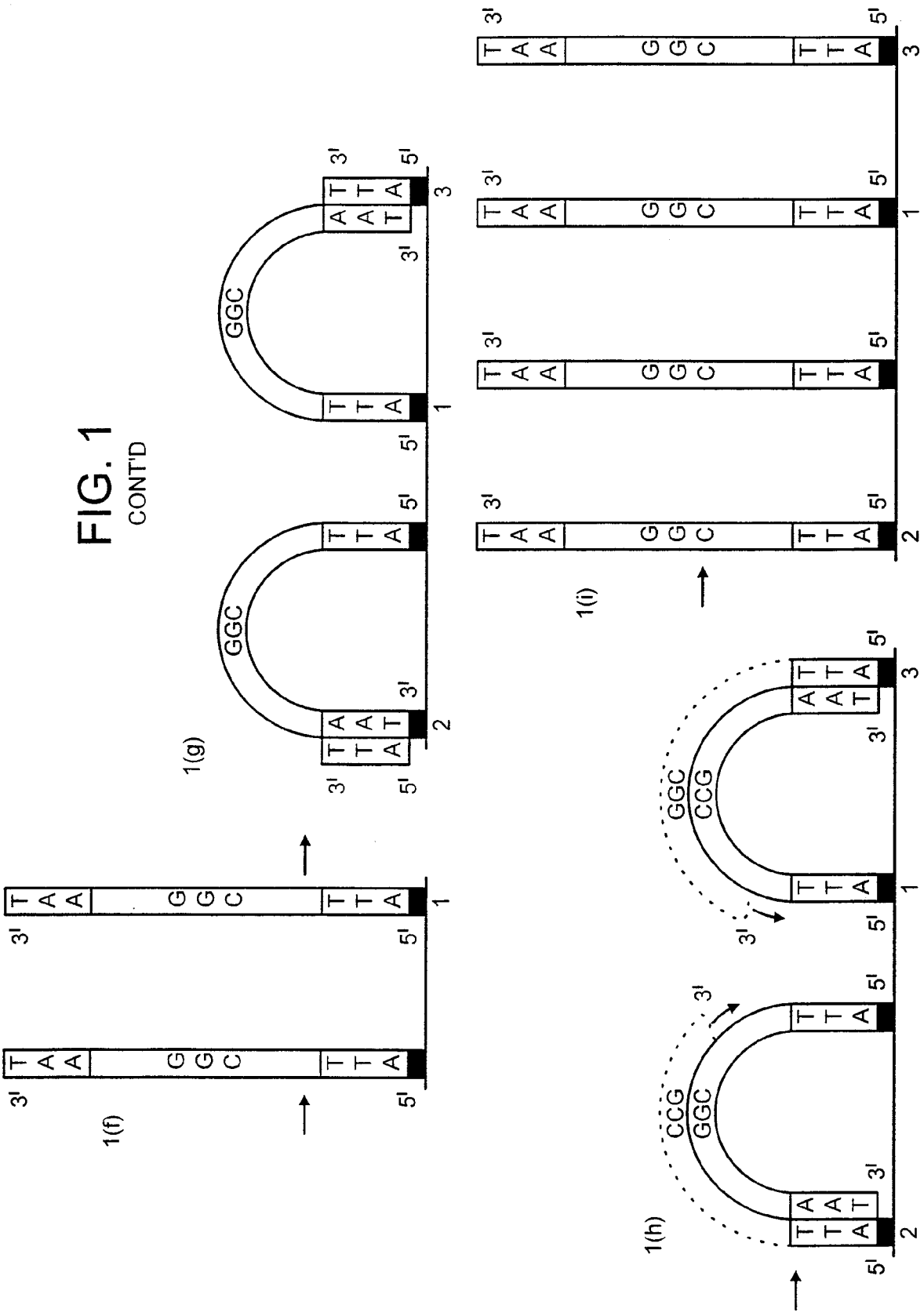


FIG. 1

IDT Exhibit 1020.9576

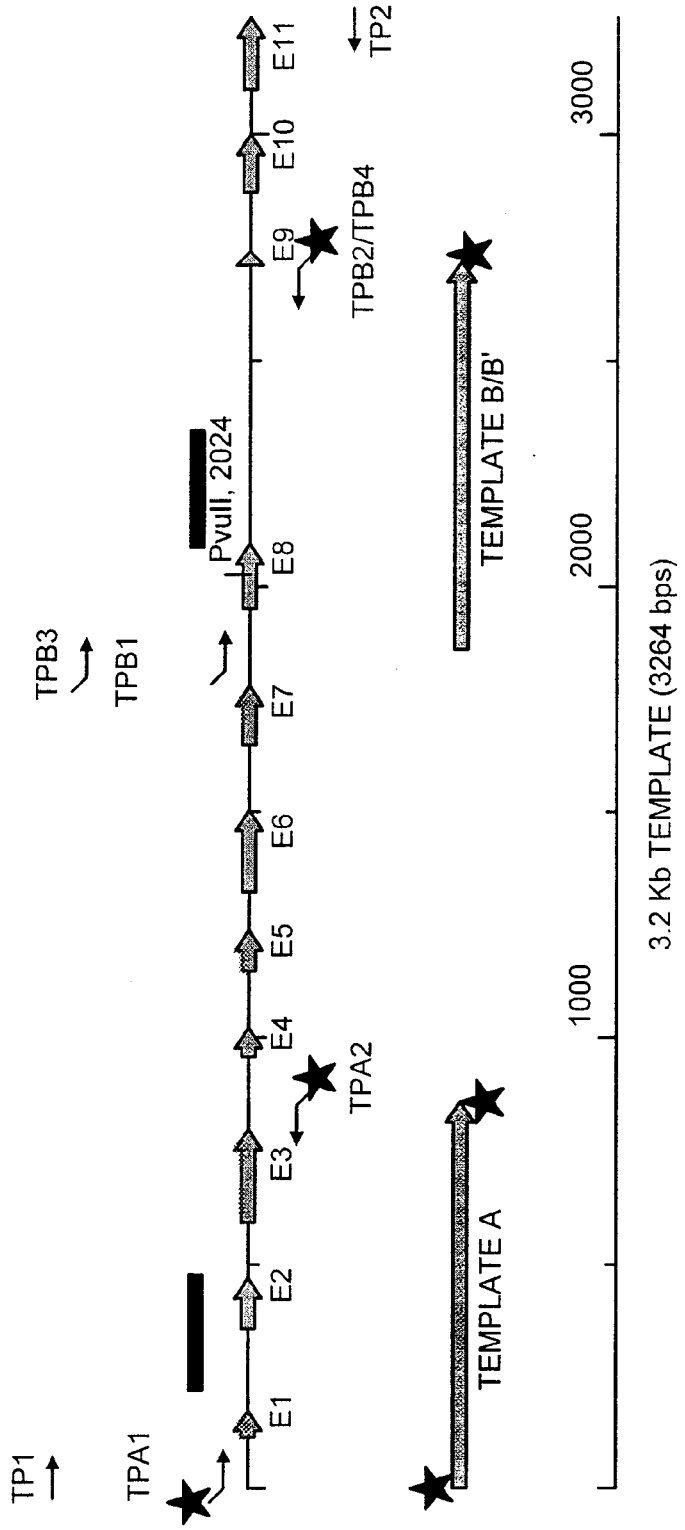
FIG. 1
CONT'D



IDT Exhibit 1020.9577

FIG. 2a

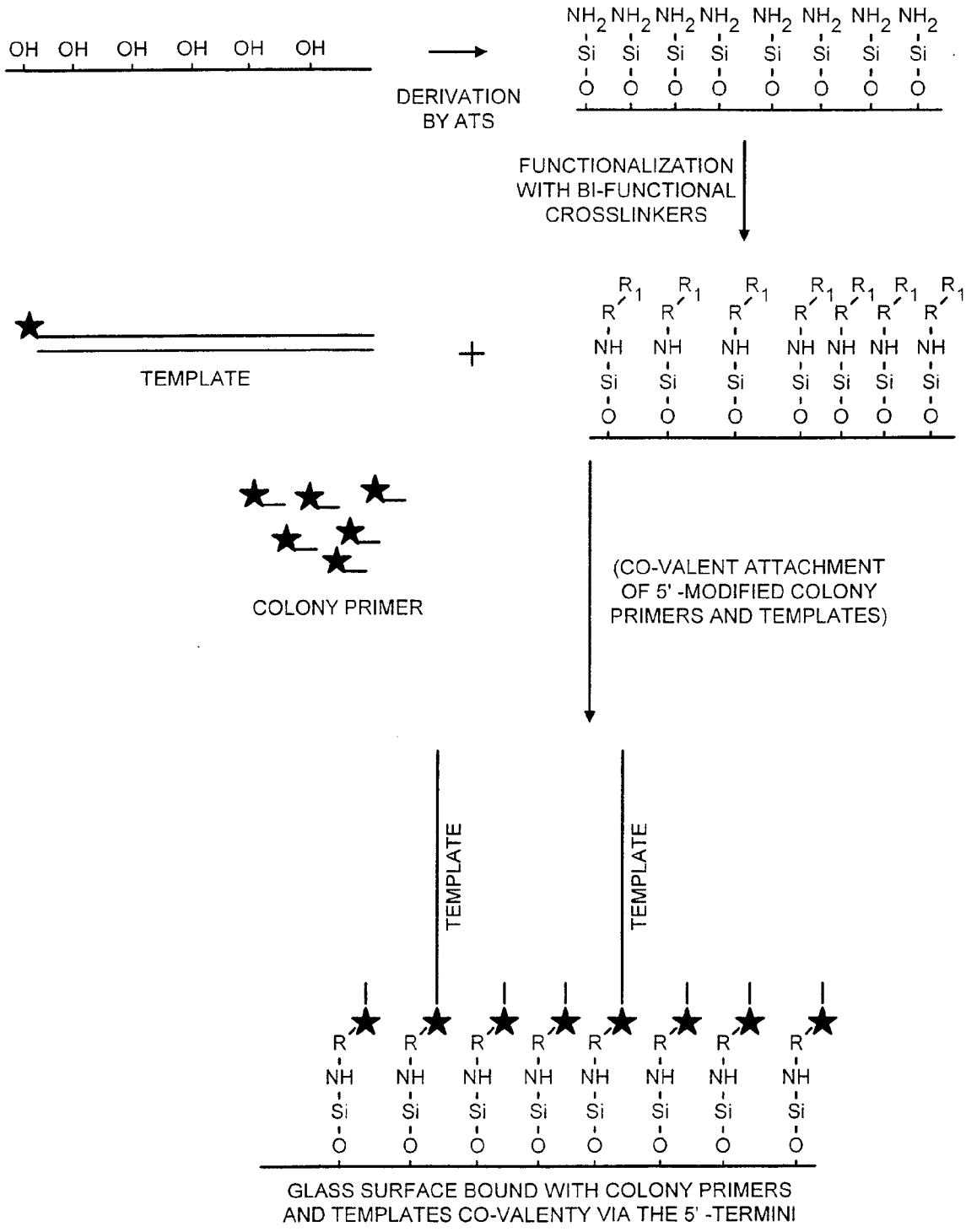
PREPARATION OF TEMPLATES CONTAINING COLONY PRIMER SEQUENCES
(LOCATION OF PRIMERS ON RAGE)



IDT Exhibit 1020.9578

FIG. 2b

CHEMICAL ATTACHMENT OF COLONY PRIMERS AND COLONY TEMPLATES TO THE SURFACE



IDT Exhibit 1020.9579

FIG. 3

COLONIES FORMED WITH TEMPLATE AND COLONY PRIMER ATTACHED TO THE SUPPORT

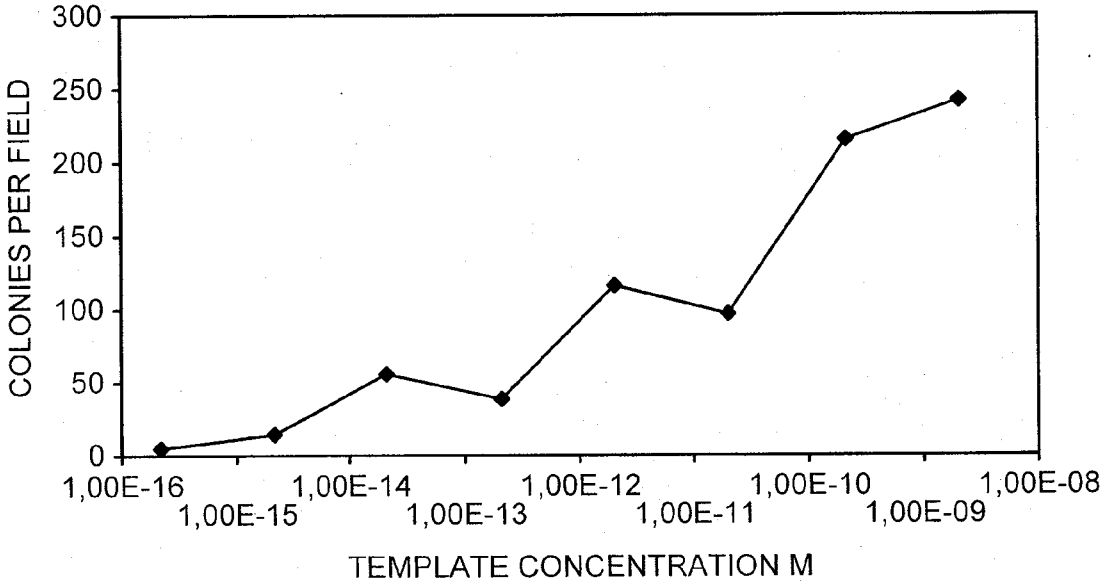
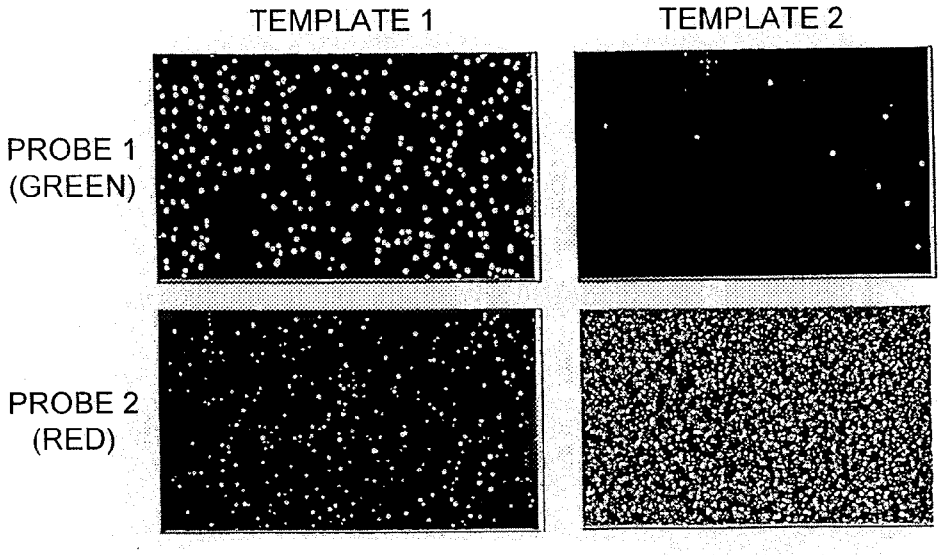


FIG. 4a

SIMULTANEOUS DETECTION OF TWO INDEPENDENT DNA COLONIES



IDT Exhibit 1020.9580

FIG. 4b

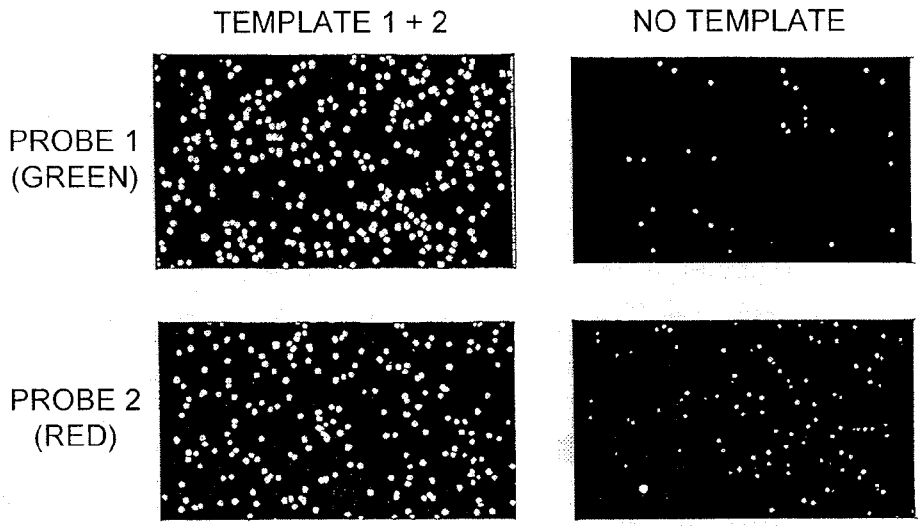
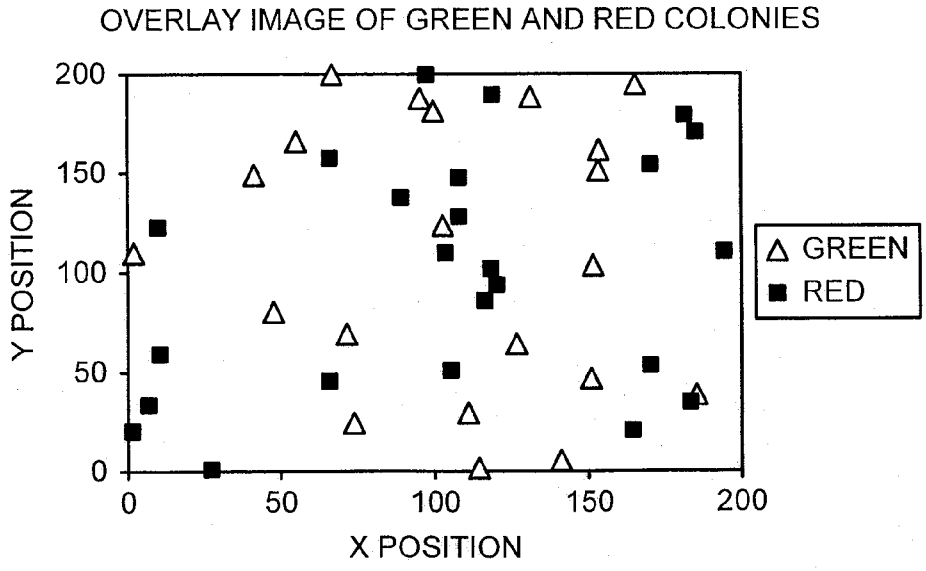


FIG. 4c



IDT Exhibit 1020.9581

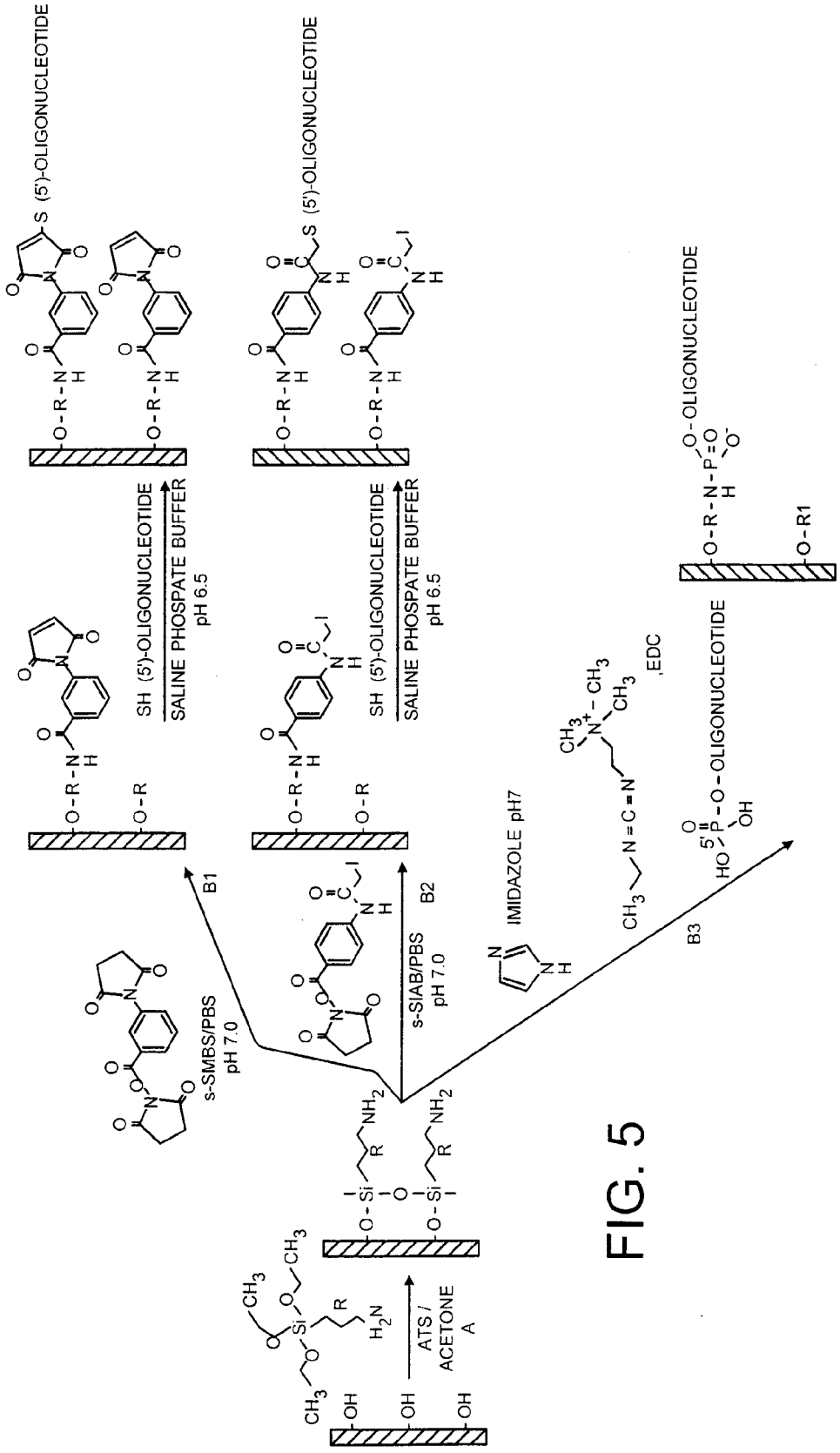
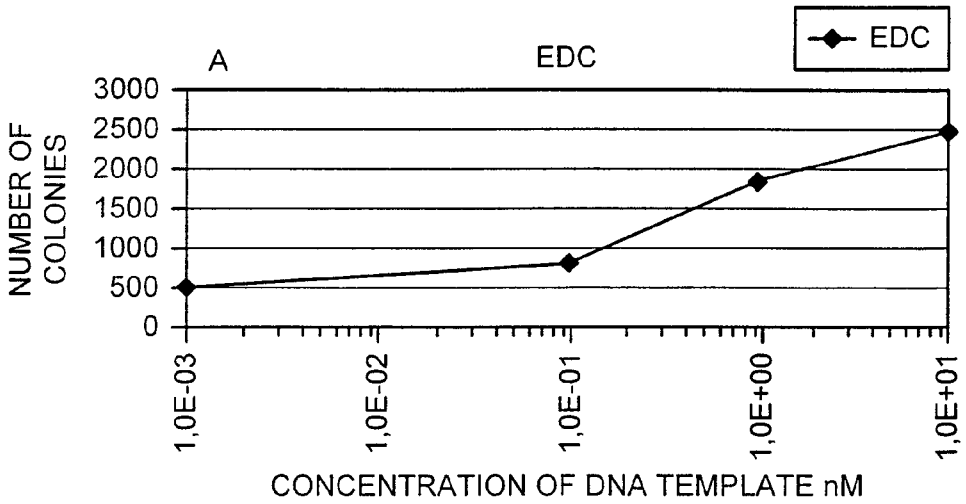


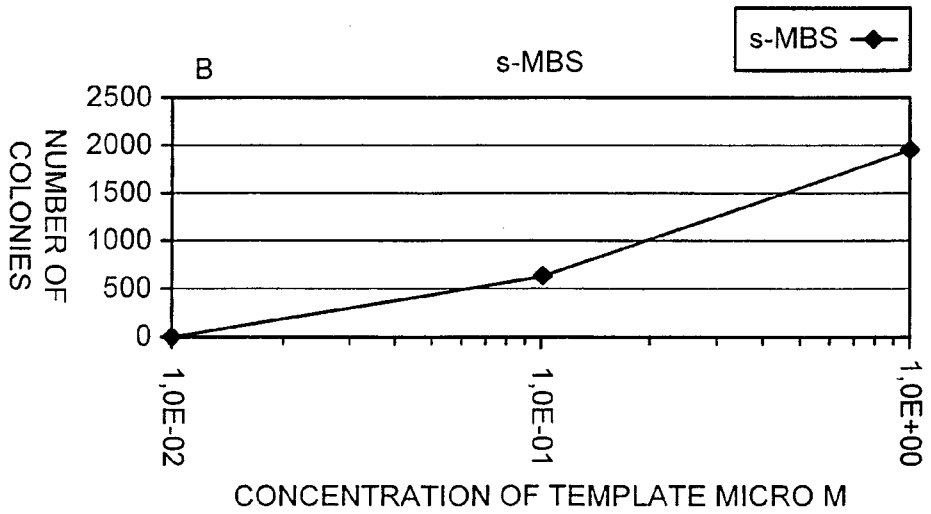
FIG. 5

IDT Exhibit 1020.9582

FIG. 6

DNA COLONY GENERATION ON GLASS:
EFFECT OF TEMPLATE CONCENTRATION





IDT Exhibit 1020.9583

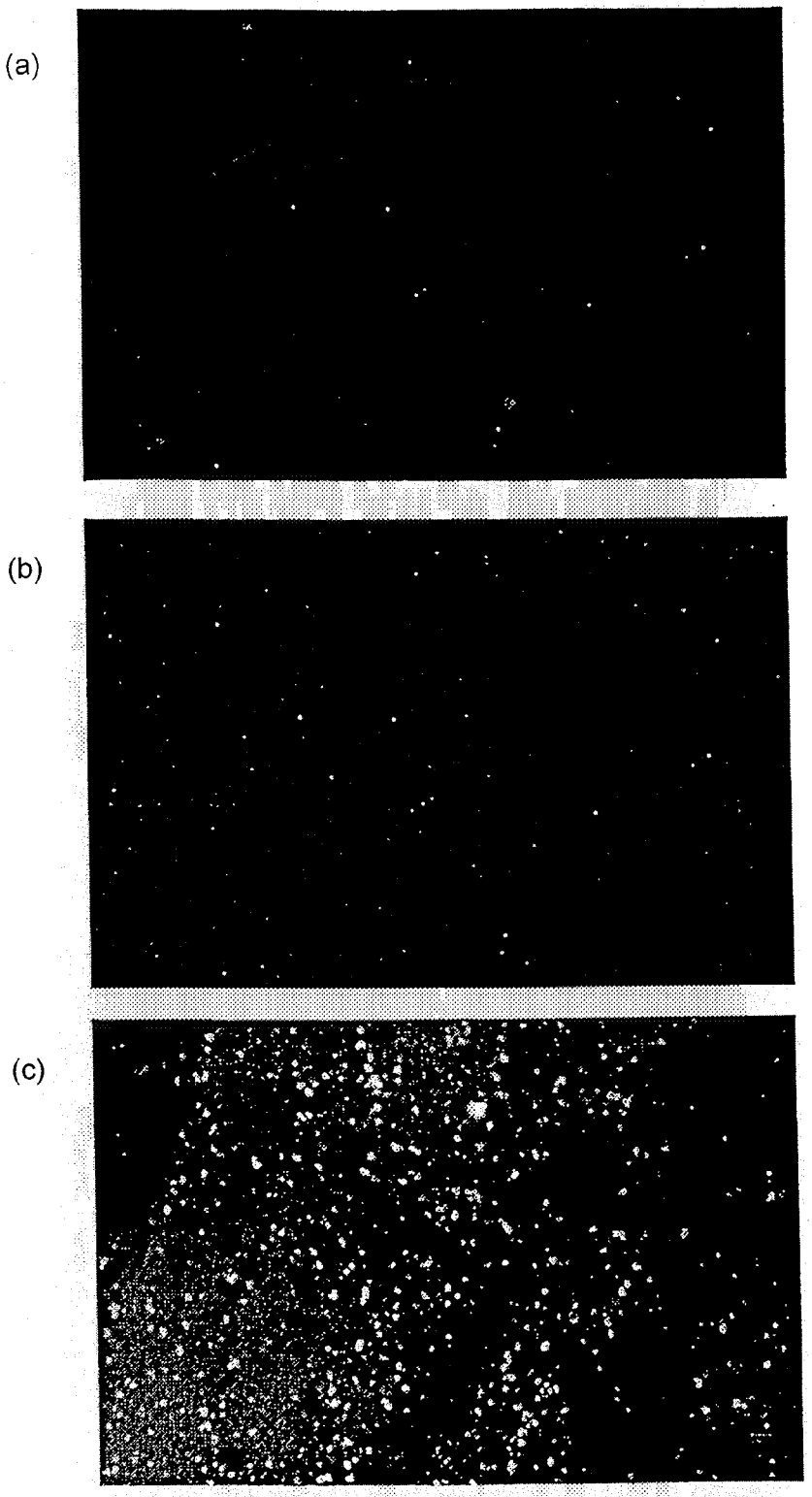
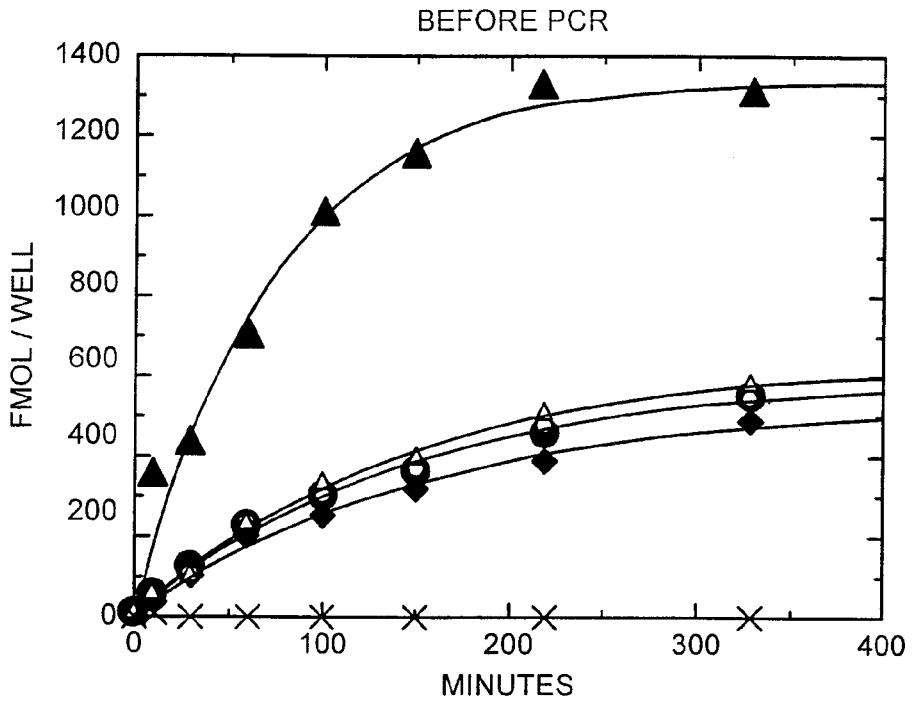
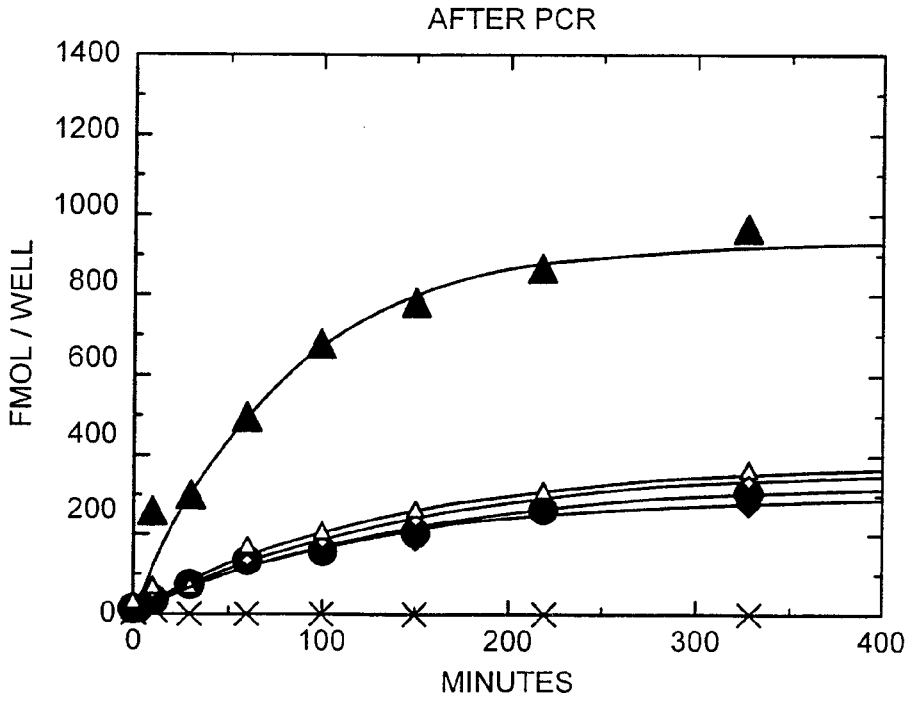


FIG. 7

IDT Exhibit 1020.9584

FIG. 8





IDT Exhibit 1020.9585

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Applied Research Systems ARS Holding N.V.
(B) STREET: 14 John B. Gorsiraweg
(C) CITY: Curacao
(E) COUNTRY: Netherlands Antilles
(F) POSTAL CODE (ZIP): None

(ii) TITLE OF INVENTION: Methods of nucleic acid amplification and sequencing

(iii) NUMBER OF SEQUENCES: 19

(iv) 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 #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AGAAGGAGAA GGAAAGGGAA AGGG 24

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

IDT Exhibit 1020.9586

CACCAACCCA AACCAACCCA AACC 24

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAGGCCAGAA CAGTTCAAGG 20

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCTGTGACAA GACGACTGAA 20

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTTTTTTTTT CACCAACCCA AACCAACCCA AACC 34

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TTTTTTTTTT AGAAGGAGAA GGAAAGGGAA AGGG 34

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TTTTTTTTTT CACCAACCCA AACCAACCCA AACC 34

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTTTTTTTTT AGAAGGAGAA GGAAAGGGAA AGGG 34

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AGAAGGAGAA GGAAAGGGAA AGGGTTTTTT TTTTTTTTTT NN 42

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGAAGGAGAA GGAAAGGGAA AGGGGG 26

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AGAAGGAGAA GGAAAGGGAA AGGGGCGGCC GCTCGCCTGG TTCTGGAAGA CA 52

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

IDT Exhibit 1020.9589

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AGAAGGAGAA GGAAAGGGAA AGGGCCTGTG ACAAGACGAC TGAA 44

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTTTTTTTTT AGAAGGAGAA GGAAAGGGAA AGGGGCGGCC GCTGAGGCCA GTGGAAGTCA 60

GA 62

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TTTTTTTTTT CACCAACCCA AACCAACCCA AACCGAGCTC AGGCTGAGGC AGGAGAATTG 60

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AGAAGGAGAA GGAAAGGGAA AGGGGAGCTG AGGAGGAAGA GAGG 44

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AGAAGGAGAA GGAAAGGGAA AGGGGCGGCC GCTCGCCTGG TTCTGGAAGA CA 52

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:11
(D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:13
(D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:15
(D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:17
(D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:19

(D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:21
(D) OTHER INFORMATION:/mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TTTTTTTTTT NSNSNSNSNS NS 22

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CGACAGCCGG AAGGAAGAGG GAGC 24

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GAGAAGGAAA GGGAAAGG 18

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/GB 99/03248

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68

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)
IPC 7 C12Q

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 practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 98 44151 A (FARINELLI LAURENT ;KAWASHIMA ERIC (CH); MAYER PASCAL (CH); GLAXO G) 8 October 1998 (1998-10-08) the whole document ---	1,5,19, 24-31
E	WO 98 44152 A (FARINELLI LAURENT ;MAYER PASCAL (CH); KAWASHIMA ERIC H (CH); GLAXO) 8 October 1998 (1998-10-08) the whole document ---	1,5,19, 24-31
X	WO 98 36094 A (MOSAIC TECHNOLOGIES) 20 August 1998 (1998-08-20) cited in the application page 11, line 27 - line 30; figure 2 ---	1,5,19, 24-31

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° 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 document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step 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
11 February 2000	21/02/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Osborne, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

IDT Exhibit 1020.9593

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/GB 99/03248

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 04404 A (MOSAIC TECHNOLOGIES INC ;WHITEHEAD FOR BIOMEDICAL RESEA (US); ADAM) 15 February 1996 (1996-02-15) the whole document	1,5,19, 24-31
A	WO 95 33073 A (TEPNEL MEDICAL LTD ;MINTER STEPHEN JOHN (GB)) 7 December 1995 (1995-12-07) the whole document	1,5, 24-31
A	WO 94 02634 A (UNIV SOUTH AUSTRALIA ;ADELAIDE CHILDREN S HOSPITAL (AU); HARRIS RA) 3 February 1994 (1994-02-03) claim 1	1
X	EP 0 665 293 A (HAMAMATSU PHOTONICS KK) 2 August 1995 (1995-08-02) cited in the application	19
A	the whole document	1
A	STERKY F ET AL: "Direct sequencing of bacterial artificial chromosomes (BACs) and prokaryotic genomes by biotin-capture PCR" JOURNAL OF BIOTECHNOLOGY,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 60, no. 1-2, 5 February 1998 (1998-02-05), pages 119-129, XP004115549 ISSN: 0168-1656 the whole document	1,4-7
A	EP 0 543 484 A (JAPAN RES DEV CORP ;YOKOI HARUHIKO (JP)) 26 May 1993 (1993-05-26) see abstract	

1

IDT Exhibit 1020.9594

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte tional Application No
PCT/GB 99/03248

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9844151 A	08-10-1998	AU 6846698 A	22-10-1998
		EP 0972081 A	19-01-2000
		AU 6846798 A	22-10-1998
		EP 0975802 A	02-02-2000
		WO 9844152 A	08-10-1998
WO 9844152 A	08-10-1998	AU 6846698 A	22-10-1998
		AU 6846798 A	22-10-1998
		EP 0972081 A	19-01-2000
		EP 0975802 A	02-02-2000
		WO 9844151 A	08-10-1998
WO 9836094 A	20-08-1998	AU 6146098 A	08-09-1998
WO 9604404 A	15-02-1996	US 5641658 A	24-06-1997
		CA 2196604 A	15-02-1996
		EP 0784701 A	23-07-1997
		JP 10505492 T	02-06-1998
WO 9533073 A	07-12-1995	AU 707391 B	08-07-1999
		AU 2572895 A	21-12-1995
		CA 2191426 A	07-12-1995
		CN 1152946 A	25-06-1997
		EP 0763135 A	19-03-1997
		FI 964735 A	27-01-1997
		JP 10500862 T	27-01-1998
		NZ 287118 A	28-01-1999
		US 5922574 A	13-07-1999
		ZA 9504365 A	02-04-1996
WO 9402634 A	03-02-1994	AU 698934 B	12-11-1998
		AU 4551193 A	14-02-1994
		CA 2140877 A	03-02-1994
		DE 69327326 D	20-01-2000
		EP 0656068 A	07-06-1995
		US 5849544 A	15-12-1998
EP 0665293 A	02-08-1995	JP 7203998 A	08-08-1995
EP 0543484 A	26-05-1993	CA 2077135 A	01-03-1993
		JP 7255482 A	09-10-1995
		US 5665572 A	09-09-1997

IDT Exhibit 1020.9595

Espacenet

Bibliographic data: EP0843735 (A1) — 1998-05-27

GLYCOSYLASE MEDIATED DETECTION OF NUCLEOTIDE SEQUENCES AT CANDIDATE LOCI

Inventor(s): MCCARTHY THOMAS VALENTINE [IE]; VAUGHAN PATRICK MARTIN [IE] ± (MCCARTHY, THOMAS, VALENTINE, ; VAUGHAN, PATRICK, MARTIN)

Applicant(s): FORFAS TRADING AS BIORESEARCH [IE]; UNIV CORK [IE] ± (FORFAS (TRADING AS BIORESEARCH IRELAND), ; UNIVERSITY COLLEGE, CORK, ; UNIVERSITY COLLEGE CORK-NATIONAL UNIVERSITY OF IRELAND, CORK)

Classification: - international: ***C12N15/09; C12Q1/68;*** (IPC1-7): C12Q1/68
- cooperative: **C12Q1/6827; C12Q1/6834; C12Q1/6853; C12Q1/689;** C12Q2600/156

Application number: EP19950942776 19951221 Global Dossier

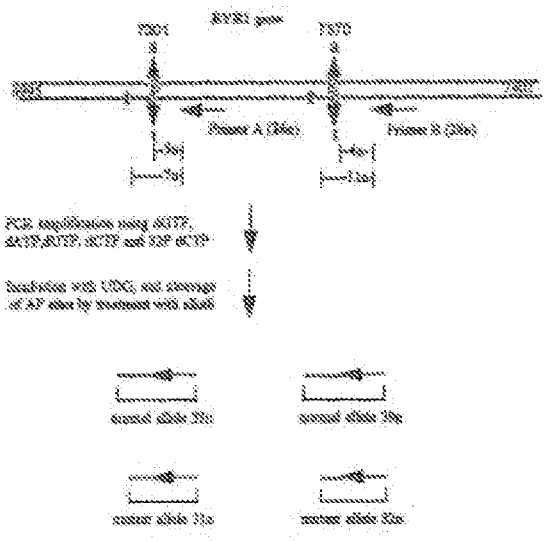
Priority number(s): WO1995IE00067 19951221 ; IE19950000526 19950711

Also published as: EP0843735 (B1) EP0843735 (B2) AT220725 (T) AU4399596 (A) AU708821 (B2) more

Abstract not available for EP0843735 (A1)
Abstract of corresponding document: WO9703210 (A1)

A method for rapidly detecting the presence or absence of a particular nucleic acid sequence at a candidate locus in a target nucleic acid sample comprises the steps of: i) introducing a modified base which is a substrate for a DNA glycosylase into said candidate locus at one or more preselected positions; ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site; iii) cleaving phosphate linkages at abasic sites generated in step ii); and iv) analysing the cleavage products of step iii) so as to identify in said target nucleic acid sequence the presence or absence of said particular nucleic acid sequence at said candidate locus. The method has particular application for detecting specific mutations in a DNA sample, including the detection of multiple known mutations in DNA. Large throughput of sample can be achieved rapidly and easily.

IDT Exhibit 1020.9596



IDT Exhibit 1020.9597

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 843 735 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
17.07.2002 Bulletin 2002/29

(21) Application number: **95942776.6**

(22) Date of filing: **21.12.1995**

(51) Int Cl.[7]: **C12Q 1/68**

(86) International application number:
PCT/IE95/00067

(87) International publication number:
WO 97/03210 (30.01.1997 Gazette 1997/06)

(54) GLYCOSYLASE MEDIATED DETECTION OF KNOWN NUCLEOTIDE SEQUENCES

GLYCOSYLASE VERWENDENDE DETEKTION VON BESTIMMTEN NUKLEINSÄURESEQUENZEN

DETECTION DE SEQUENCES NUCLEOTIDIQUES CONNUES INDUITE PAR LA GLYCOSYLASE

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **11.07.1995 IE 950526**

(43) Date of publication of application:
27.05.1998 Bulletin 1998/22

(73) Proprietors:
- **Forfas (trading as Bioresearch Ireland)**
  Wilton Place, Dublin 2 (IE)
- **University College Cork-National University of Ireland, Cork**
  Cork (IE)

(72) Inventors:
- **McCARTHY, Thomas, Valentine**
  Montenotte, Cork (IE)
- **VAUGHAN, Patrick, Martin**
  Frankfield, Cork (IE)

(74) Representative: **Ryan, Anne Mary**
c/o Anne Ryan & Co.
60 Northumberland Road
Ballsbridge Dublin 4 (IE)

(56) References cited:
WO-A-92/01814

- **METHODS IN MOLECULAR BIOLOGY, vol. 9, 1991, HUMANA PRESS INC., CLIFTON, NJ, US, pages 51-68, XP002004234 D.R. BENTLEY ET AL.: "Rapid methods for detection of polymorphic markers in genomic DNA"**
- **NUCLEIC ACIDS RESEARCH, vol. 20, no. 4, 25 February 1992, IRL PRESS LIMITED,OXFORD,ENGLAND, pages 771-775, XP002004235 W.T. PU AND K. STRUHL: "Uracil interference, a rapid and general method for defining protein-DNA interactions involving the 5-methyl group of thymines: The GCN4-DNA complex" cited in the application**
- **NUCLEIC ACIDS RESEARCH, vol. 21, no. 15, 25 July 1993, IRL PRESS LIMITED,OXFORD,ENGLAND, pages 3437-3443, XP002004236 P.R. DEVCHAND ET AL.: "Uracil-DNA glycosylase as a probe for DNA interactions" cited in the application**
- **ANALYTICAL BIOCHEMISTRY, vol. 206, no. 1, 1 October 1992, pages 91-97, XP000311343 AYOUB RASHTCHIAN ET AL: "URACIL DNA GLYCOSYLASE-MEDIATED CLONING OF POLYMERASE CHAIN REACTION-AMPLIFIED DNA: APPLICATION TO GENOMIC AND CDNA CLONING"**
- **PROGRESS IN NUCLEIC ACID RESEARCH AND MOLECULAR BIOLOGY, vol. 22, 1963, NEW YORK, US, pages 135-192, XP002004237 T. LINDAHL: "DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair"**
- **GENOMICS, vol. 14, no. 2, October 1992, ACADEMIC PRESS INC., NY, US, pages 249-255, XP000571637 A.-L. LU AND I.-C. HSU: "Detection of single DNA base mutations with mismatch repair enzymes" cited in the application**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 0 843 735 B1

IDT Exhibit 1020.9598

Description

Technical Field

[0001] This invention relates to a method for rapidly detecting the presence or absence of a particular nucleic acid sequence at a candidate locus in a target nucleic acid sample. In particular, the invention relates to a method for detecting specific mutations in a DNA sample.

Background Art

[0002] Detection of specific sequences at candidate loci in target nucleic acid samples is highly important for several reasons relating to diagnosis of inherited disorders and of infectious diseases. Detection of multiple different mutations is necessary for screening for the presence of specific genetic disorders. Detection of specific sequences in amplified DNA samples significantly enhances current DNA diagnostic methods for the detection of organisms of interest and especially of infectious organisms. Current techniques for the detection of sequences at candidate loci are either cumbersome, lacking in specificity, difficult to optimise and use or poorly adaptable to high sample throughput.

[0003] There are several methods known for the detection of a particular nucleic acid sequence at a candidate locus in a target nucleic acid sample. Details of these methods are as follows.

1) Restriction enzyme analysis

[0004] Detection of a particular DNA sequence by restriction enzyme analysis. Restriction enzymes cleave DNA at specific sequences. For example, the enzyme EcoRI cleaves double stranded DNA at the sequence GAATTC. Thus, by checking a candidate locus in a particular DNA sample for cleavage with EcoRI, one may determine whether the sequence GAATTC is present or absent at the candidate locus. The appearance or disappearance of a restriction site from a candidate locus indicates that one or more of the bases in the sequence at the candidate locus has been altered. Thus, the creation or loss of a restriction site at a candidate locus can involve the alteration of any base or bases of the respective restriction site. Therefore, the appearance or disappearance of a restriction site at a candidate locus may not inform the investigator of the exact sequence change. However, there are a number of problems with the use of restriction enzymes for the detection of specific sequences at candidate loci. These include:

a) the presence or absence of a specific sequence at a candidate locus regularly does not occur at a restriction site;

b) since different restriction enzymes have different recognition sequences, different restriction enzymes are regularly required for the detection of the presence or absence of a particular sequence at different candidate loci; and

c) because different enzymes can be required for the detection of the presence or absence of a particular sequence at different candidate loci, a high throughput is difficult to achieve with this approach and automation is difficult.

2) DNA sequencing

[0005] DNA sequencing allows detection of a particular sequence at any candidate locus. The main methods of DNA sequencing are the Sanger method (Sanger, F. and Coulson, A.R. (1975) J. Mol. Biol. *94*, 444-448), also known as the dideoxy method or chain termination method. and the Maxam Gilbert method (Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564), also known as the chemical method.

[0006] While DNA sequencing is the ultimate way of determining if a particular sequence is present or absent at a particular candidate locus, it suffers from a number of drawbacks as follows:

a) it is a cumbersome and difficult method for routine detection of the presence or absence of a particular sequence at different candidate loci;

b) full size sequencing gels are required for determination of the presence or absence of a particular sequence at a given candidate locus;

c) the DNA sample for analysis needs to be of high quality in order to obtain good quality DNA sequences;

d) DNA sequencing of directly amplified DNA samples is regularly problematic;

e) sequencing gels can often be difficult to read;

f) high throughput with a high success rate is difficult to achieve;

g) some DNA sequences are more difficult to obtain than others; and

h) resolution of multiple DNA fragments of different size is necessary to detect the presence or absence of a specific sequence at a candidate locus.

3) Uracil interference

[0007] A method has been described whereby uracil is incorporated into an amplified DNA molecule random-

IDT Exhibit 1020.9599

ly and at a low level. This is achieved by amplifying the DNA in the presence of the normal DNA precursor nucleotides and dUTP. The ratio of dTTP to dUTP is chosen so that in the amplification process dUTP is occasionally incorporated opposite an adenine residue on the template strand while dTTP is incorporated opposite adenine residues at a much higher frequency. This results in a population of products bearing a low level of uracil residues randomly distributed throughout the amplified molecules. Treatment of the amplified products with uracil glycosylase and cleavage of the abasic site results in cleavage of the molecules at the position of incorporation of the uracil residues. Because the uracil was incorporated randomly opposite adenine residues at a low level, different molecules will be cleaved at different points depending on where the uracil residues were incorporated. Thus labelling of one of the primers used in the amplification process and separation of all of the cleavage products on a DNA sequencing gel produces a ladder of fragments that allows the determination of the total number of positions of uracil incorporation in one strand of the amplified DNA sample (Tu, W. T. and Struhl, K. Nuc. Acids Res., (1992) *20,* 771-775. Devchand, P.R. *et al.,* Nuc. Acids Res. (1993) *21*, 3437-3443).

[0008] The main application of the approach outlined has been for DNA footprinting (a method used to identify the bases in DNA to which particular proteins bind).

[0009] The uracil incorporation method can be used to determine the location of the total number of uracil residues in an amplified DNA sample. However,

a) it would be a cumbersome and difficult method for rapid detection of an uracil residue at a specific candidate locus;

b) full size sequencing gels would be required for determination of an uracil residue at a specific candidate locus;

c) sequencing gels can often be difficult to read/interpret; and

d) resolution of multiple DNA fragments of different size would be necessary to detect an uracil residue at a specific candidate locus.

4) Use of mismatched nucleotide glycosylases

[0010] Two accounts of the use of mismatched nucleotide glycosylases have been reported for detection of point mutations (Lu, A-L and Hsu, I-C, Genomics (1992) *14*, 249-255 and Hsu, I-C., *et al*, Carcinogenesis (1994) *14*, 1657-1662). The glycosylases in question are the *E. coli* Mut Y gene product which releases the mispaired adenines of A/G mismatches efficiently and A/C mismatches inefficiently by regular glycosylase action and the human thymidine DNA glycosylase which cleaves at G/T mismatches. These enzymes have been used for mutation detection in amplified heteroduplex DNA molecules where mismatches are present. Labelling of one of the primers used in the amplification reaction permits detection of the position of the mismatch after glycosylase treatment, cleavage of the abasic site and resolution of the fragments by gel electrophoresis.

[0011] There are several problems with this method as follows:

a) the method is dependent on the formation of heteroduplex molecules where a mismatched based pair is formed at the site in question. Thus, to detect a mutation in a homozygous sample, an external probe must be provided and hybridization carried out to generate the mismatch;

b) the method permits detection of the position of the mismatch but not necessarily the sequence at the mismatch; and

c) not all mismatches are recognised with equal efficiency.

5) Other methods based on cleavage of mismatched base pairs

[0012] Several other methods based on the cleavage of mismatched base pairs have been described for detecting point, deletion and insertion type mutations. These include chemical cleavage at mismatched base pairs, heteroduplex detection based on the slower migration thereof during gel electrophoresis relative to homoduplex molecules which migrate faster, RNAse cleavage of mismatches in RNA:DNA hybrids and cleavage of mismatches by enzymatic means.

[0013] All of the above methods can detect mutations in heteroduplex molecules and allow the approximate position of the mutation in the nucleic acid molecule to be determined (except in the case of the heteroduplex electrophoresis retardation method which only informs of the presence of a mutation in the sample). However, these methods only work on heteroduplex DNA and do not allow one to deduce what specific sequence is present at a candidate locus.

6) Ligase chain reaction

[0014] The ligase chain reaction (LCR) is a probe amplification method that can be used for detecting the presence or absence of a particular target sequence at a candidate locus. It utilises the enzyme DNA ligase to join two pairs of oligonucleotides that hybridise adjacent to one another on the denatured target DNA strands. The enzyme forms a phosphodiester link between the two oligonucleotides, provided that the oligonucleotides at the junction correctly hybridised with the template. Thus an exact match between the oligonucleotides and

IDT Exhibit 1020.9600

the target sequence at the junction permits ligation of the oligonucleotides resulting in the formation of a larger product which is the cumulative size of both the oligonucleotides. Multiple cycles of annealing, ligation and denaturation results in the exponential amplification of the larger product. Thus detection of the larger product indicates the presence of a sequence at the candidate locus while absence of the larger product means that there were differences between the oligonucleotide sequences and the sequence of the template DNA.

[0015] While this method has good potential for detecting particular sequences at candidate loci and also offers high throughput of samples, it does not allow one to deduce what specific sequence is present at a candidate locus. In addition, the method requires considerable effort to optimise the process.

7) ARMS method

[0016] By designing appropriate primers for use in the polymerase chain reaction, it is possible to detect the presence or absence of a specific sequence at a candidate locus. In this method known as the amplification refractory method (ARMS), primers are designed so that amplification of a target sequence only occurs if there is a perfect match between the 3' end of the primers and the target sequence. Thus if a pair of primers are designed so that one primer is complementary to a given sequence of the target sample while the partner primer is designed so that its 3' end is complementary to the wild type sequence at the candidate locus, then this pair of primers will only produce a product on amplification if the wild type sequence is present at the candidate locus. If a third partner primer is designed so that its 3' end is complementary to the mutant sequence at the candidate locus, then this primer if used with the primer complementary to the given sequence of the target sample will only produce a product on amplification if a mutant sequence is present at the candidate locus. This method suffers from a number of problems as follows:

a) three primers are required to determine whether wild type or mutant sequence is present at the candidate locus;

b) the method does not allow one to deduce what specific sequence is present at a candidate locus; and

c) the annealing conditions for the ARMS method have to be precise, thus the method is difficult to transfer in many cases and has to be optimised for each mutation investigated.

[0017] Thus, it will be appreciated that there is a need in the nucleic acid diagnostics field for a robust method for detection of specific sequences at candidate loci that allows rapid and high throughput of samples.

[0018] Methods in Molecular Biology, Vol. 9, 1991, p51-68 describes various rapid methods for detecting polymorphic markers in DNA. The polymorphisms are detected by amplification of the DNA surrounding and including the locus of interest and mismatch detection at that locus. One of the methods involves cleaving the DNA phosphate linkages at sites of mismatches following modification and subsequently analysing the cleavage products so as to identify the presence or absence of a particular polymorphism in the genomic DNA. The alternative methods described therein rely on the differential hybridisation and/or extension of oligonucleotide primers to sites of polymorphisms dependent on whether a mismatched base pair is generated during the hybridisation step. These methods rely on analysing the hybridised or amplified products so as to identify the presence or absence of a particular polymorphism in the genomic DNA. The methods described do not involve the introduction of a modified base, which is a substrate for a DNA glycosylase, into the amplified DNA and do not involve the excision of such a modified base by the DNA glycosylase.

Disclosure of Invention

[0019] The invention provides a method for rapidly detecting the presence or absence of a particular nucleic acid sequence at a candidate locus in a target nucleic acid sample which comprises the steps of:

i) introducing a modified base which is a substrate for a DNA glycosylase into said candidate locus at one or more preselected positions;

ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site;

iii) cleaving phosphate linkages at abasic sites generated in step ii); and

iv) analysing the cleavage products of step iii) so as to identify in said target nucleic acid sequence the presence or absence of said particular nucleic acid sequence at said candidate locus.

[0020] The method according to the invention offers significant advantages over existing methods in that a single enzyme and a single process can be used to detect multiple known mutations in DNA. Thus, a large throughput of sample can be achieved rapidly and easily as hereinafter demonstrated.

[0021] The method in accordance with the invention enables one to investigate a target nucleic acid to determine if a particular sequence, such as a gene mutation is present or absent at a particular location known as the "candidate locus" herein.

[0022] Preferably, the candidate locus is amplified using normal DNA precursor nucleotides and at least one

IDT Exhibit 1020.9601

modified precursor nucleotide.

[0023] The term "amplifying" as used herein refers to any *in vitro* process for increasing the number of copies of a nucleotide sequence or sequences. Amplification of a target sample results in the incorporation of precursor nucleotides into the DNA being amplified. Typically, amplification of a target sample is carried out using appropriate primers in the polymerase chain reaction (PCR). Amplification of a target sample may be also carried out using the ligase chain reaction (LCR) and a variation of the LCR which employs a short PCR step (PLCR). Precursor nucleotides in the case of a DNA amplification process refer to the deoxyribonucleotides dATP, dCTP, dGTP and dTTP herein referred to as "normal" DNA precursor nucleotides. Modified precursor nucleotide(s) as used herein refers to a modified nucleotide or nucleotides that can be incorporated into a nucleic acid so that a substrate base or bases (glycosylase substrate base) is generated which is recognised by a DNA glycosylase enzyme.

[0024] The amplification will typically involve amplifying a target nucleic acid sample using a combination of normal DNA precursor nucleotides and one or more modified precursor nucleotide(s) where the modified precursor nucleotide replaces one of the normal precursor nucleotides. The incorporation of a modified precursor nucleotide into the amplified product generates one or more glycosylase substrate base(s) at one or more positions recognised by a DNA glycosylase enzyme in the amplified product. A particular sequence may be present at the candidate locus in all or a portion of the target nucleic acid sample, or may be absent from the target nucleic acid sample.

[0025] Further, preferably, at least one of the primers for the amplification is positioned adjacent to the candidate locus.

[0026] Thus, suitably primers for amplification purposes are designed such that one of the primers is positioned adjacent (referred to as the "adjacent primer" herein) to the candidate locus so that during the amplification process, the position of incorporation of the first modified precursor nucleotide into the extended adjacent primer will be at, or distal to, the candidate locus depending on the particular sequence present at the candidate locus. The other primer of a pair of primers as used herein is referred to as the "distal primer" as hereinafter defined.

[0027] Design of the adjacent and distal primers so that the modified precursor nucleotide is incorporated at the candidate locus alone or not incorporated at all, if a particular sequence is present or absent in the amplified target nucleic acid sample, permits cleavage of the target strand bearing the glycosylase substrate base into two fragments.

[0028] Amplification of the target strand of the target sample can be achieved using the adjacent oligonucleotide primer which anneals to the complementary region of the complementary target strand of the target nucleic acid sample. Primer extension of the adjacent primer in the amplification process results in the incorporation of precursor nucleotides and modified precursor nucleotides in a 5' to 3' direction. The amplified DNA strand generated through the extension of the adjacent primer in the amplification process is referred to as the "target strand" herein. The amplified DNA strand generated through the extension of the other primer (referred to as the "distal primer" herein) in the amplification process is referred to as the "complementary target strand" herein. Amplification of the complementary target strand of the target sample can be achieved using the distal oligonucleotide primer which anneals to the complementary region of the target strand of the target nucleic acid sample. Primer extension of the distal primer in the amplification process results in the incorporation of precursor nucleotides and modified precursor nucleotides in a 5' to 3' direction. Amplification of both strands of the target sequence can be achieved using both the adjacent and distal primers. In the PCR process, repeated cycles of amplification are performed. The size of the amplified fragment in this case is delineated by the position of annealing of the adjacent and distal primers to the strands of the target nucleic acid sample.

[0029] For detection of a particular sequence such as a mutation at the candidate locus in a target sample, primers for amplification purposes are designed so that the adjacent primer is positioned close to the candidate locus so that during the amplification process, the position of the first modified precursor nucleotide incorporated into the extended adjacent primer will be at, or distal to, the candidate locus depending on whether a particular sequence is present or absent at the candidate locus. In cases where the incorporation of a sole modified precursor nucleotide at the candidate locus is desirable, the primers are designed so that all of the bases in the primers promote preferential incorporation of nucleotides other than the key modified precursor nucleotide (the key modified precursor nucleotide being the modified precursor nucleotide to be incorporated at the candidate locus depending on the presence or absence of a particular sequence) in the newly synthesised DNA complementary to the primer sequences. For example, in the case where dUTP is the modified precursor nucleotide, the primers are synthesised so that the adenine bases are replaced by cytosine, guanine, thymine, inosine or modified bases (other than uracil) which preclude incorporation of uracil residues in the newly synthesised DNA complementary to the primer sequences. In the case where dITP is the modified precursor nucleotide, the primers are synthesised so that the cytosine bases are replaced by guanine, thymine, adenine, uracil or modified bases (other than inosine) which preclude incorporation of inosine residues in the newly synthesised DNA complementary to the primer sequences. In cases where cleavage at specific points in the primers is desirable in addition to cleavage at the candidate locus on the target strand, the primers are synthesised so

IDT Exhibit 1020.9602

that one or more glycosylase substrate base(s) is/are present in the primer at a defined position or positions. Treatment of the primer with the appropriate glycosylase after the amplification process will result in cleavage of the primers at a specific position or positions. Design of the primers in such a fashion facilitates detection of specific fragments of the target strand indicative of the presence or absence of a particular sequence at the candidate locus and serves to reduce the size of primers if desirable so that they do not interfere with the detection process.

[0030] Suitably at least one primer is labelled when an amplification method is used in accordance with the invention. Labelling of one of the primers prior to the amplification process allows detection of the amplified target or complementary strand alone. Labelling of the primers can be performed by a variety of means including addition of a radioactive, fluorescent, or detectable ligand to the primer during or post primer synthesis. The use of a labelled precursor nucleotide (i.e. a radioactive precursor nucleotide, or a precursor nucleotide with a linked fluorescent or detectable ligand group) in the amplification process facilitates detection of the target stand and the complementary target strand and any DNA fragments arising from the glycosylase mediated cleavage process that bear the incorporated label. DNA staining methods such as silver or ethidium bromide staining facilitates detection of all of the fragments generated as a result of the glycosylase mediated cleavage of the amplified target nucleic acid sample after separation of the fragments. Alternatively, detection of the amplified target and complementary strand and fragments generated as a result of the glycosylase mediated cleavage can be accomplished using appropriate nucleic acid hybridisation probes.

[0031] The modified base can be introduced by chemical modification of an existing base at the candidate locus.

[0032] Several methods exist where treatment of DNA with specific chemicals modify existing bases so that they are recognised by specific DNA glycosylase enzymes. For example, treatment of DNA with alkylating agents such as methylnitrosourea generates several alkylated bases including N3-methyladenine and N3-methylguanine which are recognised and excised by alkyl purine DNA-glycosylase. Treatment of DNA with sodium bisulfite causes deamination of cytosine residues in DNA to form uracil residues in the DNA which can be cleaved by uracil DNA-glycosylase.

[0033] Thus, from knowledge of the prior art, bases present at the candidate locus of an amplified target nucleic sample can be converted into glycosylase recognisable substrates by chemical means. For example, a cytosine present at the candidate locus can be readily converted into an uracil, thereby rendering the amplified sample susceptible to uracil DNA-glycosylase cleavage at the candidate locus. If the adjacent primer is synthesised so that it contains 5-methylcytosine rather than cytosine in such a case, the primer will be resistant to uracil DNA-glycosylase mediated cleavage since deamination of 5-methylcytosine generates a thymine residue rather than an uracil residue.

[0034] Preferably, the modified base is excised (or cleaved) by means of a DNA glycosylase enzyme.

[0035] Thus, suitably following the amplification process, the amplified product is treated with a suitable DNA glycosylase enzyme which recognises and releases the glycosylase substrate bases present in the amplified target sample and consequently generates apurinic or apyrimidinic sites in the amplified target nucleic acid sample.

[0036] In the case where the modified precursor nucleotide is dUTP, the glycosylase substrate base uracil will be generated in the amplified target nucleic acid sample. Addition of uracil DNA-glycosylase to the sample releases the uracil from the sample. In the case where the modified precursor nucleotide is dITP, the glycosylase substrate base hypoxanthine will be generated in the amplified target nucleic acid sample. Addition of alkylpurine DNA-glycosylase to the sample releases the hypoxanthine from the sample. Release of the glycosylase substrate bases from the amplified target nucleic acid sample results in an apyrimidinic site in the case of uracil and an apurinic site in the case of hypoxanthine.

[0037] Glycosylase mediated cleavage of the target strand of an amplified target nucleic acid sample as a result of a particular sequence being present or absent at the candidate locus in the amplified target nucleic acid also permits detection by immobilisation methods. In this instance one of the primers usually the adjacent primer is synthesised with a "capture" agent attached which allows immobilisation of the target strand to a solid matrix. The primers are designed so that the modified precursor nucleotide is incorporated at the candidate locus alone in the amplified target nucleic acid sample. The amplification process is carried out in the presence of a labelled precursor nucleotide resulting in the incorporation of label into the extended adjacent primer distal (3') to the candidate locus. Immobilisation of the amplified target strand is achieved by incubating the amplified target strand with a solid matrix bearing a molecule which binds the capture agent specifically. Removal of the complementary target strand is then performed by washing with a denaturating agent and removing all of the non-immobilised material. If a glycosylase substrate base is present at the candidate locus, release of the labelled portion of the target strand will occur in the glycosylase mediated cleavage process. Removal of the released fragment and determination of the level of label released or remaining immobilised is diagnostic of the presence or absence of the particular sequence at the candidate locus. Alternatively, if labelling of the target strand is not desirable, detection may be performed by use of specific nucleic acid hybridisation probes. Such probes can be readily designed to detect the complete immobilised target strand, the portion of the target

IDT Exhibit 1020.9603

strand released after the glycosylase mediated cleavage process or the portion of the target strand remaining after the cleavage process.

[0038] This glycosylase mediated cleavage process allows the detection of any particular sequence at a candidate locus in any target sample provided the DNA sequence surrounding the mutation site is known. Sequence information of at least 15 to 20 nucleotides on each side of the candidate locus is required in order to design appropriate primers for the amplification process.

[0039] The use of a DNA glycosylase enzyme which recognises a modified base has not previously been used directly for the detection of a particular sequence at a candidate locus in an amplified target nucleic acid sample.

[0040] A main application of the glycosylase mediated cleavage process is that it permits detection of the presence or absence of a particular sequence at a candidate locus by the aforementioned immobilisation methods.

[0041] A suitable "capture" agent is biotin. Thus, one of the primers is suitably synthesised with biotin attached which allows immobilisation of the target strand to streptavidin coated on, or linked to a solid matrix so as to allow immobilisation of the target strand. The primers are designed so that the modified precursor nucleotide is only incorporated in the target strand at the candidate locus alone if a particular sequence is present or absent at the candidate locus.

[0042] Suitably one of the precursor nucleotide(s) is also labelled. As indicated above the amplification process is preferably carried out in the presence of a labelled precursor nucleotide, such as an alpha P^{32} dNTP, a fluorescent dNTP or digoxygenin dNTP resulting in the incorporation of the label into the extended adjacent primer. The primer bearing the capture groups is also designed so that the labelled precursor nucleotide(s) can only be incorporated distal to (3') the primer in question and the candidate locus i.e. the labelled precursor nucleotide cannot be incorporated between the candidate locus and the adjacent primer.

[0043] Preferably the phosphate linkages at the abasic sites are cleaved by a treatment selected from alkali treatment or other chemical treatment, heat treatment and treatment with an enzyme.

[0044] The process of strand cleavage resulting from the release of glycosylase substrate base by DNA glycosylase action followed by abasic site cleavage is referred to herein as glycosylase mediated cleavage. The presence or absence of a particular sequence at the candidate locus results in the incorporation, or lack of incorporation of a modified precursor nucleotide at the candidate locus in the amplified target nucleic acid sample. Thus if a modified precursor nucleotide is incorporated at the candidate locus, cleavage of the target strand will occur at the candidate locus through the glycosylase mediated cleavage process and this cleavage point will be the closest cleavage point to the 3' end of the extended adjacent primer. If a modified precursor nucleotide is not incorporated at the candidate locus, the closest cleavage point to the 3' end of the extended adjacent primer will occur at the first point distal to the candidate locus where a modified precursor nucleotide is incorporated. Therefore the observed length of the extended adjacent primer following the glycosylase mediated cleavage process will be diagnostic of the presence or absence of the particular sequence at the candidate locus.

[0045] The preferred treatment is alkali at high temperature, or with an enzyme which cuts specifically at apurinic or apyridimic sites, such as *E.coli* endonuclease IV. Both of these treatments cleave the apurinic or apyridimic site to completion on the 5' side. In a case where there is a sole glycosylase substrate base at the candidate locus in the amplified target strand, glycosylase mediated cleavage cuts the target strand at a single position yielding two fragment strands. Manipulation of the design of the primer sequences used permits amplification of products bearing glycosylase substrate bases at any desired position(s) on the target and/or complementary strand in addition to the candidate locus on the target strand and facilitates subsequent analysis of the glycosylase cleaved amplified target nucleic acid sample.

[0046] The presence or absence of a particular sequence at the candidate locus determines whether or not a glycosylase substrate base is incorporated at the candidate locus in the target strand. Thus different diagnostic cleavage patterns are produced by glycosylase mediated cleavage of an amplified target strand depending on whether the particular sequence is present or absent at the candidate locus. Primers may also be designed so that glycosylase substrate bases are absent from the target strand if a particular sequence is present at the candidate locus. In such a case, the target strand is resistant to glycosylase mediated cleavage.

[0047] The products/cleavage pattern resulting from the glycosylase mediated cleavage of the amplified target nucleic acid sample may be discerned by existing DNA sizing methods such as polyacrylamide gel electrophoresis, agarose gel electrophoresis or high performance liquid chromatography (HPLC).

[0048] Thus, the size of the extended adjacent primer when an amplification method is used in accordance with the invention can be determined after denaturation by existing DNA sizing methods such as denaturating polyacrylamide gel electrophoresis. Labelling of the adjacent primer prior to the amplification process facilitates detection of the cleaved extended adjacent primers alone while the use of a labelled precursor nucleotide in the amplification process or DNA staining methods facilitated detection of all of the fragments generated as a result of glycosylase mediated cleavage of the amplified target nucleic acid sample.

[0049] The modified base used is preferably uracil or hypoxanthine.

IDT Exhibit 1020.9604

[0050] Thus. the preferred modified precursor nucleotides are dUTP and dITP which when incorporated into DNA generate the glycosylase substrate bases uracil and hypoxanthine respectively. The modified precursor nucleotide dUTP is a base sugar phosphate comprising the base uracil and a sugar phosphate moiety. The modified precursor nucleotide dITP is a base sugar phosphate comprising the base hypoxanthine and a sugar phosphate moiety. Uracil in DNA is recognised specifically by uracil DNA-glycosylase (UDG) and released from DNA. Uracil DNA glycosylase also recognises certain other uracil related bases when present in DNA. Hypoxanthine is recognised specifically by alkylpurine DNA glycosylase (ADG) and released form DNA. This enzyme also recognises and releases N3 methyl adenine. N3 methyl guanine, O2 methyl cytosine and O2 methyl thymine when present in DNA. Amplification of a target DNA sequence using the precursor nucleotides dATP, dCTP and dGTP and the modified precursor nucleotide dUTP results in an amplified DNA where thymine is replaced by uracil. The uracil is incorporated in the newly synthesised DNA strand at positions complementary to adenine residues in the template DNA strand during the amplification process. Amplification of a target DNA sequence in the presence of the precursor nucleotides dATP, dCTP and dTTP and the modified precursor nucleotide dITP results in an amplified DNA where guanine is preferentially replaced by hypoxanthine. The hypoxanthine is preferentially incorporated opposite cytosines in the template DNA strand in the amplification process when the other precursor nucleotides are not limiting.

[0051] Any DNA or RNA from any source can serve as a target sample.

[0052] Preferably, the target nucleic acid sample is DNA. For amplification purposes, RNA is first converted into cDNA by reverse transcription. A particular sequence may be present at the candidate locus in all or a portion of the target nucleic acid sample, or may be absent from the target nucleic acid. The target nucleic acid sample may be homozygous or heterozygous for the presence or absence of a particular sequence at the candidate locus which is one of the advantages of the method according to the invention over some of the prior art methods discussed above.

[0053] The DNA which can be used in accordance with the method of the invention can be single stranded, homoduplex or heteroduplex DNA.

[0054] Some of the advantages of the present invention relative to specific prior art methods are as follows.

[0055] The invention has advantages over the restriction enzyme analysis method given that a minimum of two glycosylases are required to identify all possible sequence changes due to mutations such as point, insertion or deletion type mutations at a candidate locus and one glycosylase can be used to detect 10 out of all 12 possible point mutations.

[0056] The method according to the invention is significantly faster than DNA sequencing in that full size DNA sequencing type gels are not necessary for sizing the cleavage products and the number of cleavage products that need to be detected to determine if a specific sequence is present or absent in a target nucleic acid is small. Usually, the appearance or disappearance of a single additional DNA fragment in addition to the wild type fragment will be sufficient to diagnose the presence or absence of the particular DNA sequence in the target nucleic acid. By contrast in DNA sequencing, resolution of multiple DNA fragments of different size is necessary to detect the presence or absence of a specific sequence at a candidate locus. The method according to the invention allows the detection of the presence or absence of a specific sequence at a candidate locus by determining whether a DNA fragment is released or not from an immobilised target nucleic acid. This is not possible with DNA sequencing.

[0057] The method according to the invention also represents a significant improvement over the uracil interference method. The method according to the invention differs from the uracil interference method in that the modified nucleotide introduced into the target nucleic acid sample is introduced at all preselected positions whereas in the case of the uracil interference method uracil is incorporated randomly and at a low level in the amplified molecules. Also in the case of the method according to the invention the modified nucleotide introduced replaces a particular DNA precursor, while in the uracil interference method a ratio of modified nucleotide to normal nucleotide is used. Other advantages of the method according to the invention relative to the uracil interference method are that the method according to the invention is significantly faster, full size DNA sequencing gels are not necessary for sizing the cleavage products, and the number of cleavage products that need to be detected is small. Usually, the appearance or disappearance of a single additional DNA fragment in addition to the wild type fragment will be sufficient to diagnose the presence or absence of the particular DNA sequence in the target nucleic acid. By contrast in the uracil interference method, resolution of multiple DNA fragments of different size would be necessary to detect an uracil residue at a specific candidate locus. The method according to the invention allows the detection of the presence or absence of a specific sequence at a candidate locus by determining whether a DNA fragment is released or not from an immobilised target nucleic acid. This is not possible with the uracil interference method.

[0058] Advantages of the method according to the invention relative to use of mismatched nucleotide glycosylases are that external probes are not required to detect the presence or absence of a particular sequence (a mutation or otherwise) in a homozygous or heterozygous state. In the method according to the invention certain glycosylases such as uracil DNA-glycosylase work on single stranded DNA so that single stranded

IDT Exhibit 1020.9605

DNA can be investigated. Furthermore, the modified bases used in accordance with the invention are recognised and removed efficiently by the glycosylase enzymes which recognise them. Thus the invention offers significant advantages over the mismatch cleavage method outlined about in that the provision of external probes is not necessary, no hybridisation steps are required, single stranded DNA can be used and high throughput of samples can be achieved rapidly and easily.

[0059] Finally, the method according to the invention has significant advantages over the ARMS method in that only one amplification reaction is necessary to determine whether wild type or mutant sequence is present at the candidate locus. The annealing conditions for the ARMS method have to precise, thus the method is difficult to transfer in many cases and has to be optimised for each mutation investigated. By contrast the method according to the invention is robust, and is optimised easily and enables one to achieve a high throughput of samples.

Brief Description of Drawings

[0060]

Fig. 1 is a schematic representation of procedure A) described in Example 1;

Fig. 2 is a schematic representation of procedure B) described in Example 1;

Fig. 3 is a schematic representation of the procedure described in Example 2;

Fig. 4 is a schematic representation of the procedure described in Example 3;

Fig. 5 is a schematic representation of the procedure described in Example 4;

Fig. 6 is a schematic representation of the procedure described in Example 5;

Fig. 7 is a schematic representation of the procedure described in Example 6;

Fig. 8 is a schematic representation of the procedure described in Example 7; and

Fig. 9 is a schematic representation of the procedure described in Example 8.

Modes for Carrying Out the Invention

[0061] The invention will be further illustrated by the following Examples.

Example 1

[0062] The method according to the invention was used to detect the presence of a G to A base substitution mutation causal of malignant hyperthermia at position 1021 of the human skeletal ryanodine receptor gene (RYR1) in a patient heterozygous for the mutation. The sequence of steps is depicted in Fig. 1. In this case, the target nucleic acid was DNA extracted from a patient with malignant hyperthermia, the candidate locus was nucleotide 1021 in the human RYR1 gene and the objective was to determine whether the particular sequence base pair G:C or A:T was present at the candidate locus in either RYR1 allele in the patient. The lower strand shown in Fig. 1 is the target strand and the presence or absence of a T (U) nucleotide at the candidate locus was determined. The upper strand is the complementary target strand.

[0063] In the context of the Example the normal allele refers to the RYR1 allele bearing the normal sequence and the mutant allele refers to the RYR1 allele bearing the mutant sequence at position 1021. The DNA sequence surrounding the mutation site is shown in Fig. 1, with the mutation site indicated by a bold upper case letter.

A) The sequence of the target nucleic acid at the candidate locus region and the sequence of the adjacent and distal primers (primers contain standard nucleotides (dG, dA, dT and dC)) are shown in Fig. 1. Six pmoles of the adjacent primer was end-labelled by incubation with 1 unit of polynucleotide kinase, appropriate buffer and 1 uCiγ^{32}P ATP (3000Ci/mmol) for 30 min at 37°C. The target nucleic acid sample was amplified by PCR as follows: the reaction mix for PCR contained 200 ng genomic DNA from the affected patient, 0.2 mM dATP, dCTP, dGTP and dUTP, 1.5mM $MgCl_2$, 50mM KCl, 10mM Tris-HCl pH 9.0,0.1% TritonX-100, 6 pmoles of each primer in a total volume of 19μl. The reaction mix was then overlaid with an equal volume of mineral oil and a hot start PCR was performed wherein the reaction mix was heated to 94°C for 5 min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 20μl). 30 Cycles of 94°C for 60 sec., 59°C for 60 sec. and 72°C for 60 sec. were carried out in thermocycler followed by removal of the aqueous reaction mixture to a separate microtube. The reaction mixture bearing the amplified target nucleic acid was then treated with exonuclease I to digest the primers not extended in the amplification step. This was achieved by incubating 2 μl of the PCR reaction with 0.5 units of exonuclease I at 37°C for 30 min. The exonuclease was subsequently heat inactivated by incubating the reaction at 80°C for 15 min.

Uracil DNA glycosylase (0.05 units) was then added and the incubation continued at room tem-

IDT Exhibit 1020.9606

perature for 30 min. Following Uracil DNA glycosylase treatment, the apurinic and apyridimic (AP) sites generated in the amplified product were cleaved to completion by adding NaOH to a final concentration of 0.25M and heating the mixture for 15 min at 95°C. The reaction was then neutralised by addition of Tris base to 30mM final concentration.

An equal volume of formamide loading dye (90% formamide, 0.025% Bromophenol blue, 0.025% Xylene cylanol) was added to the sample which was then heated at 85°C for 5 min. All of the sample was then loaded onto a 20% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 400 volts for size analysis of the cleaved products in the sample. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hours at -70°C. A labelled set of oligonucleotides (20mer, 22mer, 24mer) was used as markers. Analysis of the autoradiographed products showed that two cleavage products were present in approximately equal amounts. One product was 22 nucleotides (n) in length, corresponding to the normal allele, while the second product corresponding to the mutant allele, was 20n in size, as hereinafter described.

The 20n product could only have been generated if an uracil residue was incorporated at position 1021 in the target strand. Thus the detection of a 20n fragment is diagnostic of the presence of a T residue at position 1021. In the normal allele the first uracil incorporated should be at position 1019 and should result in a 22n product. Thus the detection of a 22n fragment is diagnostic of the absence of a T residue at position 1021 and the presence of a T residue at position 1019. Both the 20n and 22n product were detected in roughly equal amounts indicating that the target nucleic acid was from an individual heterozygous for the G1012A mutation. If the patient was homozygous normal, then only a 22n product should be detected. If the patient was homozygous affected, then only the 20n product should be detected. The relative intensity of the 22n product to the 20n product allows one to determine the relative levels of the normal and mutant allele in a target nucleic acid sample. This is especially useful for analysing complex samples where there may be a large difference between the levels of a particular normal and mutant allele.

B) Redesign of the adjacent primer so that the 3' end of the primer was at position 1023, 1024 or 1025 and so that the primer maintained an appropriate optimal length (by adjusting the 5' endpoint appropriately) is permissible. For instance, if the adjacent primer is moved 3 nucleotides in the 5' direction, the primer sequence would be 5'CTG CAC GAA GCA CAG TGA CT 3'. Thus a product of 23n would be generated from the mutant allele and a product of 25n would be generated from the normal allele. There would not be any need to remove the non utilised primers since the adjacent primer would be 20n in length. This modification is depicted in Fig. 2.

Example 2

[0064] In certain instances, it is advantageous to design the primers for amplification in accordance with the invention so that a greater and/or a clearer size difference exists between the glycosylase cleaved products diagnostic of the presence or absence of a particular sequence at a candidate locus. This can be achieved by altering the sequence of one or both of the primers so that the position of incorporation of the glycosylase recognisable substrate base(s) distal to the adjacent primer and the candidate locus is altered. This can be achieved by synthesising the primers so that some or all of the residues promoting incorporation of glycosylase substrate base(s) in the newly synthesised DNA are replaced by nucleotides which do not promote the incorporation of the glycosylase substrate base(s).

[0065] The protocol used in the present Example was exactly the same as that described in Example 1 except that the distal primer was synthesised with an inosine residue at the penultimate 3' position (position 1019) and exonuclease treatment was not carried out to remove the nonutilised primer as this was not necessary in this case. Thus, during amplification of the target strand a cytosine residue rather than an uracil residue was incorporated opposite position 1019 in the newly synthesised DNA. As a result, glycosylase mediated cleavage of the amplified products resulted in a 23n fragment arising from the mutant allele and a 28n fragment arising from the normal allele as shown in Fig. 3.

[0066] Greater differences may be achieved by preventing the incorporation of additional or all uracil residues at potential positions distal to the candidate locus to the extent where cleavage of the target strand only occurs at the candidate locus if an uracil residue is present at the locus. The same approach can also be used to prevent incorporation of additional or all uracil residues at potential positions opposite the adjacent primer on the complementary target strand.

Example 3

[0067] Cleavage of the N-glycosidic bond between a base and the DNA backbone releases the base from the DNA and results in the generation of an apurinic or apyrimidinic site (AP site). The phosphodiester bond on the 3' side of the AP site is alkali labile and is also susceptible to cleavage at neutral pH by β-elimination. Thus, treatment of a DNA sample bearing an AP site with alkali or by heating causes cleavage of the phosphodiester

IDT Exhibit 1020.9607

bond of the DNA backbone and results in a DNA terminus with a 5' phosphate moiety and a terminus with a 3' deoxyribosephosphate. Subsequent removal of the 3' terminal deoxyribosephosphate can be achieved by a variety of methods including treatment with the enzyme exonuclease III or endonuclease IV. It was observed that heating of amplified DNA fragments bearing one or more AP sites for 15 min. at 95°C after uracil DNA-glycosylase cleavage resulted in cleavage of the AP site and generated products with and without a 3' terminal deoxyribosephosphate in about equal proportions. This differential cleavage may be exploited to facilitate accurate detection of the products of the glycosylase mediated cleavage process.

[0068] The protocol of Example 2 was repeated except that following uracil DNA-glycosylase treatment of the amplified target nucleic acid. the AP sites generated in the amplified product were cleaved by heating the mixture for 15 min. at 95°C and because no NaOH was added, it was not necessary to neutralise the reaction. Following completion of the protocol as described in Example 1, analysis of the autoradiograph showed four cleavage products (a 28n product and a product about 1 nucleotide larger which is the 28n product plus 3'-terminal deoxyribosephosphate (28n+) and a 23n product and a product about 1 nucleotide larger which is the 23n product plus 3'-terminal deoxyribosephosphate (23n+)) in approximately equal amounts and the nonutilised labelled primer. This modification is depicted in Fig. 4.

[0069] The 23n and 23n+ products could only have been generated if an uracil residue was incorporated at position 1021 in the target strand. Thus the detection of a 23n and 23n+ fragment is diagnostic of the presence of a T residue at position 1021. In the normal allele the first uracil incorporated should be at position 1016 and should result in a 28n product. Thus the detection of a 28n and 28n+ product was diagnostic of the absence of a T residue at position 1021 and the presence of a T residue at position 1019. In this Example, both the 23n, 23n+, 28n, and 28n+ products were detected in roughly equal amounts indicating that the target nucleic acid was from an individual heterozygous for the G1021A mutation. This approach is especially useful for a multiplex approach since conditions can be readily designed where the appearance of fragments which are larger than the nonutilised primer by a single 3'-terminal deoxyribosephosphate are diagnostic of the presence or absence of a particular sequence at a candidate locus.

Example 4

[0070] Glycosylase mediated cleavage of a particular sequence at a candidate locus in a target strand of a target nucleic acid sample also permits detection of cleavage products using solid phase or immobilised technology. This feature of the invention permits high throughput of samples and avoids the use of time consuming gel electrophoresis analysis. Any candidate locus may be investigated for the presence or absence of a particular sequence by this technology. The main areas for this application are in detection of gene mutations in a target sample and identification of specific organisms.

[0071] This Example and Examples 5 and 6 illustrate application in both areas as follows:

[0072] A procedure was used which was similar to that used in Example 1 in that the target nucleic acid was DNA extracted from a patient with malignant hyperthermia, the candidate locus was nucleotide 1021 in the human RYR1 gene and the objective was to determine if an A:T base pair was present at the candidate locus in one or both of the RYR1 alleles in the patient. The procedure is depicted in Fig. 5. In this case, the lower strand was the "target strand" and the upper strand was the "complementary strand". The presence or absence of a T(U) nucleotide at the candidate locus on the target strand was determined.

[0073] The adjacent and distal primers used were the same as those used in Example 1 except that the adjacent primer is biotinylated at the 5' end, and in the distal primer, all of the A residues were replaced by inosines. As a result of replacement of the A residues, the amplified target strand would only have contained a single uracil residue if an A:T base pair was present at the candidate locus. Thus cleavage of the target strand by uracil DNA-glycosylase mediated cleavage could only occur at a single position (1021) if a uracil residue was present. Therefore, cleavage of the amplified target strand in this instance is diagnostic of the presence of an A:T base pair at position 1021 in the target nucleic acid sample.

[0074] The adjacent primer was biotinylated at the 5' end during synthesis. The target nucleic acid sample was amplified by PCR as follows: the reaction mix for PCR contained 200 ng genomic DNA from the affected patient, 0.2 mM dCTP, dGTP, dATP and dUTP, 1uCi alpha^{32}P dCTP (3000Ci/mmol), 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl pH 9.0, 0.01% Triton X-100. 6 pmoles of each primer in a total volume of 9 μl. The reaction mix was then overlaid with an equal volume of mineral oil and a hot start PCR was performed wherein the reaction mix was heated to 80°C for 5 min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 25 μl). 30 Cycles of 94°C for 60 sec., 59°C for 60 sec. and 72°C for 60 sec. were carried out in a thermocycler.

[0075] Immobilisation of the product was carried out by adding 1 μl of the aqueous PCR reaction to a tube, labelled tube A, containing 20 μl streptavidin coated magnetic beads (Dynal beads. Dynal is a trade mark) and 19μl sterile water and incubating at room temperature for 10 minutes. The tube was then placed in a magnetic stand which caused the beads with the attached biotin labelled PCR product to accumulate against one side of the tube. The supernatant containing unincorporated label and unused distal primer and nucleotide triphosphates was removed and the beads with trapped

IDT Exhibit 1020.9608

PCR amplified DNA were washed several times in buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA and 1M NaCl. The amplified DNA was then denatured by adding 40 µl of 0.1M NaOH. The complementary strand was removed in the supernatant and the target strand remained attached to the streptavidin coated magnetic beads. The beads were again washed with 40 µl of 0.1M NaOH followed by a wash in the aforementioned buffer. The immobilised target strand was then treated with 0.05 units uracil DNA-glycosylase and cleaved at the AP-site by adding alkali to 0.25M NaOH final concentration. The released material was then removed to another tube, labelled tube B, and monitored for the presence of ^{32}P by scintillation counting. The tube containing the cleaved immobilised PCR product was also monitored for the presence of ^{32}P.

[0076] Monitoring for the present of ^{32}P by scintillation counting demonstrated that following uracil DNA-glycosylase cleavage, about half of the radioactive material was released to tube B. This demonstrated that about half of the immobilised PCR product was cleaved due to the presence of an U residue at the candidate locus. Thus it could be concluded that the target sample was heterozygous for the A:T base pair at position 1021 in the RYR1 gene.

Example 5

[0077] Glycosylase mediated cleavage with immobilisation was used to detect the presence of pathogenic organisms of the *Mycobacterium tuberculosis* complex.

[0078] The MPB70 gene is specifically found in *Mycobacterium tuberculosis* complex organisms and several reports have been published where detection of the MPB70 gene by PCR or other means can be used successfully to diagnose the presence of the complex. In this example, specific amplification of a section of the MPB70 gene was carried out so that a single U and label was incorporated in the PCR product. Thus, immobilisation of the product, followed by cleavage with uracil DNA-glycosylase should have released labelled product which could readily be detected.

[0079] As shown in Fig. 6, specific primers were designed to amplify a section of the MPB70 gene so that the amplified section of the target strand delineated by the primers contained several guanine, cytosine or adenine nucleotides and a single uracil residue. To achieve this, the distal primer was designed so that no uracils were incorporated opposite it on the target strand. Following amplification using an adjacent primer with a capture agent and a labelled nucleotide precursor, the amplified target strand was immobilised and cleaved with uracil DNA-glycosylase. Thus a labelled fragment of the target strand was released by this process and could be readily detected.

[0080] More specifically, the objective in this Example was to detect the presence of an amplified fragment of MPB70 (nucleotide 472-540) by monitoring the cleavage of an uracil residue at position 496, incorporated during extension of the adjacent primer. Primers were designed so that a GCA rich sequence (target strand) within the MPB70 gene was amplified using dGTP, dCTP, dUTP and dATP. The distal primer was designed so that inosines (or mismatches) were substituted for adenine residues thereby ensuring that the only uracil present in the target strand of the amplified product was at position 496. The adjacent primer was biotinylated at the 5' end. The sequence of the target nucleic acid at the candidate locus region and the sequence of both the adjacent primer and the distal primer are shown in Fig. 6.

[0081] The target nucleic acid sample was amplified by PCR as follows: the reaction mix for PCR contained 200 ng genomic DNA (or 1 µl of a 1/200 dilution of a PCR amplified region of the MPB70 gene). 0.2 mM dATP, dCTP, dGTP and dUTP, and 1 µCi $\alpha^{32}P$ dCTP (3000/Ci/mmol), 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl pH 9.0, 0.1% Triton X-100, 6 pmoles of each primer in a total volume of 9 µl. The reaction mix was then overlaid with 20 µl of mineral oil and a hot start PCR was performed wherein the reaction mix was heated to 94°C for 5 min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 10 µl). 30 Cycles of 94°C for 60 sec., 58°C for 60 sec. and 72°C for 60 sec. were carried out in a thermocycler followed by removal of the aqueous reaction mixture to a separate microtube.

[0082] Immobilisation of the product was carried out by adding 1 µl of the aqueous PCR reaction to a tube, labelled tube A, containing 20 µl streptavidin coated magnetic beads (Dynal beads) and 19 µl sterile water and incubating at room temperature of 10 min. The tube was then placed in a magnetic stand which caused the beads with the attached biotin labelled PCR product to accumulate against one side of the tube. The supernatant containing unincorporated label and unused distal primer and nucleotide triphosphates was removed and the beads with trapped PCR amplified DNA were washed several times in buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA and 1M NaCl. The amplified DNA was then denatured by adding 40 µl of 0.1 M NaOH. The complementary strand was removed in the supernatant and the target strand remained attached to the streptavidin coated magnetic beads. The beads were again washed with 40 µl of 0.1M NaOH followed by a wash in the aforementioned buffer. The immobilised target strand was then treated with 0.05 units uracil DNA-glycosylase and cleaved at the AP-site by adding alkali to 0.25 M NaOH final concentration. The released material was then removed to another tube, labelled tube B, and monitored for the presence of ^{32}P by scintillation counting. The tube containing the cleaved immobilised PCR product is also monitored for the presence of ^{32}P.

[0083] Monitoring for the presence of ^{32}P by scintillation counting demonstrated that following uracil DNA-glycosylase cleavage, the radioactive material was re-

IDT Exhibit 1020.9609

leased to tube B. This demonstrated that the immobilised PCR product was cleaved due to the presence of an U residue at the candidate locus. Thus it could be concluded that the target sample originated from *Mycobacterium tuberculosis* complex.

Example 6

[0084] The procedure of Example 5 was repeated except that the reaction was carried out using the adjacent primer designed in a different way. In this case, the adjacent primer was designed so that no uracils were incorporated into the target stand during amplification (shown in Fig. 7). Thus, in this case, the amplified product should have been resistant to cleavage with uracil DNA-glycosylase and so the label should not have been released to tube B. In this case. monitoring for the presence of ^{32}P by scintillation counting demonstrated that the label in tube B remained immobilised after incubation with uracil DNA-glycosylase. No label was released to tube B. This application is especially useful for verifying that the correct DNA is immobilised since spurious PCR products generated would be expected to have uracil residues incorporated and thus, label should be released to tube B. The combination of Examples 5 and 6 allows the investigator to diagnose the presence or absence of *Mycobacterium tuberculosis* complex with a high degree of confidence.

Example 7

[0085] Amplification of specific fragments of DNA of a specific size using preselected primers is commonly used for diagnostic purposes in the DNA diagnostics area. One of the limitations of such a diagnostic approach is that artefactual amplification products of the same size as the diagnostic product in question can arise and can potentially result in misdiagnosis. Thus, a method which increases the confidence limit or verifies that a diagnostic amplified fragment is correct has significant value in the diagnostic area since it improves the accuracy of diagnosis and is superior to diagnosis by amplification alone.

[0086] Amplification of specific fragments of the MPB70 gene found in *Mycobacterium tuberculosis* complex organisms using specific primers have been used to diagnose the presence of these organisms in samples. In this Example, the method according to the invention was used to check that a 116bp and a 130bp fragment of the MPB70 gene amplified from *Mycobacterium tuberculosis* complex using specific primers was correct. First of all, PCR was employed using specific primers and a *Mycobacterium tuberculosis* complex DNA sample so that a 116bp and 130 bp product was amplified from the MPB70 gene. The primers and amplification conditions were designed so that when the 116bp and 130bp amplification products were cleaved with uracil DNA-glycosylase, fragments of a specific size should be generated verifying that the initial amplification products were correct.

[0087] The amplification strategy is shown in Fig. 8. As can be seen, primer A (5'ggcctcggtgcagggaatgtc3') and B (5'ccaggtttacttgcggattga3') were selected to amplify a 116bp region of the MPB70 gene encompassing nucleotide 320 to nucleotide 436. The primers were also selected so that the first uracil incorporated was 11 nucleotides downstream of the 3' end of the primer A and 2 nucleotides downstream of primer B. Similarly, primer C and D were selected to amplify a 130bp region of the MPB70 gene encompassing nucleotide 551 to nucleotide 681. The first uracil incorporated in this case was 26 nucleotides downstream of the 3' end of the primer C and 2 nucleotides downstream of primer D. Amplification of the 116bp and 130bp of the MPB70 gene using labelled primers, followed by cleavage with uracil-DNA-glycosylase should result in a 32n and a 23n fragment from the 116bp PCR product and a 47n and a 23n fragment from the 130bp PCR product.

[0088] The 116bp product was amplified from a *Mycobacterium tuberculosis* complex DNA sample by PCR as follows. The reaction mix for PCR contained 200 ng *Mycobacterium tuberculosis* complex DNA. 0.2mM dATP.dCTP, dGTP and dUTP, 1.5mM $MgCl_2$, 50mM KCl, 10mM Tris-HCl pH 9.0, 0.1%Triton X-100 and 6 pmoles of primer A and primer B in a total volume of 19µl. In the first amplification reaction mix, primer A was end-labelled using γ^{32} ATP and polynucleotide kinase, while primer B was labelled in the second amplification. The reaction mixes were then overlaid with an equal volume of mineral oil and a hot start PCR was performed wherein the reaction mixes were heated to 94°C for 5 min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 20µl). Thirty cycles of 94°C for 60 sec., 57°C for 60 sec. and and 72°C for 60 sec. were carried out in a thermocycler followed by removal of the aqueous reaction mixes to separate microtubes. The reaction mixes bearing the amplified target nucleic acids were then treated with exonuclease I to digest the primers not extended in the amplification step. This was achieved by incubating 3µl of the PCR reaction mixes with 0.5 units of exonuclease I at 37°C for 30 min. The exonuclease was subsequently heat inactivated by incubating the reaction at 80°C for 15 min.

[0089] Uracil DNA glycosylase (0.05 units) was then added to 3µl of the above reaction mixes and incubated at 37°C for 30 min. Following Uracil DNA glycosylase treatment, the apyrimidinic sites generated in the amplified products were cleaved to completion by adding NaOH to a final concentration of 0.05M and heating the mixtures for 15 min at 95°C. The reaction mixes were then neutralised by addition of Tris base to 0.03M final concentration. An equal volume of formamide loading dye (90% formamide, 0.025% Bromophenol blue, 0.025% Xylene cylanol) was added to the mixes which were then heated to 85°C for 5 min. 5µl of each mix was then loaded onto a 20% denaturing (7M urea) polyacr-

IDT Exhibit 1020.9610

ylamide gel and electrophoresis was carried out for 3 hours at 400 volts for size analysis of the cleaved products in the sample. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hours at -70°C.

[0090] Analysis of the autoradiographed products showed the following cleavage pattern. In the first reaction where primer A was end-labelled. a product of 32n was observed. The 32n product could only have been generated if an uracil residue was incorporated 12 nucleotides downstream of the 3' end of the primer A. In the second reaction where primer B was end-labelled, a product of 23n was observed. The 23n product could only have been generated if an uracil residue was incorporated 3 nucleotides downstream of the 3' end of primer B.

[0091] The 130bp product was amplified from a *Mycobacterium tuberculosis* complex DNA sample by PCR exactly as outlined above except that primer C (5'catcctgacctaccacgtagt3') and primer D (5'gtcggcgttaccgaccttgag3') were used in place of primer A and primer B. Analysis of the autoradiographed products showed the following cleavage pattern. In the reaction where primer C was end-labelled, a product of 47n was observed. The 47n product could only have been generated if an uracil residue was incorporated 27 nucleotides downstream of the 3' end of primer A. In the reaction where primer D was end-labelled, a product of 23n was observed. The 23n product could only have been generated if an uracil residue was incorporated 3 nucleotides downstream of the 3' end of primer D.

[0092] The amplification process described here can be performed independently with separately labelled primers as described or as a multiplex reaction where all four labelled primers are included in a single amplification reaction. In the latter case, labelled products of 47n, 32n and 23n are detectable in a single lane following cleavage with uracil-DNA-glycosylase, electrophoresis and autoradiography.

[0093] The probability of an uracil being present at a particular location in a DNA fragment is 1/4 or 0.25. The probability of an uracil not being present at a location is 3/4 or 0.75. In the case of primer A, it was extended 11 nucleotides before it was cut. The probability of such an extension occurring randomly is 0.75^{11}=0.042. The extended primer cut at the 12 nucleotide incorporated showing that an uracil was present at this location. The probability of the uracil occurring at that particular position was 0.25. Thus overall, the probability of an eleven base pair extension product occurring randomly would be 0.042 x 0.25 = 0.0105. Similarly, the probability of primer B extending by two bases would be 0.75^2 x 0.25 = 0.14. Overall, the probability of a random 116bp product cleaving to a 32n (11 nucleotide extension of the 21n primer A) and a 23n (2 nucleotide extension of the 21n primer B) product would be 0.0105 x 0.14 = 0.0014. Similar calculations performed for the 130bp product where primer C and D are extended by 26n and 2n respectively show that the probability of such extension products arising randomly is 0.000019. Taking both PCR products together, the probability of the observed products occurring randomly is 0.000000027. Thus, this Example demonstrates the application of the method according to the invention in the verification of PCR products from *Mycobacterium tuberculosis.* Thus the application of the method according to the invention in this manner offers a rapid and accurate means for verification of amplified nucleic acids from diagnostic or other purposes.

Example 8

[0094] The fact that Uracil-DNA-glycosylase is active on both single stranded and double stranded DNA allows the use of this enzyme in the analysis of linear amplification products bearing uracils residues.

[0095] Linear amplification of single stranded DNA can offer significant advantages over exponential DNA amplification in several instances. However, linear amplification suffers from some drawbacks since single stranded DNA generated by linear amplification is generally less amenable to rapid analysis by comparison with double stranded DNA. For example, single stranded DNA is generally not cleaved by restriction enzymes. Furthermore, exponential amplification of genomic DNA using a specific pair of primers in PCR results in the production of a double stranded DNA product of a specific size facilitating analysis by sizing technology. The 5' and 3' ends of the fragment are defined by the forward and reverse primer. By contrast linear amplification of genomic DNA produces fragments with defined 5' ends but with undefined 3' ends. In this Example, we demonstrate the use of the method according to the invention to generate single stranded fragments of a specific size following linear amplification. Amplification of a nucleic acid fragment bearing a candidate locus is usually required to determine if a particular sequence is present or absent at the candidate locus. Often it is necessary or desirable to examine multiple candidate loci in amplified DNA for the presence of particular sequences. This is especially desirable for detection of mutations in genes causing human disease. Thus, any method allowing simultaneous amplification of multiple candidate loci followed by simultaneous analysis of the candidate loci is advantageous by comparison with single locus analysis. Amplification of a nucleic acid fragment(s) bearing multiple candidate loci can be achieved by linear amplification or exponential amplification methods. However. the analysis of the multiple loci is often rate limiting requiring several independent approaches or full scale DNA sequencing. In this Example we also demonstrate the application of the method to simultaneously determine if particular mutations are present or absent at two specific candidate loci.

[0096] The method according to the invention was used to simultaneously detect the presence or absence of a G to A base substitution mutation causal of malig-

IDT Exhibit 1020.9611

nant hyperthermia at the candidate loci 7301 and 7370 in the human skeletal ryanodine receptor gene (RYR1) in two patients heterozygous for either mutation. The sequence of steps is illustrated in Fig. 9. In this case, the target nucleic acid was a DNA fragment amplified by PCR from skeletal muscle cDNA from two patients (A and B) with malignant hyperthermia. The lower strand shown in Fig. 9 is the target strand. Primer A (26n, 5'ggcttggattagatgcatctctggtg3') was designed so that the first uracil incorporated into the extended primer in a linear amplification of the target DNA would be positioned at the sixth nucleotide incorporated in the case of the mutant allele and at the eighth nucleotide in the case of the normal allele. Primer B (28n, 5'gaattccaaggtcctccaagggcacaag3') was designed so that the first uracil incorporated would be positioned at the fifth nucleotide incorporated in the case of the mutant allele and at the twelfth nucleotide in the case of the normal allele. Both primers were extended in the same direction by linear amplification. The presence or absence of a T(U) at the candidate loci was determined by cleavage with uracil-DNA-glycosylase, gel electrophoresis and autoradiography. The target nucleic acid can be genomic DNA or any amplified section of DNA. In this Example, the target nucleic acid was generated by amplifying the 6995 to 7402 region of the RYR1 gene from cDNA using PCR.

[0097] Primers A and B were end-labelled with $\gamma^{32}P$ as described in Example 1. The target nucleic acid from patient A and patient B was treated with an equal volume of exonuclease I (0.5 units/μl) at 37°C for 30 min to degrade primers not utilised in the PCR. The exonuclease was subsequently inactivated by incubation at 80°C for 15 min. This treatment was necessary to prevent any exponential amplification in subsequent steps. This step is not necessary when genomic DNA is the target nucleic acid. 1 μl of the target nucleic acid samples were incubated independently with 6 pmol of labelled primer A and labelled primer B, 0.2 mM dATP, dCTP, dGTP and dUTP, 2mM $MgSO_4$. 10mM KCl, 20mM Tris-HCl pH 8.8, 10mM$(NH_4)_2$ SO_4, 0.1% Triton X-100 in a total volume of 19μl. The reaction mixes were then overlaid with an equal volume of mineral oil and a hot start to the linear amplification process was performed wherein the reaction mix was heated to 94°C for 5 min prior to addition of 0.5 units of Vent DNA Polymerase (exo) (bringing the total volume to 20μl). Thirty cycles of 94°C for 60 sec., 55°C for 60 sec. and 72°C for 60 sec. were carried out in a thermocycler followed by removal of the aqueous reaction mixture to a separate microtube.

[0098] Uracil-DNA-glycosylase (0.05 units) was then added to 5μl of the above reactions and incubated at 37°C for 30 min. Following uracil-DNA-glycosylase treatment, the apyrimidinic sites generated in the amplified product were cleaved to completion by adding NaOH to a final concentration of 0.05M and heating the mixtures for 15 min at 95°C. The reactions were then neutralised by addition of Tris base to 0.03M final concentration. An equal volume of formamide loading dye (90% formamide. 0.025% bromophenol blue. 0.025% xylene cylanol) was added to the samples which were then heated to 85°C for 5 min. 5μl of the samples were then loaded onto a 20% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3 hours at 400 volts for size analysis of the cleaved products in the sample. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hours at -70°C.

[0099] Analysis of the sample from patient A showed cleavage products of 39n, 33n and 31 n in addition to the 26n and 28n non utilised primers. This pattern of cleavage products could only be produced if patient A was homozygous for the normal allele of the RYR1 gene at candidate locus 7370 and heterozygous for the G7301A mutation. Analysis of the sample from patient B showed cleavage products of 39n, 33n and 32n in addition to the 26n and 28n non utilised primers. This pattern of cleavage products could only be produced if patient B was homozygous for the normal allele of the RYR 1 gene at candidate locus 7301 and heterozygous for the G7370A mutation. Thus, the application of the method according to the invention has permitted simultaneous analysis of two candidate loci for the presence or absence of specific mutations. In this Example, two candidate loci have been investigated. The method can readily be applied to multiple candidate loci in genomic or amplified target DNA as long as the primers chosen support linear amplification but do not permit exponential amplification.

[0100] Generation of amplified nucleic acid fragments of specific size is regularly used in diagnosis of infectious agents and in the verification of amplified nucleic acid molecules. In this Example, the use of the method according to the invention generated single stranded fragments of a specific size following linear amplification. This is a useful application of the method since it permits the use of linear amplification rather than exponential amplification to generate amplified fragments of a specific size.

Claims

1. A method for rapidly detecting the presence or absence of a particular nucleic acid sequence at a candidate locus in a target nucleic acid sample which comprises the steps of:

   i) introducing a modified base which is a substrate for a DNA glycosylase into said candidate locus at one or more preselected positions;

   ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site;

IDT Exhibit 1020.9612

iii) cleaving phosphate linkages at abasic sites generated in step ii); and

iv) analysing the cleavage products of step iii) so as to identify in said target nucleic acid sequence the presence or absence of said particular nucleic acid sequence at said candidate locus.

2. A method according to Claim 1, wherein the candidate locus is amplified using a combination of normal DNA precursor nucleotides and at least one modified precursor nucleotide.

3. A method according to Claim 2, wherein at least one of the primers for the amplification is positioned adjacent to the candidate locus.

4. A method according to Claim 3, wherein the at least one primer is labelled.

5. A method according to any one of Claims 2-4, wherein at least one of the precursor nucleotide(s) is labelled.

6. A method according to any preceding claim, wherein the modified base is introduced by chemical modification of an existing base at the candidate locus.

7. A method according to any one of Claims 1-6, wherein the modified base is excised by means of a DNA glycosylase enzyme.

8. A method according to Claim 7, wherein the DNA-glycosylase enzyme is uracil DNA-glycosylase.

9. A method according to Claim 7 or 8, wherein the substrate for the DNA glycosylase enzyme is immobilised.

10. A method according to any preceding claim, wherein the phosphate linkages at the abasic sites are cleaved by treatment with alkali.

11. A method according to any one of Claims 1-9, wherein the phosphate linkages at the abasic sites are cleaved by a heat treatment.

12. A method according to any one of Claims 1-9, wherein the phosphate linkages at the abasic sites are cleaved by treatment with an enzyme.

13. A method according to Claim 12, wherein the enzyme is an enzyme which cleaves specifically at apurinic or apyrimidinic sites.

14. A method according to any preceding claim, wherein the modified base is uracil or hypoxanthine.

15. A method according to any preceding claim, wherein the target nucleic acid sample is DNA.

16. A method according to Claim 15, wherein the DNA is single stranded, homoduplex or heteroduplex DNA.

17. A method according to any one of Claims 1-14, wherein the target nucleic acid sample is RNA.

18. A method according to Claim 17, when dependent on Claim 2, wherein the RNA is converted into a cDNA by reverse transcription prior to the amplification step.

19. A method according to any preceding claim, wherein the cleavage products are detected by hybridisation with specific nucleic acid probes.

Patentansprüche

1. Verfahren zum schnellen Detektieren der Anwesenheit oder Abwesenheit einer bestimmten Nukleinsäuresequenz in einem Kandidaten-Lokus (candidate locus) in einer Zielnukleinsäurenprobe, das die Schritte umfasst:

   (i) Einführen einer modifizierten Base, die ein Substrat für eine DNA-Glycosylase ist, in den Kandidaten-Lokus an einer oder mehrerer vorher ausgewählter Positionen;
   (ii) Ausschneiden der modifizierten Base mit Hilfe der DNA-Glycosylase, um eine abasische Stelle zu erzeugen;
   (iii) Spalten der Phosphatbrücken an den in Schritt (ii) erzeugten abasischen Stellen; und
   (iv) Analysieren der Spaltungsprodukte von Schritt (iii), um in der Zielnukleinsäuresequenz die Anwesenheit oder Abwesenheit der bestimmten Nukleinsäuresequenz in dem Kandidaten-Lokus zu identifizieren.

2. Verfahren nach Anspruch 1, bei dem der Kandidaten-Lokus unter Verwendung einer Kombination von normalen DNA-Vorläufemukleotiden und mindestens einem modifizierten Vorläufernukleotid amplifiziert wird.

3. Verfahren nach Anspruch 2, bei dem mindestens einer der Primer für die Amplifikation benachbart zu dem Kandidaten-Lokus positioniert ist.

4. Verfahren nach Anspruch 3, bei dem mindestens ein Primer markiert ist.

5. Verfahren nach einem der Ansprüche 2 bis 4, bei dem mindestens eines der Vorläufernukleotide

IDT Exhibit 1020.9613

markiert ist.

6. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die modifizierte Base durch chemische Modifikation einer bestehenden Base in den Kandidaten-Lokus eingeführt wird.

7. Verfahren nach einem der Ansprüche 1 bis 6, bei dem die modifizierte Base mittels eines DNA-Glycosylaseenzyms ausgeschnitten wird.

8. Verfahren nach Anspruch 7, bei dem das DNA Glycosylaseenzym Uracil-DNA Glycosylase ist.

9. Verfahren nach Anspruch 7 oder 8, bei dem das Substrat für das DNA Glycosylaseenzym immobilisiert ist.

10. Verfahren nach einem vorhergehenden Ansprüche, bei dem die Phosphatbrücken an den abasischen Stellen durch Behandlung mit Alkali gespalten werden.

11. Verfahren nach einem der Ansprüche 1 bis 9, bei dem die Phosphatbrücken an den abasischen Stellen durch eine Hitzebehandlung gespalten werden.

12. Verfahren nach einem der Ansprüche 1 bis 9, bei dem die Phosphatbrücken an den abasischen Stellen durch Behandlung mit einem Enzym gespalten werden.

13. Verfahren nach Anspruch 12, bei dem das Enzym ein Enzym ist, das spezifisch bei apurinen oder apyrimidinen Stellen spaltet.

14. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die modifizierte Base Uracil oder Hypoxanthin ist.

15. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Zielnukleinsäureprobe DNA ist.

16. Verfahren nach Anspruch 15, bei dem die DNA einzelsträngige, Homoduplexoder Heteroduplex-DNA ist.

17. Verfahren nach einem der Ansprüche 1 bis 14, bei dem die Zielnukleinsäureprobe RNA ist.

18. Verfahren nach Anspruch 17, wenn abhängig von Anspruch 2, bei dem die RNA vor dem Amplifizierungsschritt in eine cDNA durch reverse Transkription überführt wird.

19. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Spaltprodukte durch Hybridisierung mit spezifischen Nukleinsäureproben detektiert werden.

Revendications

1. Méthode permettant de détecter rapidement la présence ou l'absence d'une séquence d'acide nucléique donnée au niveau d'un locus présumé dans un échantillon d'acide nucléique cible, qui comprend les étapes consistant:

   i) à introduire dans ledit locus présumé, dans une ou plusieurs positions pré-sélectionnées, une base modifiée qui est un substrat pour une ADN glycosylase ;

   ii) à exciser la base modifiée au moyen de ladite ADN glycosylase de façon à générer un site abasique;

   iii) à cliver les liaisons phosphate au niveau des sites abasiques générés dans l'étape ii) ;et

   iv) à analyser les produits de clivage de l'étape iii) de façon à identifier dans ladite séquence de l'acide nucléique cible la présence ou l'absence de ladite séquence d'acide nucléique donnée au niveau dudit locus présumé.

2. Méthode selon la revendication 1, dans laquelle le locus présumé est amplifié à l'aide d'une combinaison de nucléotides précurseurs d'ADN normaux et d'au moins un nucléotide précurseur modifié.

3. Méthode selon la revendication 2, dans laquelle au moins l'une des amorces de l'amplification est disposée de façon adjacente au locus présumé.

4. Méthode selon la revendication 3, dans laquelle au moins l'une des amorces est marquée.

5. Méthode selon l'une quelconque des revendications 2 à 4, dans laquelle au moins l'un des nucléotides précurseurs est marqué.

6. Méthode selon l'une quelconque des revendications précédentes, dans laquelle la base modifiée est introduite par modification chimique d'une base existante au niveau du locus présumé.

7. Méthode selon l'une quelconque des revendications 1 à 6, dans laquelle la base modifiée est excisée à l'aide d'une enzyme ADN glycosylase.

8. Méthode selon la revendication 7, dans laquelle l'enzyme ADN glycosylase est l'ADN glycosylase uracile.

IDT Exhibit 1020.9614

9. Méthode selon la revendication 7 ou 8, dans laquelle le substrat de l'enzyme ADN glycosylase est immobilisée.

10. Méthode selon l'une quelconque des revendications précédentes, dans laquelle les liaisons phosphate au niveau des sites abasiques sont clivées par traitement avec un alcali.

11. Méthode selon l'une quelconque des revendications 1 à 9, dans laquelle les liaisons phosphate au niveau des sites abasiques sont clivées par traitement thermique.

12. Méthode selon l'une quelconque des revendications 1 à 9, dans laquelle les liaisons phosphate au niveau des sites abasiques sont clivées par traitement avec une enzyme.

13. Méthode selon la revendication 12, dans laquelle l'enzyme est une enzyme qui effectue un clivage spécifique au niveau des sites apuriniques ou apyrimidiniques.

14. Méthode selon l'une quelconque des revendications précédentes, dans laquelle la base modifiée est l'uracile ou l'hypoxanthine.

15. Méthode selon l'une quelconque des revendications précédentes, dans laquelle l'échantillon d'acide nucléique cible est de l'ADN.

16. Méthode selon la revendication 15, dans laquelle l'ADN est un ADN simple brin, homoduplex ou hétéroduplex.

17. Méthode selon l'une quelconque des revendications 1 à 14, dans laquelle l'échantillon d'acide nucléique cible est de l'ARN.

18. Méthode selon la revendication 17 lorsque celle-ci est dépendante de la revendication 2, dans laquelle l'ARN est transformé en un ADNc par transcription inverse avant l'étape d'amplification.

19. Méthode selon l'une quelconque des revendications précédentes, dans laquelle les produits de clivage sont détectés par hybridation avec des sondes d'acides nucléiques spécifiques.

IDT Exhibit 1020.9615

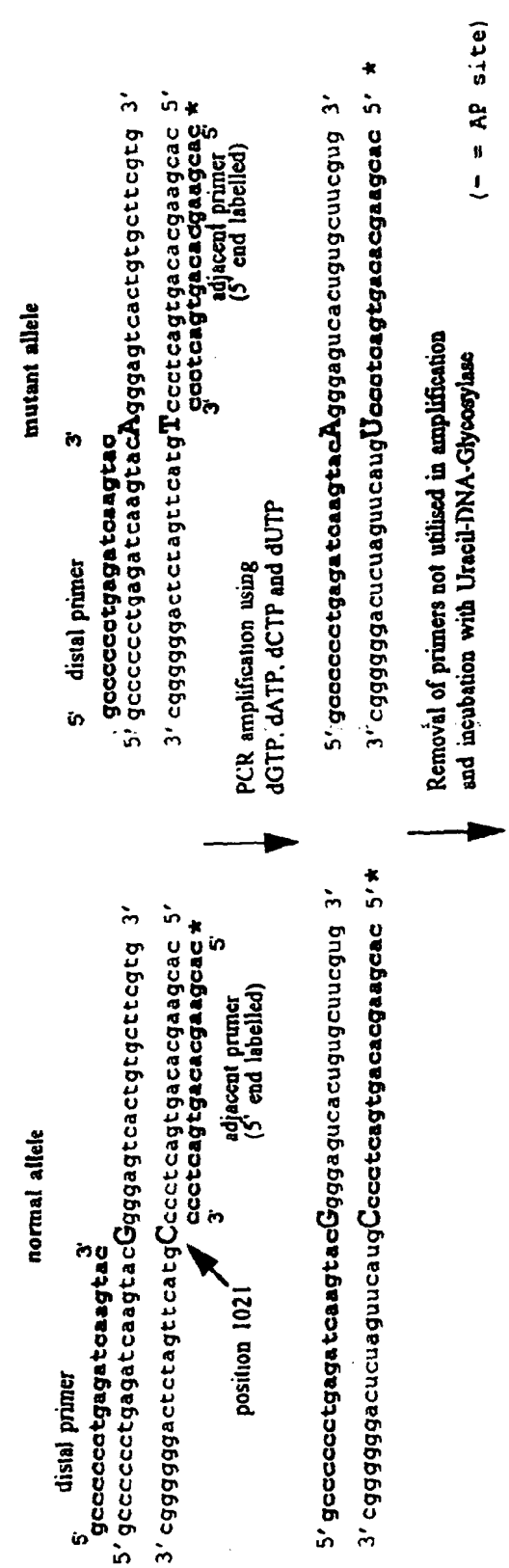


Fig.1

IDT Exhibit 1020.9616

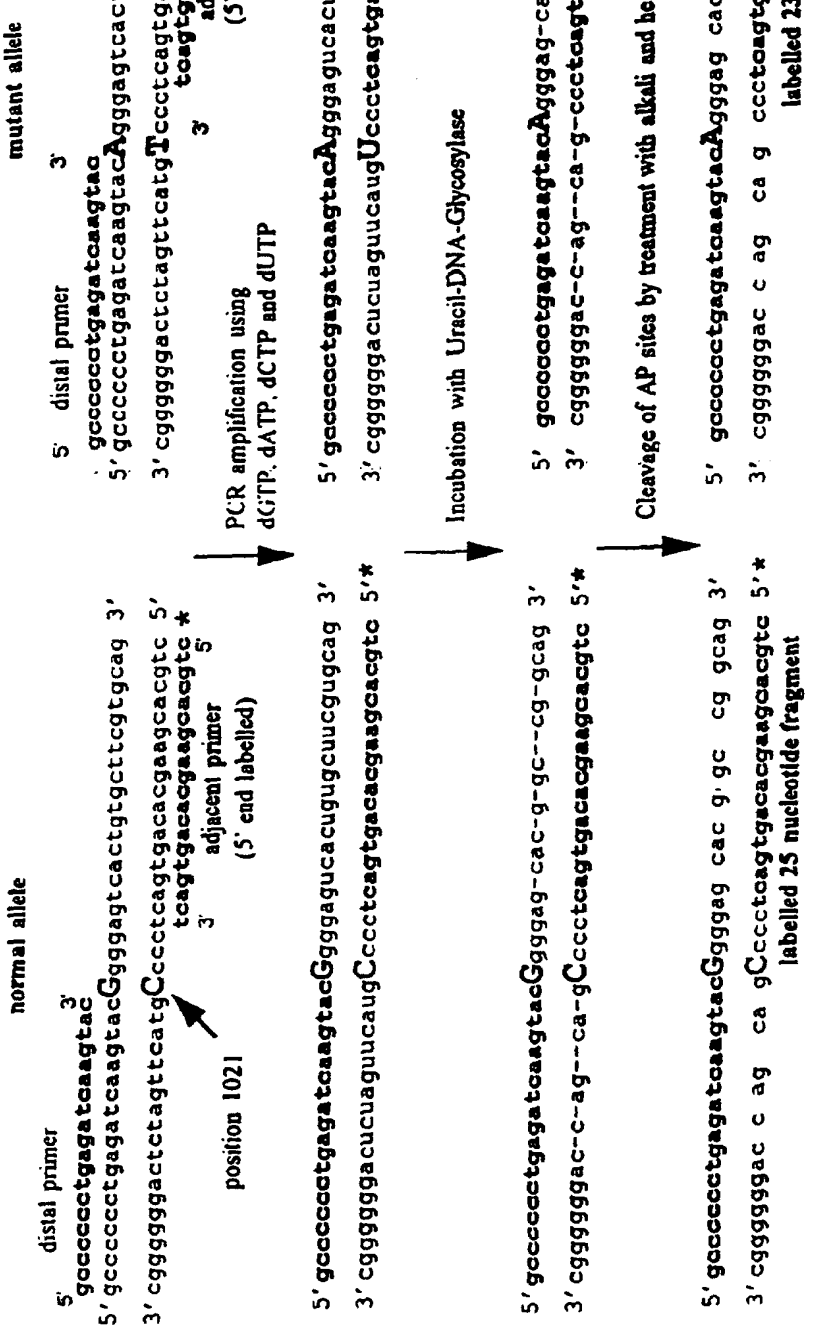


Fig.2

IDT Exhibit 1020.9617

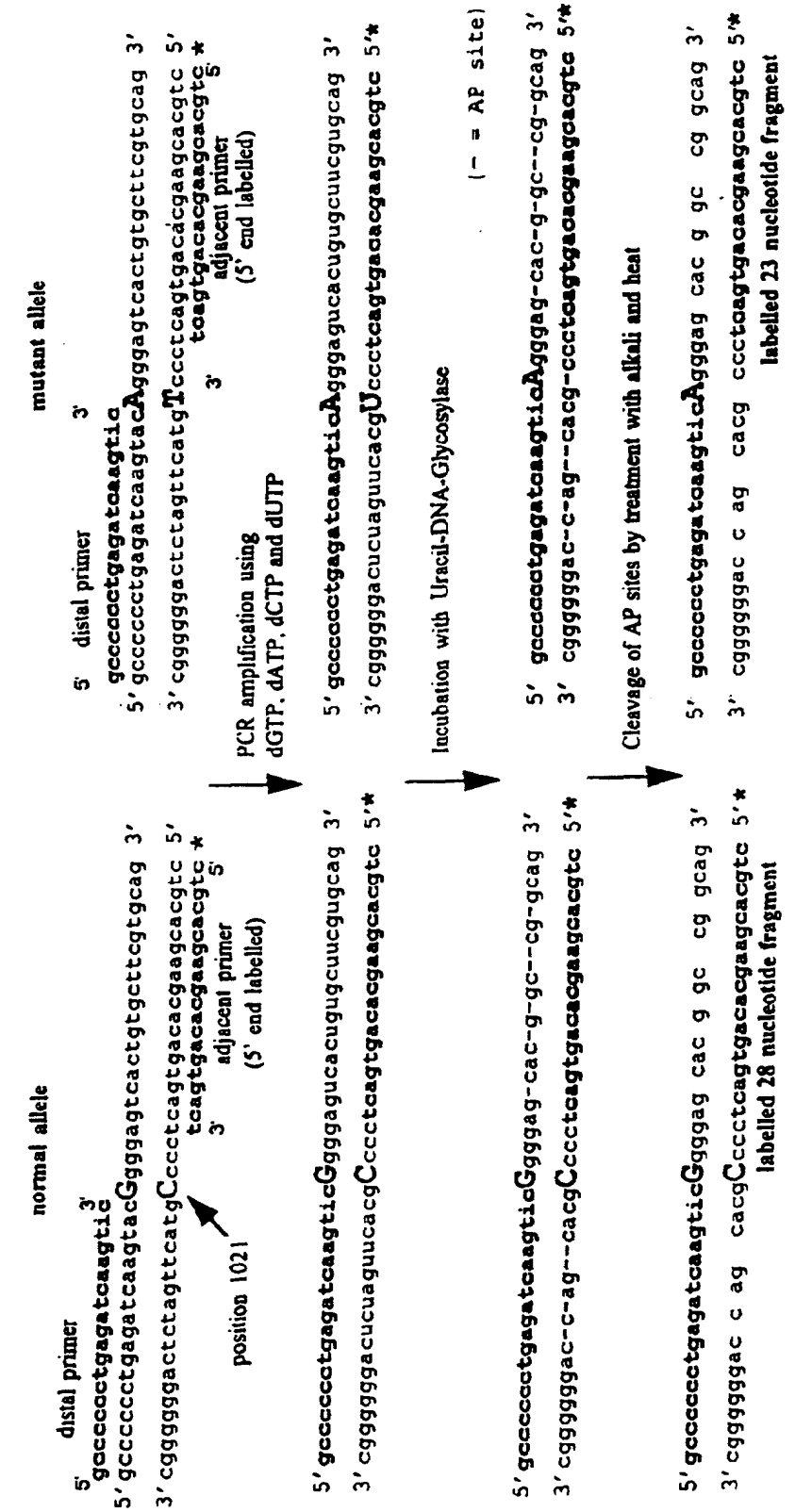


Fig.3

IDT Exhibit 1020.9618

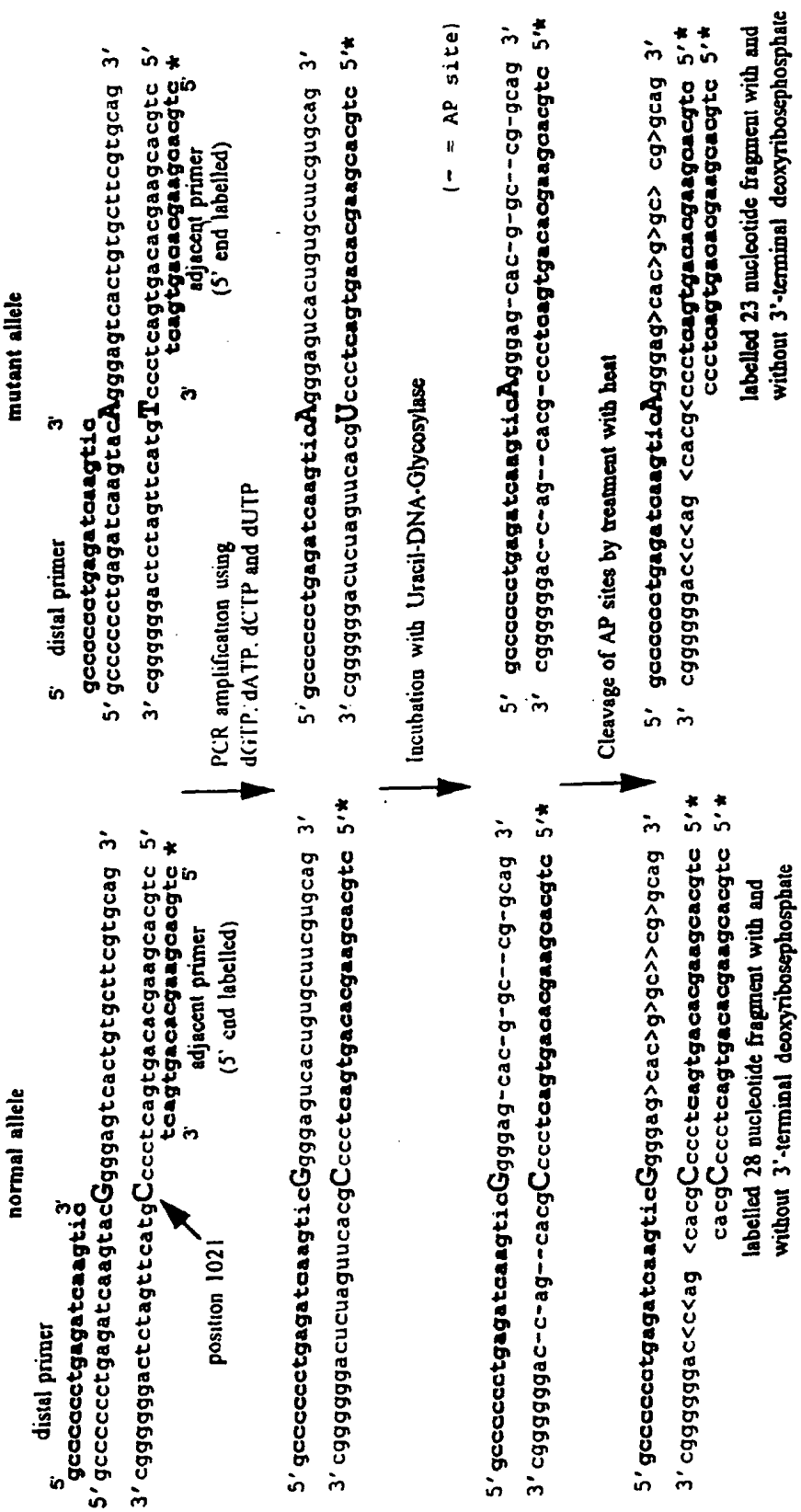


Fig.4

IDT Exhibit 1020.9619

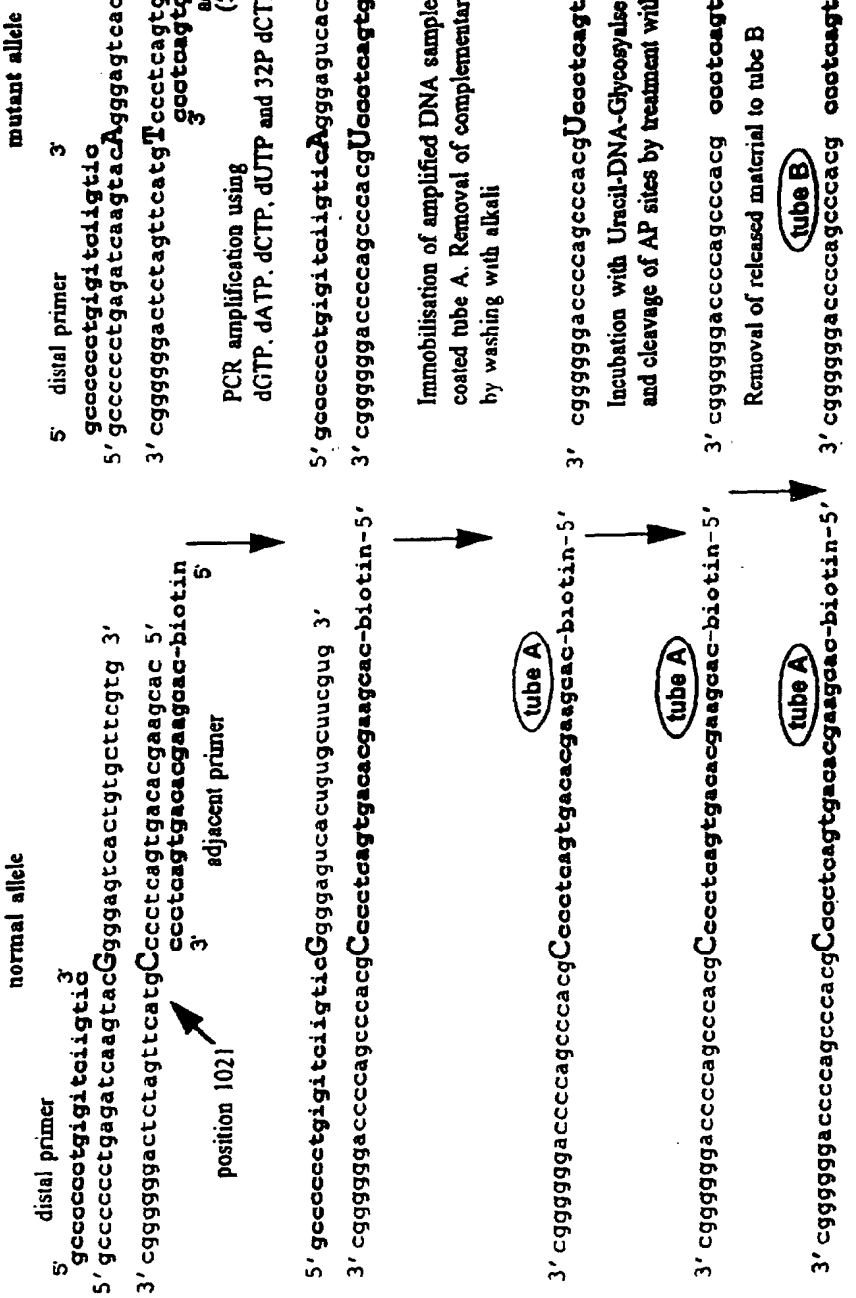


Fig.5

IDT Exhibit 1020.9620

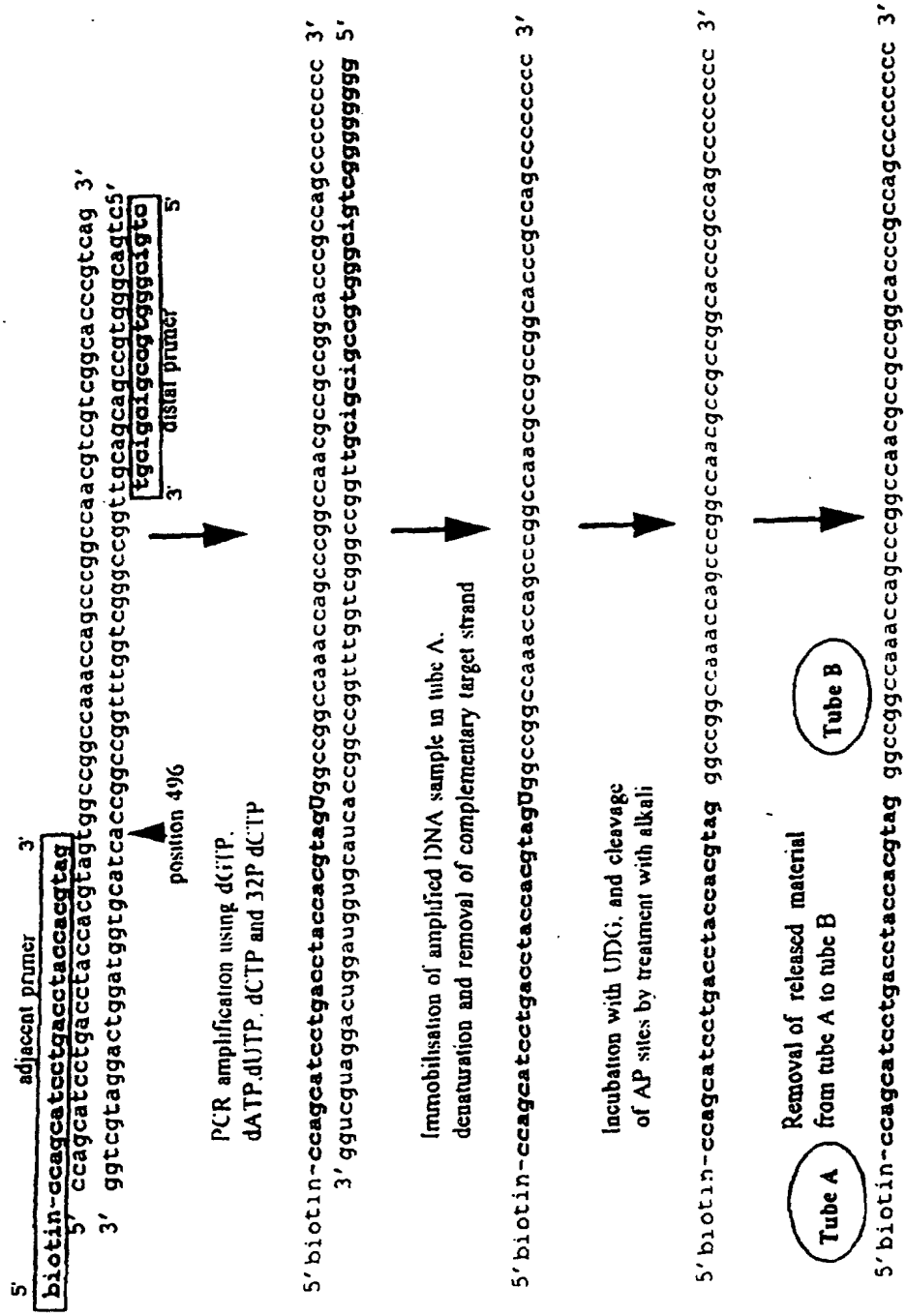


Fig.6

IDT Exhibit 1020.9621

Fig.7

IDT Exhibit 1020.9622

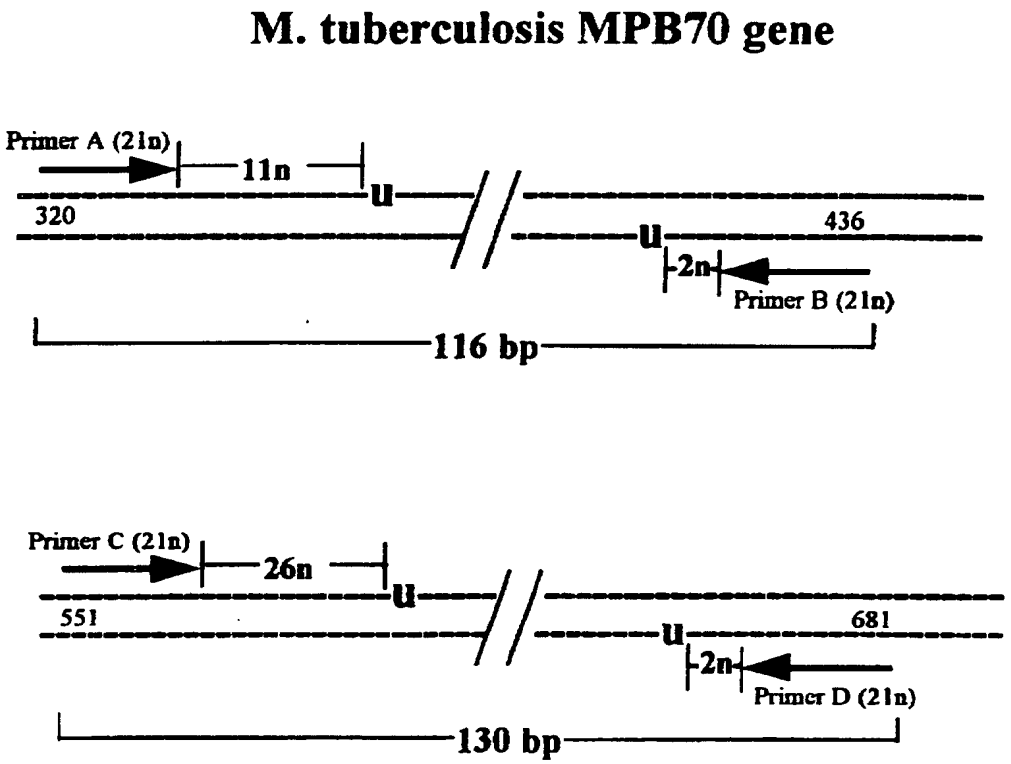


Fig.8

IDT Exhibit 1020.9623

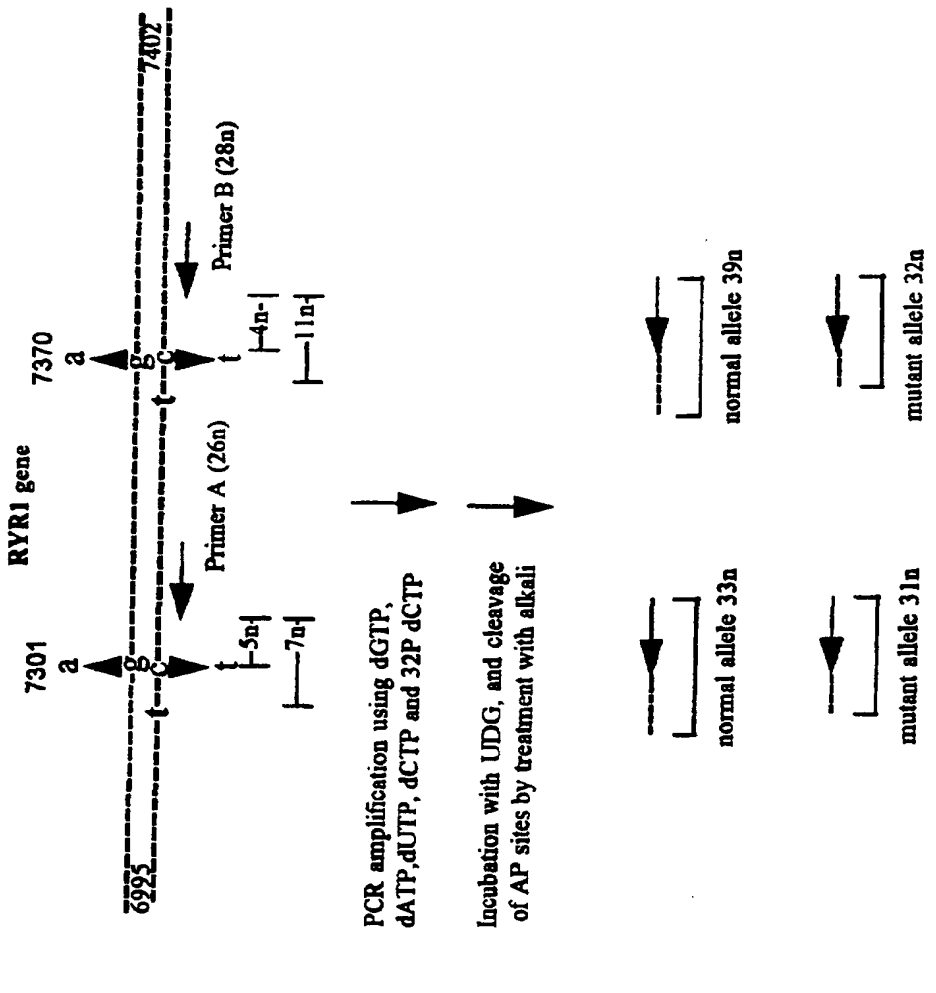


Fig. 9

IDT Exhibit 1020.9624

Espacenet

Bibliographic data: WO0229117 (A2) — 2002-04-11

METHODS AND PROBES FOR DETECTION AND/OR QUANTIFICATION OF NUCLEIC ACID SEQUENCES

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

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

Classification: - international: ***C07H21/02; C07H21/04; C12N15/09; C12P19/34; C12Q1/68; G01N33/53; G01N33/566;*** (IPC1-7): C12Q1/68
- cooperative: **C12Q1/6818** more

Application number: WO2001US31698 20011009 Global Dossier

Priority number (s): US20000238850P 20001006

Also published as: WO0229117 (A3) AU1309602 (A) AU2002213096 (B2) CA2423729 (A1) EP1356104 (A2) more

Abstract of WO0229117 (A2)

The present invention discloses nucleic acid detector probes for specific detection and/or quantification of target nucleic acid sequences and detection and/or quantification methods using these probes. In the absence of target nucleic acid sequence, a first oligonucleotide and a third oligonucleotide are bound to each other in a conformation which brings two member of an interacting moiety pair (labels) into close spatial proximity. Cooperative binding of the first oligonucleotide and a second oligonucleotide to a target nucleic acid sequence causes displacement of the third oligonucleotide from the first oligonucleotide probe resulting in separation of the two members of the interacting moiety pair (labels). The spatial separation of the moieties (labels) is detectable, and indicates the presence and/or amount of the target nucleic acid sequence. The method is useful for detection and/or quantification of a specific nucleic acid sequence as well as the detection of sequence alteration in the target nucleic acid sequence.

IDT Exhibit 1020.9625

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/29117 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US01/31698

(22) International Filing Date: 9 October 2001 (09.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/238,850 6 October 2000 (06.10.2000) US

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

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **KURN, Nurith** [US/US]; 2876 Ramona Street, Palo Alto, CA 94306 (US).

(74) Agent: **POLIZZI, Catherine, M.**; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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

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.

WO 02/29117 A2

(54) Title: METHODS AND PROBES FOR DETECTION AND/OR QUANTIFICATION OF NUCLEIC ACID SEQUENCES

(57) Abstract: The present invention discloses nucleic acid detector probes for specific detection and/or quantification of target nucleic acid sequences and detection and/or quantification methods using these probes. In the absence of target nucleic acid sequence, a first oligonucleotide and a third oligonucleotide are bound to each other in a conformation which brings two member of an interacting moiety pair (labels) into close spatial proximity. Cooperative binding of the first oligonucleotide and a second oligonucleotide to a target nucleic acid sequence causes displacement of the third oligonucleotide from the first oligonucleotide probe resulting in separation of the two members of the interacting moiety pair (labels). The spatial separation of the moieties (labels) is detectable, and indicates the presence and/or amount of the target nucleic acid sequence. The method is useful for detection and/or quantification of a specific nucleic acid sequence as well as the detection of sequence alteration in the target nucleic acid sequence.

IDT Exhibit 1020.9626

METHODS AND PROBES FOR DETECTION AND/OR QUANTIFICATION OF NUCLEIC ACID SEQUENCES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the priority benefit of the provisional patent application U.S. Serial No. 60/238,850, filed October 6, 2000, which is incorporated by reference in its entirety.

TECHNICAL FIELD

The invention relates to detection and/or quantification of nucleic acid sequences using nucleic acid detector probes.

BACKGROUND ART

The development of nucleic acid amplification methods contributed to the rapid advances in genome research, biotechnology, molecular medicine, and the like. Rapid methods for the detection and quantification of specific nucleic acid sequences, either directly or following amplification, are required for nucleic acid analysis as used in various research, clinical and industrial applications. Various homogenous or heterogeneous methods for the detection of binding of the one or more labels to the target sequence were described.

Nucleic acid hybridization is commonly used for the detection of specific sequences. Methods for the detection and quantification of nucleic acid sequences, which utilize hybridization of one or more oligonucleotide probes to a specific target sequence, usually employ means of detection of hybrid formation. Some methods utilize one or more labeled probes that are complementary to a specific region of the target nucleic acid sequence, and means of detecting binding of the one or more labels to the target sequence.

Some of the previously described detection methods employ enzyme-catalyzed reactions. These methods are based on detection of the formation of hybrids of the specific target nucleic acid and one or more oligonucleotide probes, which may be labeled. Both enzymatic cleavage and ligation of specific oligonucleotide probes hybridized to target nucleic acid sequence have been described. The enzymatic reaction products are detected directly, for example, by changes of optical properties of the labels, or by their capacity to

induce additional enzymatic reaction or become substrates for additional enzymes. The enzyme-based detection methods may be employed either for the direct detection of specific nucleic acid sequences, or may be employed for the detection of amplification products. Some of the commonly used homogeneous detection methods employing enzymes which cleave oligonucleotide probes when bound to the target nucleic acid sequence by hybridization, are CYCLING PROBE METHOD™, TAQMAN™, INVADER™, and the like. Many of these methods employ an oligonucleotide which is labeled with a pair of a fluorescent and a quencher labels, which are attached to the probe at the correct distance to ensure energy transfer from the fluorescent label to the quencher label. Cleavage of one of the probes when hybridized to the target nucleic acid sequence results in separation of the two labels which results in the generation of fluorescent signal. Various modifications of this principle have been described (Nurmi et al., 2000, Nucleic acid Res. 28, 28e; US Patent No. 5,403,711 (Walder); Nadeau et al., 1999, Anal Biochem 276: 177).

Other methods employ means of direct, non-enzymatic, detection of one or more probe-target hybrid formation. Both homogeneous and heterogeneous detection methods have been described. Heterogeneous detection methods that require extensive wash steps are inherently more complicated and are less desirable. Homogeneous detection methods are faster and are better suited for applications requiring high throughput. Hybridization of one or more labeled probes to a specific target nucleic acid sequence may result in changes in the optical properties of the one or more labeled probes. Alternatively, the association of two or more labels in a stable complex, as a result of hybridization of two or more oligonucleotide probes to a single nucleic acid target, can be detected.

Homogeneous detection of hybrid formation based on the separation of donor-acceptor label molecules has been described recently. A method utilizing a change in probe conformation upon hybridization to a target nucleic acid sequence thereby changing the distance of donor acceptor label molecules, was described by Tyagi et al. (U.S. Pat. No. 5,925,517). The probe is composed of a stem loop structure, which brings the donor-acceptor pair of labels attached to the probe in close proximity to effect fluorescence quenching. Hybridization of the probe to a specific target nucleic acid sequence results in a conformational change which is characterized by dissociation of the stem structure thereby increasing the distance between the two labels and preventing energy transfer between the two labels. Thus, hybridization of the oligonucleotide to the specific target nucleic acid sequence is detectable by a fluorescence signal of the donor label. Other variations on this

IDT Exhibit 1020.9628

concept have also been disclosed. Detection of nucleic acid by fluorescence quenching was also described by Bruce et al (EP patent application no. 98117883.3). The detector oligonucleotide probe comprises at least two oligonucleotides, which hybridize to form a partially double stranded detector. Acceptor and donor dyes are attached to the detector oligonucleotides so as to place them at close proximity to enable energy transfer. Upon hybridization of the single stranded portion of the detector probe to the target nucleic acid sequence the second oligonucleotide probe is displaced from the first oligonucleotide probe, thus increasing the distance between the donor and acceptor dye labels, causing a change in fluorescence that is detectable.

There is a need for improved nucleic acid detection and/or quantification methods. The invention provided herein fulfills this need and also provides additional benefits. These include the ease of preparation of the detector probes. The design of the interacting sequences of the detector probes are independent of the target sequence and could thus be universal. Insofar as the detector probes are not unimolecular, the design complexities of a stem loop probe are eliminated.

All references cited herein, including patent applications and publications, are incorporated by reference in their entirety.

IDT Exhibit 1020.9629

DISCLOSURE OF THE INVENTION

In one aspect of the present invention, detector oligonucleotide probes which are useful for the detection of a specific nucleic acid sequence are provided. A first oligonucleotide probe comprises a 3'-region that is hybridizable to a sequence of the target nucleic acid; a 5'-region that is not hybridizable to the target nucleic acid (under otherwise same hybridization conditions as described herein), and a label (F) that is attached to the 5'-end. A second oligonucleotide probe comprises a 5'-region, which is hybridizable to a sequence of the target nucleic acid and a 3'-region, which is not hybridizable to the target nucleic acid (under otherwise same hybridization conditions as described herein), and is hybridizable to part of the sequence of the first oligonucleotide which is not hybridizable to the target nucleic acid. A third oligonucleotide probe comprises a sequence which is not hybridizable to the target nucleic acid sequence and is hybridizable to the 5'-most sequence of the first oligonucleotide, and a label Q that is attached to its 3'-end. In some embodiments, probes with the mirror image design of the preceding oligonucleotide probes are provided. In some embodiments, a single oligonucleotide probe comprises both labels and the functions of either the first and third oligonucleotide probe, or the second and third oligonucleotide probe. Mirror images (i.e., the opposite polarity) of the probes are also provided.

In another aspect of the present invention, methods for detecting and/or quantifying nucleic acid sequences are provided using the probes described above. Detection of a target nucleic acid according to the methods of the invention comprises a) combining a sample suspected of containing said target nucleic acid sequence with a mixture containing the probes described above; optionally b) treating the mixture to render the target nucleic acid single stranded (if not already single stranded); and c) incubating the mixture under conditions which are suitable for binding of the oligonucleotide probes to said single stranded nucleic acid target, wherein binding of the first and second probes to the nucleic acid target results in the displacement of the third probe from the first probe, wherein said displacement results in the spatial separation of the labels of the first probe and the third probe, whereby a detectable and/or quantifiable signal is generated.

In another aspect of the present invention, a method of determining whether a target nucleic acid sequence is present in a sample is provided, comprising: contacting said sample with a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present,

IDT Exhibit 1020.9630

and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein: the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present; and the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence. This aspect may further comprise extension of the 3' end of the second probe along a sequence in the 5' region of the first probe using a nucleotide polymerase.

In a further aspect of the present invention is provided a method of determining whether a target nucleic acid sequence is present in a sample, comprising: contacting said sample with a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein the first probe comprises a polynucleotide comprising a 5' region which hybridizes to a first region of the target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present; and the third probe comprises a polynucleotide which hybridizes to a

IDT Exhibit 1020.9631

sequence in the 3' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

In the methods of the invention, the first and second regions of the target nucleic acid sequence may be contiguous, 1 nucleotide apart, 3 or fewer nucleotides apart, or 7 or fewer nucleotides apart. In addition, the interacting label pair may comprise donor and acceptor moieties, or the interacting label pair may comprises an enzyme, such as an enzyme and an inhibitor of said enzyme, or the interacting label pair may be capable of energy transfer, for example the interacting label pair may comprise a fluorophore and a quencher.

The methods may further comprise measuring the magnitude of the signal generated, whereby said magnitude indicates the quantity of the target nucleic acid sequence.

The target nucleic acid sequence may be attached to an analyte, or to a solid support. The target nucleic acid sequence may comprise DNA, RNA, DNA and RNA, or PNA. At least one of the probes may comprise DNA, RNA, DNA and RNA or PNA. The region of the second probe which is hybridizable to a sequence of the first probe may comprise a modified nucleotide that causes enhanced affinity to the sequence in the region of the first probe relative to an unmodified nucleotide. In one aspect, the detectable signal is of a greater magnitude than a detectable signal associated with the interacting label pair when the third probe is hybridized to the first probe. In another aspect, the detectable signal is of a lesser magnitude than a detectable signal associated with the interacting label pair when the third probe is hybridized to the first probe. In a further aspect, the method further comprises amplifying the target nucleic acid sequence. In yet a further aspect, the first probe or the second probe is allele-specific. In a still further aspect, the first probe and the second probe are allele-specific.

IDT Exhibit 1020.9632

In a further aspect, the invention provides a method of determining whether a target nucleic acid sequence is present in a sample, comprising: contacting said sample with a first probe, a second probe, a third probe, and a nucleotide polymerase, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and allowing nucleic acid polymerization, wherein: the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present, a first member of an interacting label pair; and the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present; and the third probe comprises a polynucleotide sequence which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence; and wherein the 3' end of said second probe is extended by said nucleotide polymerase along the first probe by template switching, causing dissociation of the first and second members of the interacting label pair; whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

In a further aspect, the invention provides a method of determining whether a target nucleic acid sequence is present in a sample, said method comprising: contacting said sample with a first probe and a second probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of a first nucleotide sequence in the 5' region of the first probe with a second nucleotide sequence in the 5' region of the first probe, in the absence of the target nucleic acid sequence, and hybridization of at the second probe the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein: the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present, and a 5' region comprising said first and second nucleotide sequences which hybridize to each other in the absence of the target nucleic acid sequence, and two members of an interacting label pair; the second probe comprises a

IDT Exhibit 1020.9633

polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target polynucleotide is present; wherein, when said first and second nucleotide sequences in the 5' region of the first probe hybridize, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said first probe and said second probe hybridize to the target nucleic acid sequence and said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

In yet a further aspect, the invention provides a method of determining whether a target nucleic acid sequence is present in a sample, said method comprising: contacting said sample with a first probe and a second probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of a first nucleotide sequence in the 3' region of the first probe with a second nucleotide sequence in the 3' region of the first probe, in the absence of the target nucleic acid sequence, and hybridization of at the second probe the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein: the first probe comprises a polynucleotide comprising a 5' region which is hybridizes to a first region of the target nucleic acid sequence, if present, and a 3' region comprising said first and second nucleotide sequences which hybridize to each other in the absence of the target nucleic acid sequence, and two members of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target polynucleotide is present; wherein, when said first and second nucleotide sequences in the 3' region of the first probe hybridize, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said first probe and said second probe hybridize to the target nucleic acid sequence and said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

IDT Exhibit 1020.9634

In yet another aspect of the present invention, methods for detecting changes, or mutations, in a sequence of the target nucleic acid relative to a reference nucleic acid sequence (unaltered target sequence), are provided. In one embodiment, detection of altered sequence in a target nucleic acid according to the methods of the invention comprises a) combining a test sample suspected of containing said target nucleic acid sequence (with suspected altered target sequence) with a mixture containing the probes described above, wherein the first and/or third oligonucleotide probes comprise a sequence hybridizable to the unaltered (reference) target sequence; optionally b) treating the mixture to render the target nucleic acid single stranded (if not already single stranded); and c) incubating the mixture under conditions which are suitable for binding of the oligonucleotide probes to said single stranded nucleic acid target, wherein binding of the first and second probes to the nucleic acid target results in the displacement of the third probe from the first probe, wherein said displacement results in the spatial separation of the labels of the first probe and the third probe, whereby a detectable and/or quantifiable signal is generated. In this embodiment, reduction or elimination of detectable signal in the test sample (relative to a reference sample comprising unaltered target (reference) sequence) indicates reduced or no binding of the probe(s) to the nucleic acid target in the test sample, which indicates presence of the altered sequence in the test sample. In another embodiment, detection of altered sequence in a target nucleic acid according to the methods of the invention comprises a) combining a test sample suspected of containing said target nucleic acid sequence (with suspected altered target sequence) with a mixture containing the probes described above, wherein the first and/or third oligonucleotide probes comprise a sequence hybridizable to the suspected altered target sequence; optionally b) treating the mixture to render the target nucleic acid single stranded (if not already single stranded); and c) incubating the mixture under conditions which are suitable for binding of the oligonucleotide probes to said single stranded nucleic acid target, wherein binding of the first and second probes to the nucleic acid target results in the displacement of the third probe from the first probe, wherein said displacement results in the spatial separation of the labels of the first probe and the third probe, whereby a detectable and/or quantifiable signal is generated. In this embodiment, a greater amount of detectable signal in the test sample (relative to a reference sample comprising unaltered (reference) sequence) indicates more binding of the probe(s) to the nucleic acid target in the test sample, which indicates presence of the altered sequence in the test sample.

IDT Exhibit 1020.9635

A further aspect of the present invention provides a method of determining whether a target nucleic acid sequence contains a sequence alteration relative to a reference nucleic acid sequence, said method comprising: contacting said target nucleic acid sequence with a first probe, a second probe, and a third probe, and contacting said reference nucleic acid sequence with a first probe, a second probe, and a third probe, wherein said contact of the target nucleic acid sequence with the first, second, and third probes, and said contact of the reference nucleic acid sequence with the first, second, and third probes, occur under conditions allowing hybridization of the first and second probes to the reference nucleic acid sequence, and hybridization of the first probe to the third probe, in the absence of the reference nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the reference nucleic acid sequence, wherein the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of the reference nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the reference nucleic acid sequence, if present, and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the reference nucleic acid sequence is present; and the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the reference nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the reference nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the reference nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; and comparing detectable signal generated by contacting said first, second, and third probes with the reference nucleic acid sequence with detectable signal generated by contacting said first, second, and third probes with target nucleic acid sequence; whereby reduced signal generation by the target nucleic acid sequence as compared to the reference nucleic acid sequence indicates the presence of an altered sequence in the target nucleic acid sequence relative to the reference nucleic acid sequence.

In yet a further aspect, the present invention provides a method of determining whether a target nucleic acid sequence contains a sequence alteration relative to a reference

IDT Exhibit 1020.9636

nucleic acid sequence, said method comprising: contacting said target nucleic acid sequence with a first probe, a second probe, and a third probe, and contacting said reference nucleic acid sequence with a first probe, a second probe, and a third probe, wherein said contact of the target nucleic acid sequence with the first, second, and third probes, and said contact of the reference nucleic acid sequence with the first, second, and third probes, occur under conditions allowing hybridization of the first and second probes to the reference nucleic acid sequence, and hybridization of the first probe to the third probe, in the absence of the reference nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the reference nucleic acid sequence, wherein the first probe comprises a polynucleotide comprising a 5' region which hybridizes to a first region of the reference nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of the reference nucleic acid sequence, if present, and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the reference nucleic acid sequence is present; and the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the reference nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the reference nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the reference nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; and comparing detectable signal generated by contacting said first, second, and third probes with the reference nucleic acid sequence with detectable signal generated by contacting said first, second, and third probes with target nucleic acid sequence; whereby reduced signal generation by the target nucleic acid sequence as compared to the reference nucleic acid sequence indicates the presence of an altered sequence in the target nucleic acid sequence relative to the reference nucleic acid sequence.

As is evident to one skilled in the art, in these methods binding of first and second probes to the target nucleic acid sequence competes with binding of the third probe to the second probe, thus maintaining the spatial separation of the labels of the first probe and the third probe, whereby a detectable and/or quantifiable signal is generated .

IDT Exhibit 1020.9637

In another aspect of the present invention, methods for genotype determination are provided. In these methods, the first and/or third oligonucleotide probes that are allele-specific are used.

Another aspect of the invention provides methods of detection and/or quantification of multiple target nucleic acid sequences in a sample. In these methods, two or more combinations (pairs) of interacting labels, each pair specific for a defined target nucleic acid sequence, are used.

Accordingly, the invention provides a method of determining whether one or more of a plurality of target nucleic acid sequences is present in a sample, comprising: contacting said sample with a plurality of probe sets, each set comprising a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to a target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of said target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to said target nucleic acid sequence, wherein the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of said target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of said target nucleic acid sequence, if present, and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if said target nucleic acid sequence is present; the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of said target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of said target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of said target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of said target nucleic acid sequence; and wherein each probe set comprises an interacting label pair which generates a detectable signal which is different from the signals of the interacting label pairs of every other probe set, and generation of two or more signals indicates presence of a plurality of target nucleic acid sequences.

IDT Exhibit 1020.9638

In another aspect, the invention provides a method of determining whether one or more of a plurality of target nucleic acid sequences is present in a sample, comprising: contacting said sample with a plurality of probe sets, each set comprising a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to a target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of said target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to said target nucleic acid sequence, wherein the first probe comprises a polynucleotide comprising a 5' region which hybridizes to a first region of said target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of said target nucleic acid sequence, if present, and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if said target nucleic acid sequence is present; the third probe comprises a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of said target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of said target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of said target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of said target nucleic acid sequence; and wherein each probe set comprises an interacting label pair which generates a detectable signal which is different from the signals of the interacting label pairs of every other probe set, and generation of two or more signals indicates presence of a plurality of target nucleic acid sequences.

The invention also provides compositions, kits, complexes, reaction mixtures and systems comprising various components (and various combinations of the components) used in the methods described herein.

In one aspect, the invention provides a composition comprising a first probe, a second probe and a third probe, wherein the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of a target nucleic acid sequence,

IDT Exhibit 1020.9639

if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present; the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

In another aspect, the invention provides a composition comprising a first probe, a second probe and a third probe, wherein the first probe comprises a polynucleotide comprising a 5' region which hybridizes to a first region of a target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present; the third probe comprises a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

The compositions may comprise an interacting moiety pair which comprises a fluorophore and a quencher. The compositions may further comprise a nucleotide

IDT Exhibit 1020.9640

polymerase the target nucleic acid sequence, and/or a reference nucleic acid sequence to which the target nucleic acid sequence is to be compared.

In another aspect, the invention provides kits for conducting the methods described herein. These kits, in suitable packaging and generally (but not necessarily) containing suitable instructions, contain one or more components used in the detection and/or quantification methods.

In one aspect, the invention provides a kit for determining whether a target nucleic acid is present in a sample or quantifying a target nucleic acid sequence, comprising a first probe, a second probe and a third probe, wherein the first probe comprises a polynucleotide comprising a 3' region which hybridizes to a first region of a target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present; and the third probe comprises a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

In a further aspect, the invention provides a kit for determining whether a target nucleic acid is present in a sample or quantifying a target nucleic acid sequence, comprising a first probe, a second probe and a third probe, wherein the first probe comprises a polynucleotide comprising a 5' region which hybridizes to a first region of a target nucleic acid sequence, if present, and a first member of an interacting label pair; the second probe comprises a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present; the third probe comprises a polynucleotide which hybridizes to a

IDT Exhibit 1020.9641

sequence in the 3' region of the first probe, and a second member of an interacting label pair; wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact; and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair; whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

Kits of the invention may further comprise a reference nucleic acid sequence to which the target nucleic acid sequence may be compared, and/or instructions for use of the kit to determine the presence of the target nucleic acid sequence in a sample or quantify the target nucleic acid sequence.

In another aspect, the invention provides compositions comprising any of the complexes (which are generally considered as intermediates with respect to the final products) described herein (see also the figures for schematic depictions of these various complexes).

In another aspect, the invention provides reaction mixtures (or compositions comprising reaction mixtures) which contain various combinations of components described herein.

In another aspect, the invention provides systems for effecting the detection and/or quantification methods described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic representation of probe A (first oligonucleotide; Fig. 1a), probe B (second oligonucleotide; Fig. 1b) and probe C (third oligonucleotide; Fig. 1c). As an example of an interacting label pair, "F" and "Q" are labels, such as, for example, a fluorophore and a quencher.

Figure 2 is a schematic representation of the hybrid complex of probe A and probe C ("Complex I").

Figure 3 is a schematic representation of the hybrid complex of probe A, probe C and target nucleic acid ("Complex II").

IDT Exhibit 1020.9642

Figure 4 is a schematic representation of the hybrid complex of probe A, probe B, probe C and target nucleic acid ("Complex III").

Figure 5 is a schematic representation of the hybrid complex of probe A, probe B, and target nucleic acid, with a displaced probe C.

Figure 6 is a schematic representation of the mirror image design of probe A (Fig. 6a), probe B (Fig. 6b), and probe C (Fig. 6c), depicted in the preceding figures. As an example of an interacting label pair, "F" and "Q" are labels, such as, for example, a fluorophore and a quencher.

Figures 7-10 are schematic representations of the hybrid complexes formed by the mirror image design probes depicted in Figure 6.

Figure 11 is a schematic representation of the hybrid complexes formed in the methods of detecting mutation or polymorphism.

Figure 12 is a schematic representation of the hybrid complexes formed by a probe B, a single probe that comprises both labels and the functions of both probes A and C, and target nucleic acid.

Figure 13 is a schematic representation of probes of the invention in which displacement of probe C is carried out by extension of probe B along the 5'-portion of probe A by a nucleotide acid polymerase, for example, a DNA polymerase. As an example of an interacting label pair, "F" and "Q" are labels, such as, for example, a fluorophore and a quencher.

MODES FOR CARRYING OUT THE INVENTION

The present invention discloses detector oligonucleotide probes for detection and/or quantification of target nucleic acid sequences, and detection and/or quantification methods using the probes. In the absence of target nucleic acid sequence, a first oligonucleotide and a third oligonucleotide are bound to each other in a conformation which brings members of an interacting label pair into a sufficiently close spatial proximity that the members of the pair interact. Cooperative binding of the first oligonucleotide and a second oligonucleotide to a target polynucleotide, optionally followed by polymerase-catalyzed extension of the second oligonucleotide, effects (a) displacement of the third oligonucleotide from the first oligonucleotide probe resulting in dissociation (separation) of the members of the

IDT Exhibit 1020.9643

interacting label pair , and/or (b) competition with binding of the third probe to the second probe, thus maintaining the spatial separation of the labels of the first probe and the third probe, whereby a detectable and/or quantifiable signal is generated. The spatial separation of the members of the interacting label pair (labels) is detectable, and is indicative of the presence and/or amount of the target nucleic acid sequence. Thus, the methods are useful for detection and/or quantification of a specific nucleic acid sequence as well as the detection of sequence alterations in a target nucleic acid sequence.

A variety of interacting label pairs (labels) can be used in methods and probes of the invention. Detectable signals indicative of separation of the members of the interacting label pair can be achieved in a number of ways, as is known in the art. As discussed herein, generating a signal includes any of increase, alteration, dampening or reduction, and/or elimination of a signal. For example, in some embodiments, generation of a detectable signal is dependent on changes in the spectroscopic properties of the members of the interacting label pair, wherein the spectroscopic properties are affected by the spatial proximity of the members of the interacting label pair. In other embodiments, members (labels) possess donor and acceptor properties, respectively, wherein said donor and acceptor properties are affected by the spatial proximity of the labels. Donor and acceptors generally refer to dye labels or other energy transfer pairs. In other embodiments, members may be any interacting pair that generates a signal when separated, for example, an enzyme-inhibitor pair or allosteric enzyme-suppressor pair. Signal detection may be direct or indirect.

The target nucleic acid sequence may be directly detected by the methods of the invention or may be first amplified by any of a variety of nucleic acid amplification methods known in the art.

The present invention provides numerous advantages over previously described detection and/or quantification methods. For example, the advantages of the present invention include the ease of preparation of the detector probes. The design of the interacting sequences of the detector probes is independent of the target polynucleotide sequence and could thus be universal. Insofar as the detector probes are not unimolecular, the design complexities of stem loop probes such as those known in the art are eliminated.

The methods of the invention may be used to detect and quantify target nucleic acid sequences, or determine whether a target nucleic acid is present in a sample. The methods may also be used to detect changes in target nucleic acid sequences, for example, mutations

IDT Exhibit 1020.9644

(including single nucleotide mutations), or, for example, polymorphisms. The methods may also be used to detect and quantify multiple target nucleic acid sequences.

In addition, the methods of the invention may be used in conjunction with nucleic acid amplification techniques to detect, quantify, and/or verify fidelity of amplification. The methods may be used in real time; i.e., simultaneously with amplification, in the same reaction mixture. The method may further be used for signal generation for the quantification of an analyte (i.e., a moiety other than the target nucleic acid sequence to be detected) in a sample. A known nucleic acid sequence may be attached to one member of a specific binding pair. The specific binding pair (SBP) comprises a first member of the SBP that is an analyte to be tested and the second member of the SBP is a molecular species, such as an antibody, which can specifically bind to the first member of the SBP. The second member of the SBP may be labeled by attachment of the known nucleic acid sequence. Following binding of the members of the SBP, the complex may be detected and quantified by probes and method of the invention.

General Techniques

The practice of the present 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).

Probe, oligonucleotides and polynucleotides employed in the present invention can be generated using standard techniques known in the art.

Definitions

The "target nucleic acid sequence," or sequences, or "target nucleic acid," or "target polynucleotide" (all of which are used interchangeably herein) may be any nucleic acid sequence(s) the presence and/or amount of which is desired to be known. In some

IDT Exhibit 1020.9645

embodiments, the sequence of the target nucleic acid is known. In some embodiments, e.g., mutation detection, the target nucleic acid sequence may be a sequence which is suspected of having alterations (i.e., differences) from a reference nucleic acid sequence. In these embodiments, the sequence of the target nucleic acid may or may not be known, and the "reference nucleic acid sequence" is a nucleic acid whose sequence is known and to which the target nucleic acid sequence or sequences may be compared, e.g., a wild-type sequence. In this case, the alteration in the target nucleic acid may be in a single nucleotide base or more than a single nucleotide base. Such an altered sequence can be a known polymorphic sequence, including, for example, single nucleotide polymorphism.

"Polynucleotide," or "nucleic acid," as used interchangeably herein, refer to polymers of nucleotides of any length, and include DNA and RNA. The nucleotides can be deoxyribonucleotides, ribonucleotides, modified nucleotides or bases, and/or their analogs, or any substrate that can be incorporated into a polymer by DNA or RNA polymerase. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and their analogs. If present, modification to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications include, for example, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, 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). Further, any of the hydroxyl groups ordinarily present in the sugars may be replaced, for example, by phosphonate groups, phosphate groups, protected by standard protecting groups, or activated to prepare additional linkages to additional nucleotides, or may be conjugated to solid supports. The 5' and 3' terminal OH can be phosphorylated or substituted with amines or organic capping groups moieties of from 1 to 20 carbon atoms. Other hydroxyls may also be derivatized to standard protecting groups. Polynucleotides

IDT Exhibit 1020.9646

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, sedoheptuloses, acyclic analogs and abasic nucleoside analogs such as methyl riboside. One or more phosphodiester linkages may be replaced by alternative linking groups. These alternative linking groups include, but are not limited to, embodiments wherein phosphate is replaced by P(O)S("thioate"), P(S)S ("dithioate"), "(O)NR_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. In addition, "polynucleotides" include peptide nucleic acid (PNA), lacking pentose sugar phosphate groups, in which the monomeric unit is 2-aminoethyl glycine linked by a methylenecarbonyl linkage to a base found in DNA. Not all linkages in a polynucleotide need be identical. The preceding description applies to all polynucleotides referred to herein, including RNA, DNA, and PNA, as appropriate.

"Hybridizable," as used herein, refers to the capability of two polynucleotide sequences to hybridize through complementary base pairing, under conditions used in an assay described herein; i.e., at the temperature, pH, ionic concentrations, and the like, used in carrying out the methods of the invention. "Mutually hybridizable" refers to two polynucleotide sequences that are capable of hybridizing under the conditions used in the assay.

An "interacting label pair," as used herein, refers to two members whose ability to interact with each other is dependent on their proximity to each other, wherein said interaction results in generation of a detectable signal indicative of the proximity of the moieties to each other. For example, when moieties (members) of an interacting label pair are placed in sufficient proximity to each other, they are capable of generating a detectable signal that is distinguishable from a signal associated with said moieties when not placed in said sufficient proximity. The signal generated (or generation of a signal) may be an increase in some measurable characteristic or a decrease in some measurable characteristic. An interacting label pair generally comprises two interacting moieties. However, each moiety or member of an interacting label pair may comprise more than one moiety or member. Interacting label pairs are known in the art, and described herein, for example, fluorescent dye-quencher, and receptor-ligand.

IDT Exhibit 1020.9647

By "dissociation" is meant lessening or reduction of the proximity of the members of the interacting label pair, including but not limited to separation of the two members, or failure of the two members to associate, which in turn arises from probes binding to the target nucleic acid sequence, such that detectable signal is generated (as generally compared to signal arising from interaction of the label pairs)--i.e., such that activity caused by proximate association (interaction) is affected. The signal generated (or generation of a signal) may be an increase in some measurable characteristic or a decrease in some measurable characteristic. The terms "dissociation" and "separation" and "dissociate" and "separate" are used interchangeably herein.

By "signal" is meant a measurable characteristic. The signal may increase or decrease upon dissociation of the members of the interacting label pair. For example, if the interacting label pair comprises a fluorophore and a quencher, dissociation of the members of the pair generates a detectable signal due to an increase in light energy emitted by the fluorophore in response to illumination. Or, for example, if the interacting label pair comprises subunits of an enzyme, dissociation of the members of the pair generates a detectable signal which is a decrease in the rate of the reaction catalyzed by the enzyme.

A "complex" is an assembly of components. A complex may or may not be stable and may be directly or indirectly detected.

A "system," as used herein, includes a device, apparatus or machinery (e.g., automated) for carrying out the methods of the invention.

A "portion" or "region," used interchangeably herein, of a polynucleotide or oligonucleotide is a contiguous sequence of 2 or more bases. In other embodiments, a region or portion is at least about any of 3, 5, 10, 15, 20, 25 contiguous nucleotides.

A region, portion, or sequence which is "adjacent" to another sequence directly abuts that region, portion, or sequence.

"A", "an" and "the", and the like, unless otherwise indicated include plural forms.

"Comprising" means including.

Conditions that "allow" or "permit" an event to occur or conditions that are "suitable" for an event to occur, such as hybridization, primer extension, oligonucleotide ligation 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 nucleotide sequence, temperature, and buffer conditions. These

IDT Exhibit 1020.9648

conditions also depend on what event is desired, such as hybridization, cleavage, primer extension or ligation.

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.

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.

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. This term encompasses a region which is 3' (downstream) from another region or position in the same polynucleotide or oligonucleotide (such as a probe). The "3' most nucleotide" (singular form) refers to the 3' last nucleotide of a polynucleotide or oligonucleotide. The 3' most nucleotides (plural form) includes the 3' most nucleotide and can be preferably from about 1 to about 20, more preferably from about 3 to about 18, even more preferably from about 5 to about 15 nucleotides.

The term "5'-DNA portion," "5'-DNA region," "5'-RNA portion," and "5'-RNA region," refer to the portion or region of a polynucleotide or oligonucleotide located towards the 5' end of the polynucleotide or oligonucleotide, and may or may not include the 5' most nucleotide(s) or moieties attached to the 5' most nucleotide of the same polynucleotide or oligonucleotide. This term encompasses a region which is 5' (upstream) from another region or position in the same polynucleotide or oligonucleotide (such as a probe). The "5' most nucleotide" (singular form) refers to the 5' first nucleotide of a polynucleotide or oligonucleotide. The 5' most nucleotides (plural form) includes the 5' most nucleotide and can be preferably from about 1 to about 20, more preferably from about 3 to about 18, even more preferably from about 5 to about 15 nucleotides.

"Affinity," or "binding affinity", as used herein, means a measure of the strength of bonding between two or more moieties; non-limiting examples of such bonding are hydrogen bonding, electrostatic interactions, Van der Waals interactions, and hydrophobic interactions. In particular, "affinity" or "binding affinity," when used in reference to

IDT Exhibit 1020.9649

hybridized nucleic acids, refers to hydrogen bonding between at least partly complementary nucleic acids under defined nucleic acid hybridization conditions. A convenient measure of binding affinity is the melting temperature, T_m, which is the temperature at which 50% of said two strands are in the double-stranded or hybridized form, under given hybridization conditions.

Probes and methods of the invention

Oligonucleotide probes of the invention

The present invention provides detector oligonucleotide probes, which are useful for the detection and/or quantification of a specific polynucleotide sequence.

Probes A, B, and C have various embodiments and configurations, thus the term "Probe A," "Probe B," "Probe C" generally refers to any of these embodiments, including mirror images. Further, discussion of, for example, probe design, methods, using the probes, etc., refer to exemplary probes and are generally applicable to various probe and method embodiments.

The first oligonucleotide probe, referred to herein as probe A (Figure 1a), generally comprises a 3'-region (3) that is hybridizable to a sequence of the target nucleic acid; a 5'-region that is not hybridizable to the target nucleic acid under reaction conditions in which sequence 3 and sequence 4 (see below for probe B) can hybridize to the target nucleic acid, and a member of an interacting label pair (F in Fig. 1a) that is generally, but not necessarily, attached to the 5'-end (Figure 1a). Sequence 3 is preferably at least about 60%, more preferably at least about 80%, and most preferably at least about 90% complementary to the sequence of the target nucleic acid to which it is hybridizable.

The second oligonucleotide probe, referred to herein as probe B (Figure 1b), generally comprises a 5'-region (4), which is hybridizable to a sequence of the target nucleic acid and a 3'-region (2), which is not hybridizable to the target nucleic acid under reaction conditions in which sequence 3 and sequence 4 can hybridize to the target nucleic acid. Sequence 4 is preferably at least about 60%, more preferably at least about 80%, and most preferably at least about 90% complementary to the sequence of the target nucleic acid to which it is hybridizable. Sequence 2 is hybridizable to a sequence of the region of probe A which is not hybridizable to the target nucleic acid under reaction conditions in which sequence 3 and sequence 4 can hybridize to the target nucleic acid. Sequence 2 is

IDT Exhibit 1020.9650

preferably at least about 70%, more preferably at least about 80%, and most preferably at least about 90%, complementary to the sequence of probe A to which it is hybridizable. The length and composition of sequence 2 are designed so that the affinity of the two sequences is too low to allow binding to a significant extent of the two sequences unless probe A and probe B are bound to the same nucleic acid strand of target nucleic acid. The two (mutually hybridizable) sequences of probe A and probe B do not bind to a significant extent when the length and composition of the two sequences are selected according to thermodynamic considerations so as to allow the two probes, when not bound to a target nucleic acid and at a selected incubation temperature, to bind to each other in preferably less than about 5 %, more preferably less than about 0.01%, and most preferably less than about 0.0001% of events of contact between the two probes. The appropriate thermodynamic considerations are evident to one skilled in the art, and can be determined empirically and using commonly available tools for calculation of hybridization efficiencies of various nucleic acid sequences.

The target polynucleotide sequences which are complementary to sequence 3 and sequence 4 of probe A and probe B, respectively, are in close proximity (i.e., preferably less than about 10 nucleotides apart, more preferably less than about 7 nucleotides apart, even more preferably less than about 3 nucleotides apart, and generally most preferably contiguous) to each other on the same strand of said target nucleic acid sequence.

In some embodiments, in which a nucleotide polymerase is used, probe B may or may not contain a region which overlaps and may displace or block a sequence on probe C. Accordingly, when using the method of the invention in the presence of a nucleotide polymerase, such as DNA polymerase, and components required for nucleic acid polymerization (nucleoside triphosphates and proper buffer and temperature conditions), there is generally no requirement for homology between the 3' end of probe B and the 5' end of probe C, since displacement is carried out by the extension of the hybridized probe B along the 5' portion of probe A by the polymerase. I.e., there need be little or no overlap between section 1' and section 2' of probe A and hence little or no overlap of binding of probe B and probe C to probe A. In this embodiment, because displacement is accomplished mostly or entirely by extension of probe B to displace or block probe C, the portion of probe B that hybridizes to probe A may be of lower affinity with probe A than in embodiments where probe B displaces probe C without extension. In some embodiments, probe B need not overlap with probe A. In these embodiments, the probes

IDT Exhibit 1020.9651

should be designed to allow template switch, e.g., by incorporation of a template switch oligonucleotide (TSO). Such a template switch oligonucleotide has been described in, e.g., U.S. Patent No. 6,251,639, and references therein (Kurn), and U.S. Patent Nos. 5,679,512; 5,683,879; and 6,030,774 (Laney et al.).

The third oligonucleotide probe, referred to herein as probe C (Figure 1c), comprises sequence 1, which is preferably not hybridizable to the target nucleic acid under reaction conditions in which sequence 3 and sequence 4 can hybridize to the target nucleic acid, and is hybridizable to a sequence in the 5'-region of probe A, and a member of an interacting label pair (Q in Fig. 1c) that is generally but not necessarily attached to its 3'-end. Sequence 1 is preferably at least about 60%, more preferably at least about 80%, and most preferably at least about 90% complementary to the sequence in the 5' region of probe A to which it is hybridizable. The affinity of probe A to probe C is high enough to permit binding of the two probes by hybridization of the mutually hybridizable (generally, complementary) sequences, when the two probes are in the same solution and probe A is either free or bound to said target nucleic acid sequence (complex I and complex II, respectively, see Figs. 2 and 3). As illustrated in Figures 2 and 3, moieties (labels) F and Q are in close spatial proximity when the two probes are hybridized.

Interacting label pairs. Probes A and C each comprises a member of an interacting label pair. The members interact when in close proximity, such that separation, or dissociation, of the members on the two probes results in generation of a signal. By "signal" is meant a measurable characteristic. The signal may increase or decrease upon dissociation of the members of the interacting label pair. For example, if the interacting label pair comprises a fluorophore and a quencher, dissociation of the members of the pair generates a detectable signal due to an increase in light energy emitted by the fluorophore in response to illumination. Or, for example, if the interacting label pair comprises subunits of an enzyme, dissociation of the members of the pair generates a detectable signal which is a decrease in the rate of the reaction catalyzed by the enzyme. Each member of the interacting pair may comprise one or more than one molecule or structure. The change in signal may be all-or-none (for example, if the moieties are an enzyme-inhibitor pair, where the enzyme is either active or inactive) or vary over a range (for example, if the moieties are a fluorophore-quencher pair). The change is characteristic for the moieties (labels) employed. In some embodiments, two or more kinds of interacting label pairs may be used

IDT Exhibit 1020.9652

in a single sample in order to differentiate, e.g., different target nucleic acid sequences. The detectable signal may be, e.g., a characteristic light signal that results from stimulating at least one member of a fluorescence resonance energy transfer (FRET) pair. Another example of a detectable signal is a color change that results from the action of an enzyme/suppressor pair or an enzyme/cofactor pair on a substrate to form a detectable product. In some embodiments, the signal is a reduction or absence in detectable signal.

Various combinations of moieties (labels) which are capable of producing a detectable signal which differs depending on their degree of proximity, can be used. Any combination or number of moieties (labels) which interact so as to produce a measurable change upon change in the proximity of the moieties (labels) is sufficient; hence, more than one pair of moieties (labels) may be used. Nor is it required that there be a one-to-one correspondence between members of an interacting label pair, especially where one member can affect, or be affected by, more than one molecule of the other member.

Interacting label pairs useful in the present invention are known in the art, see, e.g., U.S. Patents Nos. 5,688,648 (Mathies et. al) ; 5,340,716; 3,999,345; 4,174,384; and 4,261,968 (Ullman et al.); 4,996,143 and 5,565,322 (Heller et al.); 5,709,994 (Pease et al.); and 5,925,517 (Tyagi et al.). Examples of suitable moieties (labels) in which one member quenches another include a fluorescent label, a radioluminescent label, a chemiluminescent label, a bioluminescent label, an electrochemiluminescent label, and an enzyme-inhibitor combination. In some embodiments, the interacting moieties (labels) produce little or no signal when in close proximity, and a greater signal when separated. In other embodiments, the interacting moieties (labels) may generate a signal when in close proximity and generate less or no signal when separated. Examples of the latter such moieties (labels) are an enzyme and its cofactor and fragments or subunits of enzymes that must be close to each other for the enzyme to be active.

Various combinations of dye moieties (labels), which are capable of energy transfer when in close spatial proximity, can also be used. For example, interacting moieties (labels) (illustrated in the Figures as F and Q) may be a donor-acceptor dye pair, capable of energy transfer when in close spatial proximity. Label F may be a fluorescent dye and label Q a quencher which is able to absorb the fluorescence signal of label F by an energy transfer mechanism. Alternatively, the moieties (labels) may be ligands for reporter molecules which can interact with each other when brought in close spatial proximity, the interaction of which prevents or enables activity of one of the reporter molecules.

IDT Exhibit 1020.9653

Examples for suitable combinations of reporter groups useful for the methods of the invention are enzyme-inhibitor combination, reporter molecules which when reacting with one another form an active enzyme molecule, and the like. The dissociation of the two interacting reporter groups is detectable and indicative of the presence of one or more target nucleic acid sequence(s) in a sample, the quantity of nucleic acid target(s) in a sample or the degree of identity of the sequence of nucleic acid target(s) to that of a reference nucleic acid sequence(s).

Either probe A, or probe C, or both, may optionally incorporate more than one moiety to make up its member of the interacting label pair. The moieties may be located anywhere on the probes as long as they are capable of interacting when probe A and probe C bind together. The moieties may be attached to one end of the probe, or may be attached to the interior of the probe. In some embodiments, a moiety is located at least 1, at least 3, at least 5, or at least 7 nucleotides away from the 3' end probe A and/or the 5' end of probe C (or vice versa for the mirror image probes).

Members of the interacting label pairs may be attached to oligonucleotide probes either during or post-synthesis of the probes. The attachment of a member of an interacting label pair to the rest of the probe is preferably covalent, and means of attachment will vary depending on the probe and the member of the interacting label pair, such means being readily apparent to one of skill in the art.

Oligonucleotide probes. Labeled or unlabeled oligonucleotide probes are available commercially, and are usually prepared according to any one of a variety of methods known to those skilled in the art. An oligonucleotide can be prepared by any suitable method, including, for example, cloning and isolation of appropriate sequences using restriction enzymes and direct chemical synthesis by a method such as the phosphotriester method of Narang et al., Meth. Enzymol. (1979), 68:90-99; the phosphodiester method of Brown et al., Meth. Enzymol. 1979, 68:109-151; the diethylphosphoramidite method of Beaucage et al., Tetrahedron Lett. 1981, 22:1859-1862; and the solid support method of U.S. Pat. No.4,458,066. Methods for synthesizing labeled oligonucleotides are described in, for example, Agrawal and Zamecnik, Nucl. Acids. Res. (1990), 18(18):5419-5423; MacMillan and Vetdine, J. Org. Chem. (1990), 55:5931-5933; Pieles et al., Nucl. Acids. Res. (1989), 17(22):8967-8978; Roger et al., Nucl. Acids. Res. (1989), 17(19):7643-7651; Fisher and Watson, U.S. Pat. No. 5491063; and Tesler et al., J. Am. Chem. Soc. (1989),

IDT Exhibit 1020.9654

111:6966-6976. A review of synthesis methods is provided in, for example, Goodchild, Bioconjugate Chemistry (1990), 1(3):165-187.

In some embodiments, the displacement or blocking of probe C from its association with probe A in complex III by probe B (see Figures 4 and 5) is a consequence of the higher association affinity of sequence 2 of probe B (see Figure 1b) to its hybridizable sequence, 2', on probe A, when probe B is hybridized to target nucleic acid. The association of sequence 2 with its hybridizable sequence in probe A in complex III is thermodynamically favored insofar as this interaction is uni-molecular when probe B is hybridized to the target nucleic acid. Binding of free probe C to complex IV follows bimolecular reaction kinetics which is less favorable than the unimolecular interaction of probe A and probe B when said probes are bound to the same strand of said target nucleic acid sequence. In the absence of target nucleic acid sequence, probe A generally binds probe C and does not bind probe B. Thus in the absence of target nucleic acid sequence, probe A is generally associated with probe C to form complex I, where interacting label pair member F and interacting label pair member Q are in close spatial proximity (see Figure 2). Displacement of the portion of sequence 1 of probe C with respect to which sequence 2 has a significant degree of sequence identity from binding to probe A by sequence 2, is sufficient to cause the dissociation of probe C from probe A, leading to the dissociation of interacting label pair, or, alternatively, binding of probe A and probe B to the target nucleic acid sequence and subsequent interaction of sequences 2 and 2' blocks binding of probe C to probe A, thus maintaining the spatial separation of the labels probe A and probe C, whereby a detectable and/or quantifiable signal is generated. Moreover, the close spatial proximity of interacting label pair member F and interacting label pair member Q is maintained following association of complex I with target to form complex II, unless probe B is bound to the same target molecule to form complex III. Complex II is not detectable.

In some embodiments a portion of sequence 2 has a significant degree of sequence identity with respect to a portion of sequence 1 (see Figure 1c) of probe C. Said significant degree of sequence identity is a sequence identity between the two portions that is at least about 60%, more preferably at least about 75%, more preferably at least about 90%, and most preferably at least about 95%. The percentage of the sequence 1 of probe C which has a significant degree of sequence identity with sequence 2 of probe B required for probe B to be able to displace probe A will depend on thermodynamic considerations for the

IDT Exhibit 1020.9655

formation of Complex II and Complex IV. Probes are preferably constructed so that, under the conditions of the reaction, probe C, the 3' region of probe B, and the 5' region of probe A are not substantially hybridizable to the target nucleic acid sequence.

One factor in designing and constructing probes is the free energy parameters of hybridization of given sequences under a given set of hybridization conditions. The free energy parameters for the formation of a given hybrid may be calculated by methods known in the art (see, e.g., Tinoco et al., 1973, Improved Estimation of Secondary Structure in Ribonucleic Acids, Nature, 246, 40-41. and Freier et al., 1986, Improved free-energy parameters for predictions of RNA duplex stability, Proc. Natl. Acad. Sci. U.S.A., 83, 9373-9377; computer programs, e.g., Oligo Primer Analysis Software from Molecular Biology Insight, and references therein), and it is possible to predict, for a given target nucleic acid sequence, probe sequences for which the required free energy changes for formation of various complexes will be met. The type of nucleotides comprising probes A, B, and/or C may also contribute to adjusting or modifying the relative affinities of probes B and C for probe A. For example, the various probes may comprise RNA or peptide nucleic acid (PNA), which may have different affinities for single-stranded DNA than does DNA. Hence, probes may comprise any combination of nucleotide types. Thus, for example, in one aspect, probes A, B, and/or C may comprise RNA, in another aspect, probes A, B, and/or C may comprise DNA and RNA, in another aspect, probes A, B, and/or C may comprise PNA, in another aspect, probes A, B, and/or C may comprise DNA and PNA, in another aspect, probes A, B, and/or C may comprise RNA, and PNA, and in yet another aspect, probes A, B, and/or C may comprise RNA, PNA, and DNA.

One of skill in the art will understand that other factors affect nucleic acid hybridization affinities. For example, any and all of the guanosine-cytosine content of the probe-target and probe-probe duplexes, minor groove binders, O-methylation or other modification of nucleotides, temperature, and salt are potentially important factors in constructing probes with the requisite differences in binding energies. In addition, in some embodiments a nucleotide polymerase, for example, a DNA polymerase, is employed to extend probe B (see below), and in these embodiments the polymerase acts as a "clamp" to further enhance binding of probe B to probe A. Yet another factor affecting affinity is the proximity of probe A to probe B on the target nucleic acid, with the degree of separation of the two probes determining to some extent the degree of interaction; for example, in some cases separation by one nucleotide may be preferable to having the two probes be

IDT Exhibit 1020.9656

contiguous. The factors that determine the optimal separation, and other relevant factors, are readily apparent to those of skill in the art, and may be determined empirically. In some embodiments the probes (if hybridized to target nucleic acid sequence) are any of at least 1, at least 2, at least 3, at least 5, at least 7, at least 10, at least 15, at least 20 nucleotides apart. In other embodiments, the probes (if hybridized to target nucleic acid sequence) are any of contiguous, fewer than 1, 2, 3, 5, 7, 10, 15, 20 nucleotides apart.

The hybridization conditions chosen depend on a variety of factors known in the art, for example the length and type (e.g., RNA, DNA, PNA) of probe and target nucleic acids. General parameters for specific (i.e., stringent) hybridization conditions for nucleic acids are described in Sambrook (1989), supra , and in Ausubel (1987), supra. Useful hybridization conditions are also provided in, e.g., Tijessen, 1993, Hybridization With Nucleic Acid Probes, Elsevier Science Publishers B.V. and Kricka, 1992, Nonisotopic DNA Probe Techniques, Academic Press San Diego, Calif. The ability of two nucleotide sequences to hybridize with each other is based on the degree of complementarity of the two nucleotide sequences, which in turn is based on the fraction of matched complementary nucleotide pairs. The more nucleotides in a given sequence that are complementary to another sequence, the more stringent the conditions can be for hybridization and the more specific will be the binding of the two sequences. Increased stringency is achieved by any one or more of the following: elevating the temperature, increasing the ratio of cosolvents, lowering the salt concentration, and the like.

In some embodiments of the invention, the displacement of probe C from probe A by hybridization of probe B to the same target strand (as in complex III) can be further accomplished by extension of probe B along the 5'-portion of probe A by a by a nucleotide polymerase, preferably a DNA polymerase. A "nucleotide polymerase" is a catalyst, usually an enzyme, for forming an extension of a polynucleotide along a DNA or RNA template where the extension is complementary thereto. The nucleotide polymerase is a template dependent polynucleotide polymerase and utilizes nucleoside triphosphates as building blocks for extending the 3'-end of a polynucleotide to provide a sequence complementary with the polynucleotide template. Usually, the catalysts are enzymes, such as DNA polymerases, for example, prokaryotic DNA polymerase (I, II, or III), T4 DNA polymerase, T7 DNA polymerase, Klenow fragment, reverse transcriptase, Vent DNA polymerase, Pfu DNA polymerase, Taq DNA polymerase, and the like, derived from any source such as cells, bacteria, such as E. coli, plants, animals, virus, thermophilic bacteria,

IDT Exhibit 1020.9657

and so forth. RNA polymerases include T7 RNA polymerase, AMV polymerase, Q-beta-replicase, and so forth. The use of a nucleotide polymerase is particularly useful in cases where the method of the invention is employed for the homogeneous, i.e., simultaneous amplification, detection, and quantification of a target nucleic acid sequence . Isothermal linear amplification, or any other amplification method employing a nucleotide polymerase such as DNA polymerase, is suitable for this aspect of the present invention. This aspect of the invention can be added to other embodiments. Once probe B and probe A are hybridized to the same target strand, the 3'-end of probe B can be extended by the polymerase used in the amplification, and displace and/or block probe C (see Figure 13). The use of polymerase with strand displacement activity is preferred. When using a method of the invention in the presence of a nucleotide polymerase, such as DNA polymerase, and components required for nucleic acid polymerization (nucleoside triphosphates and proper buffer and temperature conditions), there is generally no requirement for homology between the 3' end of probe B and the 5' end of probe C, since displacement or blocking is carried out by the extension of the hybridized probe B along the 5' portion of probe A by the polymerase. That is, there need be little or no overlap between section 1' and section 2' of probe A and hence little or no overlap of binding of probe B and probe C to probe A. In this embodiment, because displacement and/or blocking is accomplished mostly or entirely by extension of probe B to displace and/or block probe C, the portion of probe B that hybridizes to probe A may be of lower affinity with probe A than in embodiments where probe B displaces probe C without extension. In some embodiments, probe B need not overlap with probe A. In these embodiments, probes should be designed to allow template switch, e.g., a template switch oligonucleotide (TSO) may be incorporated. Such a template switch oligonucleotide has been described in, e.g., U.S. Patent No. 6,251,639, and references therein (Kurn), and in Patel et al., Proc. Nat'l. Acad. Sci. USA, 1996, 93:2969-2974.

In some embodiments, the mirror image design of the detection probes are provided, as depicted in figures 5a, 5b, 5c, 6, 7 and 8. The mirror image design of a detection probe is expected to only change the probe's polarity. The ability to utilize either one of these designs provides an advantage of choice of probes in cases where sequences 3 or 4 of either probe A or probe B, respectively, in combination with the sequences (in the respective probe) that are not hybridizable to the target nucleic acid under reaction conditions in which sequence 3 and sequence 4 can hybridize to the target nucleic acid,

IDT Exhibit 1020.9658

result in internally hybridizable structures leading to secondary structures which are not compatible with the disclosed probe design features. Evaluation of suitable probe sequence can be carried out using tools well known in the art.

In some embodiments, a single probe that comprises the functions of probe A and probe C is used. For example, a single probe which comprises two interacting moieties (labels) attached to two sequence portions, respectively, wherein said two sequence portions are hybridizable to each other but, preferably, not to the target nucleic acid under conditions in which the sequences of probe A and probe B that are hybridizable to the target nucleic acid can hybridize to the target nucleic acid, and are capable of forming a hairpin loop structure, as shown in Figure 12, can be used in the methods of the invention. This design of the detector probe comprises a combination of the two probes each containing a member of an interacting label pair, described above, in the form of a unitary molecule. The two labeled sequence portions of the combination probe may be connected by a nucleotide sequence which is not related to the target sequence, or by a non-nucleotide structure. The two interacting labels are attached to nucleotides of the two hybridizable sequences so that the members of the interacting label pair are placed in close proximity when a hairpin loop structure is formed due to the hybridization of the two mutually hybridizable, sequence portions containing the members of the interacting moiety pair. For example, a hairpin loop is formed when the 5' most sequence of probe A (Figure 12) is hybridizable to an internal sequence in probe A which is, preferably, not hybridizable to the target. The linker may be peptide, sugar polymer, polyacrylamide, PEG and the like, as will be apparent to those of skill in the art.

Any combination of moieties (labels) described herein may be used in a single probe combining the functions of both probes A and C. The binding of the probe comprising the interacting moieties (labels) (for example, probe A in Figure 12) and the displacer/blocking probe (for example, probe B in Figure 12) to the target sequence, to form the tri molecular complex, and the subsequent binding of the displacer/blocking probe (such as probe B as described herein) to its hybridizable sequence on the probe comprising the interacting moieties (labels), results in opening of the hairpin loop structure through the displacement of the 5' end most sequence from hybridization to its internal hybridizable sequence (see, for example, Figure 12). Alternatively, the displacer/blocking probe (such as probe B as described herein), binds to the target nucleotide sequence and interacts with the first probe while the mutually hybridizable portions of the first probe are not

IDT Exhibit 1020.9659

hybridized, thus blocking hybridization. The interacting label pair members are thus separated. The properties of the interacting label pair members (for example, a fluorophore and a quencher) are altered as a result of the separation, and a signal is produced. The detection of the change in spatial proximity of the interacting label pair members provides for detection of the hybridization of the two probes to the target nucleic acid sequence. The signal generated by the hybridization of the two probes indicates the presence of the target sequence and is proportional to the quantity of the target nucleic acid sequence in the test sample.

In another aspect of the invention, compositions are provided which are universal reporter groups and which can be attached to any of a variety of sequences which bind to specific target polynucleotides. These groups differ from the probes previously described in that they do not contain the polynucleotide sequences which are hybridizable to the target nucleic acid sequence (3 and 4 in Fig. 1a and 1b, respectively), or, in other embodiments, only contain parts of the sequence(s) that bind to target that will be used in the assay. This is based on the fact that, although the parts of probes A and B which interact with target polynucleotide (3 and 4 in Figure 1) will vary depending on the sequence of the target nucleic acid, the parts of the probes which interact with each other (1, 2, and 1' and 2' in Figure 1) and which include interacting moieties (labels)(F and Q in Figure 1) may remain constant or nearly constant even in probes for different target nucleic acids. Hence, in this aspect, the invention provides a first probe which comprises the sections of probe A (as described herein) which interact with probes B and C (sections 1' and 2', in Figure 1) or portions thereof, and one member of an interacting label pair (F in Figure 1); a second probe which comprises the section of probe B (as described herein) which hybridizes to probe A (section 2, in Figure 1) or a portion thereof, but such hybridization is weak unless additional hybridization of an added sequence complementary to the target nucleic acid sequence occurs; and a third probe which is substantially the same as probe C (as described herein), comprising section 1 or a portion thereof (section 1 in Fig. 1c), and a second member of an interacting label pair (Q in Figure 1). As with the more specific probes which contain the 3 and 4 sections (Figure 1), the universal probes are designed so that hybridization of the second probe with a section of the first probe is weak unless the second probe also contains a sequence which hybridizes to the target nucleic acid sequence. In the case of the universal probes, this latter sequence must be added to the probe before use, and will vary depending on the nucleic acid sequence to be targeted. These probes

IDT Exhibit 1020.9660

have more general characteristics than the previously described probes. They may be attached to any of a variety of sequences which are complementary to different target nucleic acid sequences, and hence provide universal reporter probes which can be combined with unique sequences which bind specific target nucleic acid (e.g., sequence 3 for probe A in Figure 1 and sequence 4 for probe B in Figure 1). As with the larger probes, displacement or blocking of probe C by probe B may be accomplished with or without extension.

Methods of Use

The methods generally involve detection and/or quantification of a target nucleic acid sequence (including determining whether a target nucleic acid sequence is present in a sample). The target nucleic acid sequence may be part alone or of an analyte, and in some embodiments may be amplified before detection and/or quantification. Amplification techniques are well-known in the art; see, e.g., U.S. Patent Nos. 4,683,202, (Mullis) and 6,251,639 (Kurn). The target nucleic acid, or one or more of the probes of the invention, may be free in solution, or, in other embodiments, attached to a solid support, e.g., as part of a microarray. Target nucleic acid sequence(s) may be immobilized on a substrate fabricated from a material such as paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, silicon, and optical fiber. Alternatively, the target nucleic acid sequence may be 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/or cylinders. In some cases, the target nucleic acid sequence may be attached as a tag to another analyte moiety, e.g., an antibody, the presence, location, or quantity of which is desired to be known. Other examples of possible target nucleic acid sequences will be readily apparent to those of skill in the art. Target nucleic acid sequences for use in the methods of this invention can come from any sources, for example, biological samples, and may be purified and prepared for the methods of the invention using any number of methods, as is well-known in the art. Methods of rendering a double-stranded polynucleotide single-stranded, such as using heat, are known in the art.

Generally, methods of the invention comprise contacting a target nucleic acid with probe A, probe B, and probe C, as described herein, wherein binding of probes A and B to the target nucleic acid, and to each other, leads to displacement of bound probe C from probe A, thus separating interacting moieties (labels) on probe A and on probe C, resulting

IDT Exhibit 1020.9661

in a detectable signal, or, alternatively, binding of probe A and probe B to the target nucleic acid sequence blocks binding of probe C, thus maintaining the spatial separation of the labels of probe A and probe C, whereby a detectable and/or quantifiable signal is generated. In some cases, e.g., mutation detection, a decrease in, or lack of, binding of probes A and B to the target nucleic acid sequence indicates the presence of the alteration of interest, e.g., a mutation. This decrease or lack may be detected by comparison with a reference nucleic acid sequence.

In one embodiment of the invention, detection of a target nucleic acid sequence comprises a) combining a sample suspected of containing said target nucleic acid sequence with probe A, probe B and probe C, for example, provided as a mixture; optionally b) treating the mixture to render the target nucleic acid single stranded (if not already single stranded); and c) incubating the mixture under conditions which are suitable for binding of complex I and probe B to said single stranded nucleic acid target , to form complex III (Fig. 4). In this embodiment, sequence 2 of probe B will lead to displacement or blocking of probe C from its association with probe A to form complex IV, thus increasing the distance between members of an interacting label pair (illustrated in the Figures as F and Q) so as to reduce or eliminate the interaction between the two members of the interacting label pair (Fig. 5).

The dissociation of the moieties (labels) may produce an all-or-none signal (e.g., dissociation of an enzyme-inhibitor pair) or may produce a change in signal characteristic over a range (e.g., fluorescence quenching). The dissociation of the moieties (labels) is detectable and the signal generated is related to the presence of said target nucleic acid. In one embodiment, Probe B is designed to bind to said target at a sequence which is proximal to (i.e., at most about 10 nucleotides apart, or at most about 7 nucleotides apart, or at most about 3 nucleotides apart, or at most about 2 nucleotides apart, or at most about 1 nucleotide apart, or contiguous) to the sequence which is bound by probe A.

In another embodiment of the invention, the 3'-end of the displacer/blocker probe (for example, probe B as described herein), when hybridized to the target sequence and the labeled probe, can be extended by a nucleotide polymerase along a sequence of probe A such that probe C is displaced or blocked from probe A. In this embodiment, probes B and C are both bound to probe A, and there is generally little or no overlap between the binding regions for probes B and C. The extension of the displacer/blocker probe results in displacement or blocking of probe C, which carries one member of an interacting label pair,

IDT Exhibit 1020.9662

thus reducing the proximity of the members of the interacting label pair. The change in proximity of the two members of the interacting label pair is detectable. When one member is a quencher and the other is a fluorescer, the binding of the presence of the target is can be detected by the fluorescence signal of the fluorescer. In this embodiment, there need be no overlap of probe B and probe C on probe A; extension of probe B displaces or blocks probe C. Preferably, a nucleic acid polymerase (for, e.g., DNA or RNA polymers) used for extending probe B possesses strand displacement activity. Methods of extension are known in the art.

The various probes may be added simultaneously to a sample suspected of containing said target nucleic acid sequence, or may be added sequentially, or added separately at various timepoints. Such timepoints are evident to one skilled in the art, and can be determined empirically. As described above, signal generation is dependent on the binding of complex II and probe B to the same target nucleic acid sequence. A mixture of probe A and C may be added first to the sample followed by the addition of probe B. It is preferred that the concentration of probe C is greater than that of probe A to ensure the association of substantially all of probe A with probe C in the absence of target polynucleotide. For example, the molar concentration of C may be at least five-fold that of A, at least ten-fold that of A, at least one-hundred-fold that of A, or at least one-thousand-fold that of A.

The methods of the invention also provide for detection of changes, e.g., mutations, in a defined sequence of the target nucleic acid relative to a reference nucleic acid sequence, where a reference nucleic acid sequence is a known sequence to which others are to be compared, e.g., the wild-type sequence if one is searching for mutations, or the mutated sequence itself. In one embodiment, detection of sequence change, such as mutation or polymorphism, according to the method of the invention, comprises providing a sample containing said target nucleic acid, probe A, probe C and probe B, optionally treating the combination to render the target nucleic acid single stranded (for example, by heating), and further incubating the combination under conditions suitable for formation of a complex III (see Figure 11). In one embodiment, Probe B is designed to bind to said target at a sequence which is proximal to (i.e., at most about 10 nucleotides apart, or at most about 7 nucleotides apart, or at most about 3 nucleotides apart, or at most about 2 nucleotides apart, or at most about 1 nucleotide apart, or contiguous) to the sequence which is bound by probe A, but only when said sequence of the target nucleic acid which binds to

IDT Exhibit 1020.9663

probe B has a degree of sequence identity that is at least about 70%, more preferably at least about 85%, and most preferably at least about 95% that of a reference nucleic acid sequence. Any alteration (relative to a reference nucleic acid sequence) in said sequence of said target nucleic acid results in reduced affinity of probe B to said target nucleic acid sequence and reduces binding of probe B to said target nucleic acid sequence and formation of complex III. The reduction in ability to form complex III reduces displacement or blocking of probe C from complex II and subsequent signal generation. Thus signal generation is related to the degree of sequence identity of said target nucleic acid with respect to a reference nucleic acid sequence. Detection of a variety of mutation types is possible. For example, a mutated sequence can be a point mutation (substitution, deletion, single nucleotide polymorphism) or a mutation involving more than one nucleotide. The properties, e.g., length and type of nucleic acid, of the sequence of probe A, probe B and/or probe C may be adjusted to allow such detection. For example, as will be apparent to one of skill in the art, the shorter the binding section of a probe, the more likely a change in a single nucleotide will alter binding. Other factors which alter binding affinities, including those described supra, will be apparent to one of skill in the art.

Accordingly, in some embodiments, the invention provides methods for genotype determination. The methods are applicable to genotyping of any test organism, including prokaryotic and eukaryotic. The methods of the invention are applicable to the determination of either homozygote or heterozygote genotypes of a diploid organism or cell type. Signal generation, or suppression of signal generation, for a homozygote genotype is generally expected to be up to twice (preferably twice) that of a heterozygote. Various embodiments of genotyping according to methods of the invention are provided.

In one embodiment, allele-specific probe A, and/or probe B, are used. The detection of signal thus indicates the presence of a specific allele. It is possible to use different pairs of interacting moieties (labels), which produce distinguishable signals, for each of the specific alleles, thus providing a method for simultaneous detection of more than one known allele. It is possible to use the methods of the invention for the detection of any sequence alterations in the target sequence which is complementary to the target-specific sequences of probe A and/or probe B. In some embodiments, the target-specific sequence of the probe(s) is designed so that any alteration of the target sequence in this region will adversely affect the hybridization of either one, or both, of the probes to the target, thus preventing or diminishing cooperative binding of the non-target specific

IDT Exhibit 1020.9664

sequences of the probe(s) and the subsequent signal generation. The degree of sequence identity between the altered target sequence and unaltered target sequence is preferably less than about 50%, more preferably less than about 10%, even more preferably less than about 5%, and most preferably less than about 1%. In other embodiments, the sequence of the probe(s) is designed to hybridize more preferentially to a suspected altered target sequence than to an unaltered target sequence, so that a detectable signal is generated in the presence of the suspected altered target sequence.

Another embodiment of the invention provides methods of detection and/or quantification of multiple target nucleic acid sequences in a sample. These methods employ two or more sets of each of probe A, probe B and probe C, with each combination of probe A, probe B and probe C being specific for each target nucleic acid sequence. In these methods, two or more kinds of interacting label pairs (i.e., with respect to combinations of interacting label pair members F and Q as illustrated in Figure 1), each specific for a defined target nucleic acid sequence, are provided. For example, various combinations of dye moieties, which are capable of energy transfer when in close spatial proximity, can be used. Alternatively, the members of the label pair may be ligands for reporter molecules which can interact with each other when brought in close spatial proximity, the interaction of which prevents or enables activity of one of the reporter molecules. Examples for suitable combinations of reporter groups useful for the method of the invention are any of those known in the art and/or described herein, including enzyme-inhibitor combinations, reporter molecules which when reacting with one another form an active enzyme molecule, and the like. The dissociation of the two members of an interacting label pair, and subsequently the dissociation of the two interacting reporter groups, is detectable and indicative of the presence of said target nucleic acid sequence in said sample, the quantity of said nucleic acid target in said sample or the identity of said sequence of said nucleic acid target to that of a reference nucleic acid sequence.

It will be readily apparent that all of the above embodiments are suitable not only for detection but for quantification, using methods apparent to one of skill in the art, e.g., the use of standard quantities of target DNA as references against which to measure an unknown sample. For both detection and quantification, a control lacking the target nucleic acid sequence may be run and compared to the sample run under the same conditions, to determine the presence and/or amount of target nucleic acid sequence in the sample.

IDT Exhibit 1020.9665

The invention may also be used in conjunction with nucleic acid amplification techniques, as exemplified in Example 2. Any method of nucleic acid sequence amplification may be used for the amplification of a target nucleic acid sequence. These methods include PCR (Mullis et al. U.S. Pat. No. 4,582,788), isothermal exponential amplification methods such as nucleic acid sequence-based amplification (U.S. Patent No. 5,654,142), transcription-mediated amplification (U.S. Patent No. 5,766,849, or strand-displacement amplification (U.S. Patent No. 5,648,211), or isothermal linear amplification (U.S. Patent No. 6,251,639), linked linear amplification (Wallace et al., U.S. Pat. No.6,027,923), ligation-based amplification (Wu et al., Genomics 4:560, 1989). For a discussion of methods of amplification, see U.S. Patent No. 6,251,639, and references therein. The detection and quantification of the amplification products may be carried out simultaneously with the amplification reactions, or in a separate step following the amplification reaction. When the detection and quantification of the amplification products are carried out by generation of signals by displacement or blocking of a probe that is labeled with a quencher (probe C) by polymerase catalyzed probe extension, the nucleotide polymerase (e.g., DNA polymerase) employed in the amplification step may be used in the signal generation step. The quantification of amplification products is directly related to the amount of target nucleic acid in the sample. The efficiency of amplification of the target nucleic acid sequence may be separately determined. Control samples with known amounts of the target nucleic acid sequence may be subjected to amplification and quantification by the method of the present invention to produce a standard curve that can be used for the determination of the amount of the test sample. Quantification of a test nucleic acid in a sample may also be carried out by simultaneous amplification and detection of a test nucleic acid sequence and a control nucleic acid sequence that is different than the test sequence and serves as an internal control for quantification.

Compositions and Kits of the Invention

The invention also provides compositions, kits, complexes, reaction mixtures and systems comprising various components (and various combinations of the components) used in the methods described herein.

The compositions may be any component(s) (including any of the probe embodiments and any combination of the various probe embodiments), reaction mixture and/or

IDT Exhibit 1020.9666

intermediate described herein, as well as any combination. In one embodiment, the invention provides a composition comprising probe A, a composition comprising probe B, and/or a composition comprising probe C. In other embodiments, various combinations of the probes may be provided in compositions of the invention. For example, the invention may provide a mixture of probes A, B, and C; or it may provide, for example, a mixture of probes A and C, and provide probe B separately. When probes A, B, and C are provided together, probes A and C may form, under suitable reaction conditions, Complex I (see Figure 2). In the presence of target nucleic acid, all of the complexes of the invention may be present (see Figures 2-5), in varying relative amounts, depending on the affinity of the probes for the target nucleic acid and for each other, with concomitant signal generation. When probes A and C are provided together, and probe B is provided separately, under suitable reaction conditions probes A and C may combine to form Complex I (see Figure 2) or, in the presence of target nucleic acid, Complex II (see Figure 3). On addition of probe B, said compositions may, in the presence of target nucleic acid, form Complexes III and IV as well (see Figures 4 and 5, respectively), as described herein, with concomitant signal generation.

In some embodiments, probes B and C may contain portions of their sequences which overlap, each of which is hybridizable to the same portion of probe A, whereas in other embodiments B and C may contain no portions of their sequences which are homologous (see, e.g., Figure 13). In some embodiments, nucleotide polymerase, for example DNA polymerase, may be provided as a separate component, or as part of a composition. In these embodiments, suitable single nucleoside triphosphates, for example dNTPs may also be provided in order to allow extension of probes of the invention.

In some compositions, a single probe that comprises the functions of probe A and probe C is provided. This probe may be provided by itself, or in combination with probe B. For example, one aspect of the invention may comprise a composition comprising a first probe and a second probe, wherein the first probe comprises a 3' region that is hybridizable to a target polynucleotide comprising a sequence of interest, a 5' region comprising two mutually hybridizable sequences and members of an interacting moiety pair; the second probe comprises a 5' region which is hybridizable to the target polynucleotide, and a 3' region which is not hybridizable to the target polynucleotide and is hybridizable to a sequence in the 5' region of the first probe; wherein hybridization of the second probe to the target polynucleotide is associated with a decrease in extent of hybridization between

IDT Exhibit 1020.9667

the mutually hybridizable sequences of the first probe; whereby generation of a detectable signal associated with said decrease in extent of hybridization between the mutually hybridizable sequences indicates presence of the sequence of interest in the sample. This composition may further comprise components for extension of the 3' end of the second probe along a sequence in the 5' region of the first

In the compositions of the invention, interacting label pairs may be provided attached to the suitable probes (i.e., probes A and C as described herein). In other compositions, probes and/or interacting label pairs may be provided separately. Members of interacting label pairs include a fluorescent moiety, a radioluminescent moiety, a chemiluminescent moiety, a bioluminescent moiety, an electrochemiluminescent moiety, and an enzyme-inhibitor combination. In some embodiments, the members of the interacting label pair produce little or no signal when in close proximity, and a greater signal when separated. In other embodiments, the members of the interacting label pair may generate a signal when in close proximity and generate less or no signal when separated. Examples of the latter such interacting label pairs are an enzyme and its cofactor and fragments or subunits of enzymes that must be close to each other for the enzyme to be active. In the latter case, suitable reagents for attachment of the members of the interacting label pair to the appropriate probes and locations, may also be provided. Such reagents will be readily apparent to one of skill in the art.

In all compositions, probes A, B, and/or C may be designed to bind to specific nucleic acid sequences, or they may be the universal probes described above, which lack the sequences which bind to target nucleic acid sequence. Alternatively, the probes may contain part of a sequence which binds to the target nucleic acid. The remainder of the binding sequence of the probe(s) may be provided by the user in order to match the target sequence or sequences to be detected and/or quantified. Probe C, especially, is amenable to provision as a composition which is universal or nearly universal, since probe C does not contain sequences which bind to the target nucleic acid, and therefore the sequence of probe C is not substantially dependent on the sequence of the particular target nucleic acid to be detected. Suitable reagents and compounds to accomplish attachment of probe sequences to sequences which bind to target nucleic acid sequences, and/or extension of probe sequences to comprise sequences which bind to target nucleic acid sequences, may also be provided.

IDT Exhibit 1020.9668

In other embodiments, the invention provides compositions further comprising reference nucleic acid for detection and/or quantification of changes in nucleic acid. A composition may provide probe A, probe B, and/or probe C, in combination with or separate from a reference nucleic acid sequence. For example, a composition may provide probes A and C, and reference nucleic acid. Under suitable reaction conditions, probes A and C may form Complex II (see Figure 3) with the reference nucleic acid. Addition of probe B to the mixture may result in generation of a detectable signal, which may be compared to the signal generated by addition of probes A, B, and C to the target nucleic acid sequence under similar reaction conditions.

In yet other embodiments, probes of the composition may target more than one target nucleic acid. For example, a composition may contain more than one probe A, each of which targets a different nucleic acid sequence. A composition may further contain more than one probe B, each of which targets a different nucleic acid sequence. Probe C may be the same or it may be different for the various probes A and/or B provided. These compositions are useful in, for example, determining the presence or absence and/or quantity of more than one target nucleic acid sequence at a time, in determining the presence or absence and/or quantity of various alleles, or in determining the presence or absence or quantities of various polymorphisms or mutations in a sequence.

All compositions described herein may also comprise the mirror image probes depicted, for example, in Figures 6 through 10.

The compositions are generally in a suitable medium, although they can be in lyophilized form. Suitable media include, but are not limited to, aqueous media (such as pure water or buffers).

In some embodiments, the invention provides compositions comprising any of the complexes (which are generally considered as intermediates with respect to the final products) described herein (see also the figures for exemplary schematic depictions of these various complexes). Complexes are schematically depicted in Figures 2-10, and in Figure 13. As an example, one complex of the invention is a complex comprising: (a) probe A; and (b) probe C, wherein probe C is bound to probe A to from Complex I, as shown in Figure 2. Probe A may comprise a nucleic acid sequence that binds to target nucleic acid. Alternatively, probe A may comprise none or only part of a nucleic acid sequence which binds to target nucleic acid. If probe A comprises a nucleic acid sequence which binds to target nucleic acid, the complex may further comprise a nucleic acid which binds to probe

IDT Exhibit 1020.9669

A where target nucleic acid may bind, e.g., a reference nucleic acid. The complex, with or without a nucleic acid which binds to probe A, may also comprise probe B.

The invention also provides kits for conducting the methods described herein. These kits, in suitable packaging and generally (but not necessarily) containing suitable instructions, contain one or more components used in the detection and/or quantification methods, including mirror-image components. The kits may be used for any one or more of the uses described herein, and, accordingly, may contain instructions for any one or more of the following uses: determining whether a target nucleic acid sequence is present in a sample, detecting a target nucleic acid sequence, quantifying a target nucleic acid sequence, comparing target nucleic acid sequence to reference sequence(s), determining genotype, determining allele composition of target nucleic acid(s), detecting and/or quantifying multiple nucleic acid sequences, and use of the methods in conjunction with nucleic acid amplification techniques.

The kits of the invention comprise one or more containers comprising any combination of the components/reagents (e.g., probes) described herein, and the following are examples of such kits. A kit for detecting a target nucleic acid sequence may contain probe A, probe B, probe C, and optionally components for extension of probe B, if used. A kit for quantifying a nucleic acid sequence may further comprise a reference or set of references of known quantity, to which the target nucleic acid may be compared. A kit for comparing a target nucleic acid sequence with reference sequence(s) may further comprise said reference sequence(s). A kit for determining genotype and/or allele composition may further comprise a reference sequence corresponding to one of the alleles to be determined, or more than one reference corresponding to more than one allele. A kit for detecting and/or quantifying a multiple nucleic acid sequences may contain one or more sets of probe A, probe B, and/or probe C, each for a given target nucleic acid sequence, as well as optionally comprising a reference or set of references for one or more of the target sequences. Kits may also optionally include any of one or more of the enzymes described herein, substrates for enzymes (for example, when an enzyme is used a member of an interacting label pair) as well as deoxynucleoside triphosphates and/or ribonucleoside triphosphates. Kits may also include one or more suitable buffers (as 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

IDT Exhibit 1020.9670

component can be packaged in separate containers or some components can be combined in one container where cross-reactivity and shelf life permit.

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 present invention for the intended nucleic acid detection and/or quantification, and/or, as appropriate, for using the detection and quantification methods in conjunction with amplification techniques. 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 presentation invention, instructions on how to use the kit, and/or appropriate reaction conditions.

In another embodiment, the invention provides a kit for the detection or quantification of a target polynucleotide, comprising a first probe and a second probe, wherein the first probe comprises a 3' region that is hybridizable to a target polynucleotide comprising a sequence of interest, a 5' region comprising two mutually hybridizable sequences and members of an interacting moiety pair; the second probe comprises a 5' region which is hybridizable to the target polynucleotide, and a 3' region which is not hybridizable to the target polynucleotide and is hybridizable to a sequence in the 5' region of the first probe; wherein hybridization of the second probe to the target polynucleotide is associated with a decrease in extent of hybridization between the mutually hybridizable sequences of the first probe; whereby generation of a detectable signal associated with said decrease in extent of hybridization between the mutually hybridizable sequences indicates presence of the sequence of interest in the sample. This kit may further comprise a reference polynucleotide to which the target polynucleotide may be compared instructions for use of the kit to detect or quantify the target polynucleotide, and/or a nucleic acid polymerase with strand displacement activity.

The component(s) of the kits of the invention may be packaged in any convenient, appropriate packaging. The components may be packaged separately, or in one or multiple combinations. 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. In kits made up of the reagents of the present invention, probe A and probe C (or a unitary probe which combines the sequences of

IDT Exhibit 1020.9671

probe A and probe C) may be provided with or without interacting label pairs, and suitable members of an interacting label pair may be attached before assay; this is useful when several different interacting label pairs are to be used, and/or members of the interacting label pair are more subject to degradation than the rest of the probe molecule.

The invention also provides systems for effecting the detection and/or quantification methods described herein. These systems comprise various combinations of the components discussed above. For example, in some embodiments, the invention provides a system suitable for detecting target nucleic acid sequence (or quantifying target polynucleotide sequence) comprising (a) probes A, B, and C (as described herein), and (b) optionally, a polynucleotide polymerase (preferably DNA polymerase) and nucleoside triphosphates for probe extension. In some embodiments, a single probe may combine the functions of probes A and C. In some embodiments, more than one probe of any or each type may be provided in order to target multiple nucleic acid sequences. In some embodiments, the system further comprises reference nucleic acid sequence(s). In some embodiments, the system further comprises a standard or set of standards for quantifying nucleic acids. In some embodiments the system may also comprise interacting label pairs, as described herein, to be attached to probes before use. In some embodiments, the system may comprise components used for attaching a sequence complementary to the target nucleic acid sequence to universal probes which lack some or all of such sequence. In all embodiments, the system may comprise suitable buffers, enzymes, substrates, and/or apparatus and other components necessary to provide the proper reaction conditions for detection and/or quantification of target nucleic acid.

In another aspect, the invention provides reaction mixtures (or compositions comprising reaction mixtures) which contain various combinations of components (e.g., probes) described herein. In some embodiments, the invention provides reaction mixtures comprising (a) probe A, probe B, and probe C, and (b) optionally, a nucleic acid polymerase. Such reaction mixtures may further comprise reference nucleic acid(s). In addition, such reaction mixtures may contain the necessary nucleoside triphosphates for strand extension. In some embodiments, the invention provides reaction mixtures comprising a) a unitary probe combining the functions of probe A and probe C, and probe B, and (b) optionally, a nucleic acid polymerase. Such reaction mixtures may further comprise reference nucleic acid(s). In addition, such reaction mixtures may contain the necessary nucleoside triphosphates for strand extension. In all cases, reaction mixtures

IDT Exhibit 1020.9672

may further comprise buffers, enzymes, enzyme substrates, and other components necessary for detection and/or quantification of a target polynucleotide.

In one embodiment, the invention comprises a reaction mixture, comprising a first probe, a second probe and a third probe, wherein the first probe comprises a 3' region that is hybridizable to a target polynucleotide comprising a sequence of interest and a 5' region comprising a first member of an interacting moiety pair; the second probe comprises a 5' region which is hybridizable to the target polynucleotide, and a 3' region which is hybridizable to a sequence in the 5' region of the first probe; the third probe comprises a sequence which is hybridizable to a sequence in the 5' region of the first probe and a second member of the interacting moiety pair; wherein hybridization of the second probe to the target polynucleotide is associated with displacement of a sequence of the third probe from the first probe; whereby generation of a detectable signal associated with displacement of the third probe from the first probe indicates presence of the sequence of interest in the sample; and further comprising a nucleic acid polymerase with strand displacement activity.

In another embodiment, the invention comprises a reaction mixture, comprising a first probe and a second probe, wherein the first probe comprises a 3' region that is hybridizable to a target polynucleotide comprising a sequence of interest, a 5' region comprising two mutually hybridizable sequences and members of an interacting moiety pair the second probe comprises a 5' region which is hybridizable to the target polynucleotide, and a 3' region which is not hybridizable to the target polynucleotide and is hybridizable to a sequence in the 5' region of the first probe; wherein hybridization of the second probe to the target polynucleotide is associated with a decrease in extent of hybridization between the mutually hybridizable sequences of the first probe; whereby generation of a detectable signal associated with said decrease in extent of hybridization between the mutually hybridizable sequences indicates presence of the sequence of interest in the sample; and further comprising a nucleic acid polymerase with strand displacement activity.

The following Examples are provided to illustrate but not limit the present invention.

IDT Exhibit 1020.9673

EXAMPLES

Example 1

Detection of the presence of target nucleic acid sequence

The sequence of the target nucleic acid is as follows:

5′-AGATTCAGTG AAGGATAGTG CTGAATTTCC ATCTCTGAGT TCAAAATAAT
TTGAGAAAAT ATGATAGAAA TTGTGAAGTA CTAGATTTCA GAAAATATGA
TTAGAAAAAG CGTGGTACAT CTTAAAATAT AACTT TCTGG CACTGAATCC
TGCCACCTAC CATCTCCACC TCTAACATGG ATGCAGTTTA TAATTAAGAG
CTTTCAGATT TTGAGCA-3′

The underlined sequence is the sequence complementary to the corresponding sequences (underlined) of probes A and probe B.

Probe A
3′-CTTTTATACTAATCTTTTTCGCACC
CATCATCATCAT*CATCATCATGTAATCGTAATCGAAGAC*-FL

Probe C
5′- **GTAGTAGTACATTAGCATTAGCTTCTG**-Q

Probe B
3′-ATGATGATGATGTAGAATTTTATATTGTT-5′

Probe C is complementary to the 5'-end sequence of probe A. The 3'-most sequence of probe B is complementary to the sequence of probe A which is in the 5'-direction of the non-target complementary sequence (in italics).

In this example the displacement or blocking of probe C from probe A to disrupt quenching of fluorescence of probe F by quencher Q, is carried out by extension of the 3'-end of probe B hybridized to the non target complementary sequence of probe A, when probe A and probe B are hybridized to the same strand of the target nucleic acid, by a DNA polymerase.

Sample suspected of containing the test nucleic acid sequence is combined with a reaction-mixture composed of Probes A, Probe B and Probe C, dNTPs and DNA polymerase in a buffer. The buffer composition is similar to that used primer extension by

IDT Exhibit 1020.9674

DNA polymerase. The reaction mixture is heated to 95^{0}C for 3 min. for denaturation of the target nucleic acid and the mixture is further incubated at 55^{0}C for 10 min.

If the target nucleic acid sequence is present in the sample, probe A and Probe B will hybridize to the corresponding sequences on the target. Probe C hybridizes to probe A and the labels F and Q are brought to close proximity, which results in quenching of the fluorescence signal of F. Probe C is in large excess over probe A, to ensure binding of more than 99.999% of probe A. Probe A and probe B are present in amount of 1 to 10 pmoles.

Hybridization of probe A and probe B to the same strand of the target nucleic acid enhances the hybridization of the 3'-end of probe B to the corresponding non-target complementary sequence of probe A. DNA polymerase binds to the hybridized 3'-end of probe B and extends it along the 5'-end strand of probe A. The extension of probe B by a DNA polymerase results in the displacement or blocking of probe C from its hybridization to probe A, thus leading to dissociation of Q from F. The presence, and amount of the target nucleic acid is assessed by measuring the fluorescence of the reaction mixture. The fluorescence is measured either at the end of the incubation or in real time during the incubation period. The amount of target is proportional to the measured fluorescence.

Example 2

Detection and quantification of target nucleic acid sequence in a sample using the method of the present invention for the detection and quantification of products of amplification reaction

Any method of nucleic acid sequence amplification may be used for the amplification of a target nucleic acid sequence. These methods include PCR, isothermal exponential amplification methods such as NASBA, TMA or SDA, or isothermal linear amplification (U.S. Patent No. 6,251,639). The detection and quantification of the amplification products may be carried out simultaneously with the amplification reactions, or in a separate step following the amplification reaction. When the detection and quantification of the amplification products are carried out by the method described in example 1, i.e. generation of fluorescent signals by displacement or blocking of a probe that is labeled with a quencher (probe C) by polymerase catalyzed probe extension, the DNA polymerase employed in the amplification step is used in the signal generation step. The conditions for amplification of the target nucleic acid sequence to be detected or quantified are known to

IDT Exhibit 1020.9675

those skilled in the art. The quantification of amplification products is directly related to the amount of target nucleic acid in the sample. The efficiency of amplification of the target nucleic acid sequence may be separately determined. Control samples with known amounts of the target nucleic acid sequence may be subjected to amplification and quantification by the method of the present invention to produce a standard curve that can be used for the determination of the amount of the test sample. Quantification of a test nucleic acid in a sample may also be carried out by simultaneous amplification and detection of a test nucleic acid sequence and a control nucleic acid sequence that is different than the test sequence and serves as an internal control for quantification. Thus, the reaction mixture contains a control nucleic acid sequence, the one or more primers required for amplification of the control and test nucleic acid sequences, the probes required for the detection and quantification of the amplification products generated from the test and control nucleic acid sequence, the enzyme required for amplification of the test and control nucleic acid sequences, dNTPs or rNTPs, any accessory proteins, such as single stranded DNA binding protein and the like, and the buffer components. The amplification and detection of amplification products of the test and control nucleic acid sequences are carried out in the same reaction vessel. In the case of PCR amplification, the reaction mixtures are subjected to incubations at varied temperatures, as commonly known in the art. The incubations are carried out in a thermo-cycler. The incubation temperatures are selected for optimal denaturation (80 to 99^{0}C), annealing of the primers (45^{0}C to 72^{0}C) and primer extension (60^{0}C to 75^{0}C). The duration and temperature for each of the incubation steps are determined according to the sequence of the test nucleic acid and the selected primers and probes. A commonly used medium for PCR amplification comprises 1 to 5 mM $MgCl_2$, Tris buffer at pH 8.5, and 0 to 50 mM KCl. The reaction is carried out in the presence of 250 nM of each dNTP and 0.1 to 1 uM of each of the primers. The reactions are commonly carried out using Taq DNA polymerase, or other suitable thermostable DNA polymerases such as Pfu, Vent and the like.

Isothermal Single Primer Amplification (SPIA) can be also employed for amplification of the test and/or control nucleic acid and detection and quantification of the amplification products according to the method of the invention. A single composite primer, comprising a 3' DNA portion and a 5' RNA portion, is used for amplification of a defined sequence. Specific primers for the control and test nucleic acid sequences are employed. The reaction is carried out in a buffer similar to that described for the PCR reaction, with 2 to 5 mM

IDT Exhibit 1020.9676

$MgCl_2$, 0.25 to 0.5 mM dNTPs, 3 ug T4gp32 (USB) or similar ssDNA binding protein, a DNA polymerase with strong strand displacement activity, such as Bca or Bst polymerases, RNase H , and 1 to 5 mM DTT. The reaction mixtures containing the primers, probes and samples and/or controls, are first denatured by incubation at 95^0C for 2 to 5 min., and the primer(s) are allowed to anneal to the respective target by incubation at 55^0C for 5 min. The enzyme mixture is than added to the reaction tubes and the amplification and signal generation and detection is carried out by further incubation at this temperature for 30 min.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding it, it will be apparent to those skilled in the art that certain changes and modifications could be practiced.

IDT Exhibit 1020.9677

CLAIMS

What is claimed is:

1. A method of determining whether a target nucleic acid sequence is present in a sample, said method comprising:

contacting said sample with a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present; and

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

IDT Exhibit 1020.9678

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

2. A method of determining whether a target nucleic acid sequence is present in a sample, said method comprising:

contacting said sample with a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a first region of the target nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present; and

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first

IDT Exhibit 1020.9679

probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

3. The method of claim 1 or 2, wherein said first and second regions of the target nucleic acid sequence are contiguous.

4. The method of claim 1 or 2, wherein said first and second regions of the target nucleic acid sequence are 1 nucleotide apart.

5. The method of claim 1 or 2, wherein said first and second regions of the target nucleic acid sequence are 3 or fewer nucleotides apart.

6. The method of claim 1 or 2, wherein said first and second regions of the target nucleic acid sequence are 7 or fewer nucleotides apart.

7. The method of claim 1 or 2, wherein the interacting label pair comprises donor and acceptor moieties.

8. The method of claim 1 or 2, wherein the interacting label pair comprises an enzyme.

9. The method of claim 8, wherein the interacting label pair comprises an enzyme and an inhibitor of said enzyme.

10. The method of claim 1 or 2, wherein the interacting label pair is capable of energy transfer.

11. The method of claim 1 or 2, wherein the interacting label pair comprises a fluorophore and a quencher.

12. The method of claim 1 or 2, wherein the method further comprises measuring the magnitude of the signal generated, whereby said magnitude indicates the quantity of the target nucleic acid sequence.

IDT Exhibit 1020.9680

13. The method of claim 1, wherein the method further comprises extension of the 3' end of the second probe along a sequence in the 5' region of the first probe using a nucleotide polymerase.

14. The method of claim 1 or 2, wherein the target nucleic acid sequence is attached to an analyte.

15. The method of claim 1 or 2, wherein the target nucleic acid sequence is attached to a solid support.

16. The method of claim 1 or 2, wherein the target nucleic acid sequence comprises DNA.

17. The method of claim 1 or 2, wherein the target nucleic acid sequence comprises RNA.

18. The method of claim 1 or 2, wherein the target nucleic acid sequence comprises DNA and RNA.

19. The method of claim 1 or 2, wherein the target nucleic acid sequence comprises PNA.

20. The method of claim 1 or 2, wherein at least one of the probes comprises DNA.

21. The method of claim 1 or 2 , wherein at least one of the probes comprises RNA.

22. The method of claim 1 or 2, wherein at least one of the probes comprises DNA and RNA.

23. The method of claim 1 or 2, wherein at least one of the probes comprises PNA.

24. The method of claim 1 or 2, wherein the region of the second probe which is hybridizable to a sequence of the first probe comprises a modified nucleotide that

IDT Exhibit 1020.9681

causes enhanced affinity to the sequence in the region of the first probe relative to an unmodified nucleotide.

25. The method of claim 1 or 2, wherein said detectable signal is of a greater magnitude than a detectable signal associated with the interacting label pair when the third probe is hybridized to the first probe.

26. The method of claim 1 or 2, wherein said detectable signal is of a lesser magnitude than a detectable signal associated with the interacting label pair when the third probe is hybridized to the first probe.

27. The method of claim 1 or 2, wherein the method further comprises amplifying the target nucleic acid sequence.

28. The method of claim 1 or 2, wherein the first probe or the second probe is allele-specific.

29. The method of claim 1 or 2, wherein the first probe and the second probe are allele-specific.

30. A method of determining whether a target nucleic acid sequence is present in a sample, said method comprising:
contacting said sample with a first probe, a second probe, a third probe, and a nucleotide polymerase, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of the target nucleic acid sequence, and allowing nucleic acid polymerization, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present,

(b) a first member of an interacting label pair; and

(ii) the second probe comprises a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present; and

(iii) the third probe comprises

IDT Exhibit 1020.9682

(a) a polynucleotide sequence which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence;

and wherein the 3' end of said second probe is extended by said nucleotide polymerase along the first probe by template switching, causing dissociation of the first and second members of the interacting label pair;

whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

31. A method of determining whether a target nucleic acid sequence is present in a sample, said method comprising:

contacting said sample with a first probe and a second probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of a first nucleotide sequence in the 5' region of the first probe with a second nucleotide sequence in the 5' region of the first probe, in the absence of the target nucleic acid sequence, and hybridization of at the second probe the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein:

(i) the first probe comprises a polynucleotide comprising

(a) a 3' region which hybridizes to a first region of the target nucleic acid sequence, if present, and

(b) a 5' region comprising said first and second nucleotide sequences which hybridize to each other in the absence of the target nucleic acid sequence, and two members of an interacting label pair;

(ii) the second probe comprises a polynucleotide comprising

(a) a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present and

IDT Exhibit 1020.9683

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target polynucleotide is present;

wherein, when said first and second nucleotide sequences in the 5' region of the first probe hybridize, said first and second members of the interacting label pair are brought into proximity and interact; and

wherein, in the presence of the target nucleic acid sequence, said first probe and said second probe hybridize to the target nucleic acid sequence and said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

32. A method of determining whether a target nucleic acid sequence is present in a sample, said method comprising:

contacting said sample with a first probe and a second probe, under conditions allowing hybridization of the first and second probes to the target nucleic acid sequence, if present, and hybridization of a first nucleotide sequence in the 3' region of the first probe with a second nucleotide sequence in the 3' region of the first probe, in the absence of the target nucleic acid sequence, and hybridization of at the second probe the first probe when the second probe is hybridized to the target nucleic acid sequence, wherein:

(i) the first probe comprises a polynucleotide comprising

(a) a 5' region which is hybridizes to a first region of the target nucleic acid sequence, if present, and

(b) a 3' region comprising said first and second nucleotide sequences which hybridize to each other in the absence of the target nucleic acid sequence, and two members of an interacting label pair;

(ii) the second probe comprises a polynucleotide comprising

(a) a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present and

(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target polynucleotide is present;

IDT Exhibit 1020.9684

wherein, when said first and second nucleotide sequences in the 3' region of the first probe hybridize, said first and second members of the interacting label pair are brought into proximity and interact; and

wherein, in the presence of the target nucleic acid sequence, said first probe and said second probe hybridize to the target nucleic acid sequence and said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

33. A method of determining whether a target nucleic acid sequence contains a sequence alteration relative to a reference nucleic acid sequence, said method comprising:

(a) contacting said target nucleic acid sequence with a first probe, a second probe, and a third probe, and

(b) contacting said reference nucleic acid sequence with a first probe, a second probe, and a third probe,

wherein said contact of the target nucleic acid sequence with the first, second, and third probes, and said contact of the reference nucleic acid sequence with the first, second, and third probes, occur under conditions allowing hybridization of the first and second probes to the reference nucleic acid sequence, and hybridization of the first probe to the third probe, in the absence of the reference nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to the reference nucleic acid sequence, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of the reference nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a second region of the reference nucleic acid sequence, if present, and

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the reference nucleic acid sequence is present; and

IDT Exhibit 1020.9685

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;
and wherein, in the presence of the reference nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the reference nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the reference nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;
and comparing detectable signal generated by contacting said first, second, and third probes with the reference nucleic acid sequence with detectable signal generated by contacting said first, second, and third probes with target nucleic acid sequence;
whereby reduced signal generation by the target nucleic acid sequence as compared to the reference nucleic acid sequence indicates the presence of an altered sequence in the target nucleic acid sequence relative to the reference nucleic acid sequence.

34. A method of determining whether a target nucleic acid sequence contains a sequence alteration relative to a reference nucleic acid sequence, said method comprising:

(a) contacting said target nucleic acid sequence with a first probe, a second probe, and a third probe, and

(b) contacting said reference nucleic acid sequence with a first probe, a second probe, and a third probe,

wherein said contact of the target nucleic acid sequence with the first, second, and third probes, and said contact of the reference nucleic acid sequence with the first, second, and third probes, occur under conditions allowing hybridization of the first and second probes to the reference nucleic acid sequence, and hybridization of the first probe to the third probe, in the absence of the reference nucleic acid sequence,

IDT Exhibit 1020.9686

and hybridization of the second probe to the first probe when the second probe is hybridized to the reference nucleic acid sequence, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a first region of the reference nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a second region of the reference nucleic acid sequence, if present, and

(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the reference nucleic acid sequence is present; and

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the reference nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the reference nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the reference nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

and comparing detectable signal generated by contacting said first, second, and third probes with the reference nucleic acid sequence with detectable signal generated by contacting said first, second, and third probes with target nucleic acid sequence;

whereby reduced signal generation by the target nucleic acid sequence as compared to the reference nucleic acid sequence indicates the presence of an altered sequence in the target nucleic acid sequence relative to the reference nucleic acid sequence.

35. A method of determining whether one or more of a plurality of target nucleic acid sequences is present in a sample, said method comprising:

IDT Exhibit 1020.9687

contacting said sample with a plurality of probe sets, each set comprising a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to a target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of said target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to said target nucleic acid sequence, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of said target nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a second region of said target nucleic acid sequence, if present, and

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if said target nucleic acid sequence is present;

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of said target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of said target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of said target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of said target nucleic acid sequence; and

wherein each probe set comprises an interacting label pair which generates a detectable signal which is different from the signals of the interacting label pairs of every other probe set, and generation of two or more signals indicates presence of a plurality of target nucleic acid sequences.

IDT Exhibit 1020.9688

36. A method of determining whether one or more of a plurality of target nucleic acid sequences is present in a sample, said method comprising:
contacting said sample with a plurality of probe sets, each set comprising a first probe, a second probe, and a third probe, under conditions allowing hybridization of the first and second probes to a target nucleic acid sequence, if present, and hybridization of the first probe to the third probe, in the absence of said target nucleic acid sequence, and hybridization of the second probe to the first probe when the second probe is hybridized to said target nucleic acid sequence, wherein:
(i) the first probe comprises
(a) a polynucleotide comprising a 5' region which hybridizes to a first region of said target nucleic acid sequence, if present, and
(b) a first member of an interacting label pair;
(ii) the second probe comprises
(a) a polynucleotide comprising a 3' region which hybridizes to a second region of said target nucleic acid sequence, if present, and
(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if said target nucleic acid sequence is present;
(iii) the third probe comprises
(a) a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and
(b) a second member of an interacting label pair;
wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;
and wherein, in the presence of said target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of said target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of said target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;
whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of said target nucleic acid sequence; and

IDT Exhibit 1020.9689

wherein each probe set comprises an interacting label pair which generates a detectable signal which is different from the signals of the interacting label pairs of every other probe set, and generation of two or more signals indicates presence of a plurality of target nucleic acid sequences.

37. A composition comprising a first probe, a second probe and a third probe, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of a target nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present;

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

38. A composition comprising a first probe, a second probe and a third probe, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a first region of a target nucleic acid sequence, if present, and

IDT Exhibit 1020.9690

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present;

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

39. The composition of claim 36 or 37, wherein the interacting moiety pair comprises a fluorophore and a quencher.

40. The composition of claim 36, further comprising a nucleotide polymerase.

41. The composition of claim 36 or 37, further comprising the target nucleic acid sequence.

42. The composition of claim 36 or 37, further comprising a reference nucleic acid sequence to which the target nucleic acid sequence is to be compared.

IDT Exhibit 1020.9691

43. A kit for determining whether a target nucleic acid is present in a sample or quantifying a target nucleic acid sequence, comprising a first probe, a second probe and a third probe, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a first region of a target nucleic acid sequence, if present, and

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 3' region which hybridizes to a sequence in the 5' region of the first probe, if the target nucleic acid sequence is present; and

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 5' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 3' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

44. A kit for determining whether a target nucleic acid is present in a sample or quantifying a target nucleic acid sequence, comprising a first probe, a second probe and a third probe, wherein:

(i) the first probe comprises

(a) a polynucleotide comprising a 5' region which hybridizes to a first region of a target nucleic acid sequence, if present, and

IDT Exhibit 1020.9692

(b) a first member of an interacting label pair;

(ii) the second probe comprises

(a) a polynucleotide comprising a 3' region which hybridizes to a second region of the target nucleic acid sequence, if present, and

(b) a 5' region which hybridizes to a sequence in the 3' region of the first probe, if the target nucleic acid sequence is present;

(iii) the third probe comprises

(a) a polynucleotide which hybridizes to a sequence in the 3' region of the first probe, and

(b) a second member of an interacting label pair;

wherein when said third probe is hybridized to said first probe, said first and second members of the interacting label pair are brought into proximity and interact;

and wherein, in the presence of the target nucleic acid sequence, said 5' region of said first probe hybridizes to said first region of the target nucleic acid sequence and said 3' region of said second probe hybridizes to said second region of the target nucleic acid sequence and said 5' region of said second probe hybridizes to the first probe, causing dissociation of the first and second members of the interacting label pair;

whereby generation of a detectable signal caused by dissociation of the interacting label pair indicates presence of the target nucleic acid sequence.

45. The kit of claim 43 or 44, further comprising a reference nucleic acid sequence to which the target nucleic acid sequence may be compared.

46. The kit of claim 43 or 44, further comprising instructions for use of the kit to determine the presence of the target nucleic acid sequence in a sample or quantify the target nucleic acid sequence.

IDT Exhibit 1020.9693

Figure 1a: Probe A

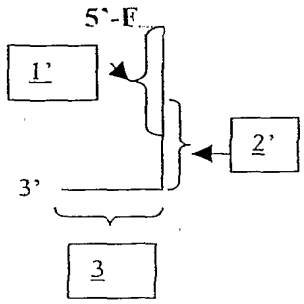


Figure 1b: Probe B

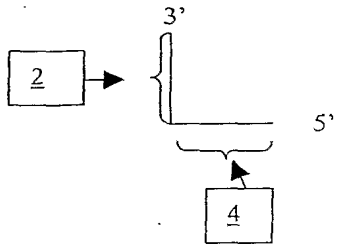


Figure 1c: Probe C

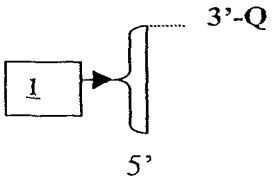


Figure 1

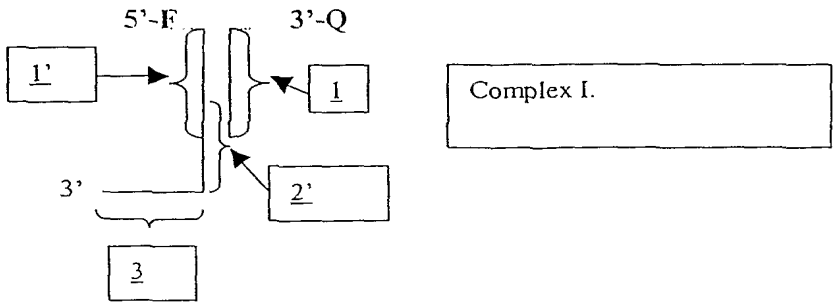


Figure 2

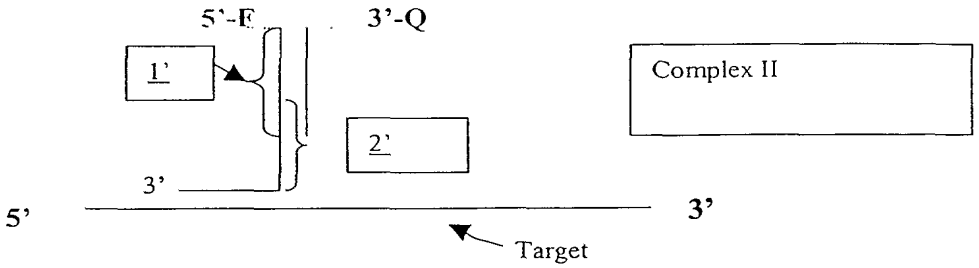


Figure 3

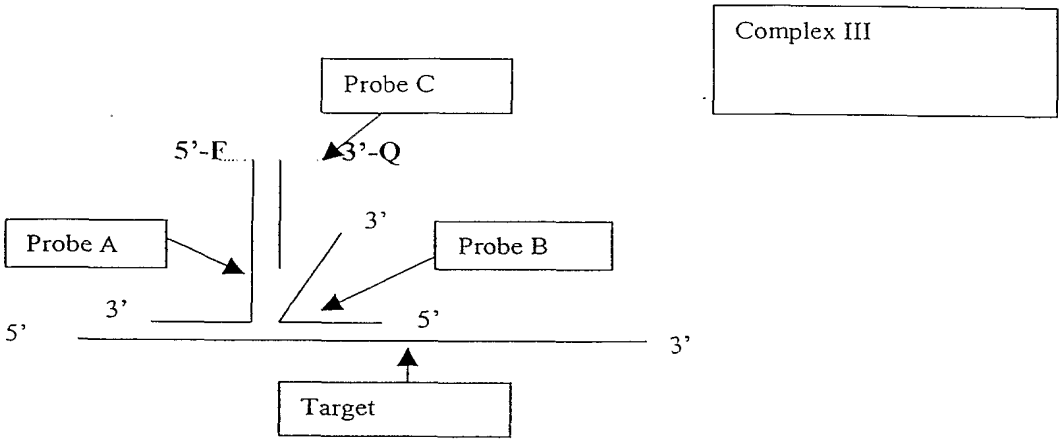


Figure 4

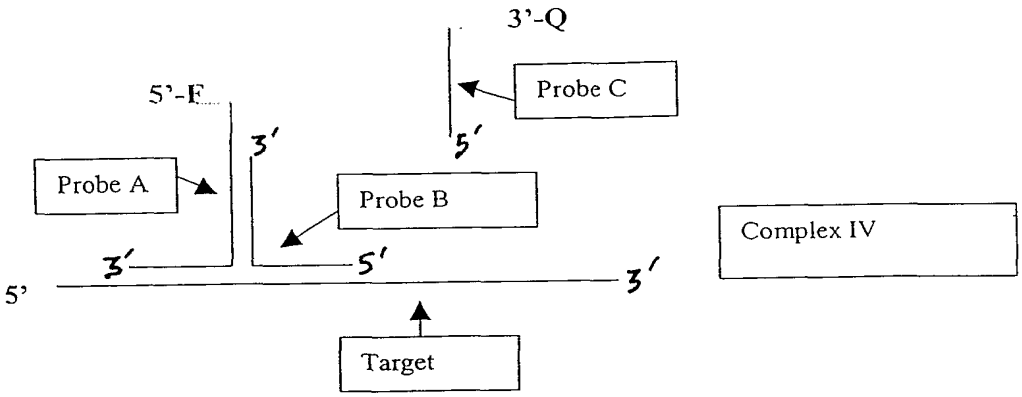


Figure 5

Figure 6a

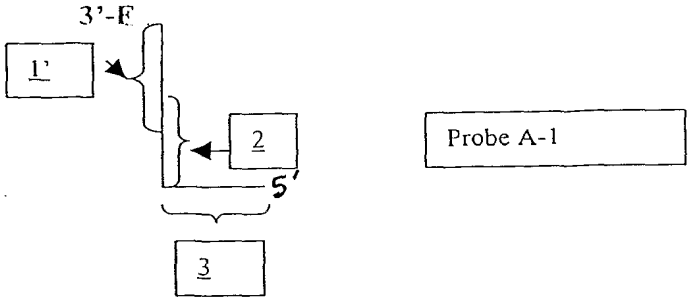


Figure 6b

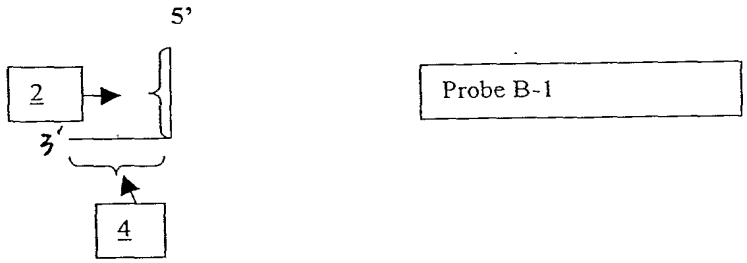


Figure 6c

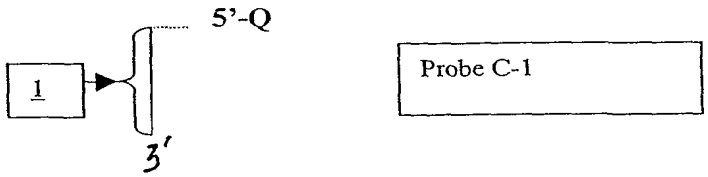
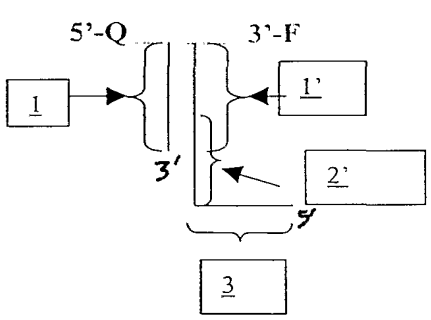


Figure 6



Complex I-1

Figure 7

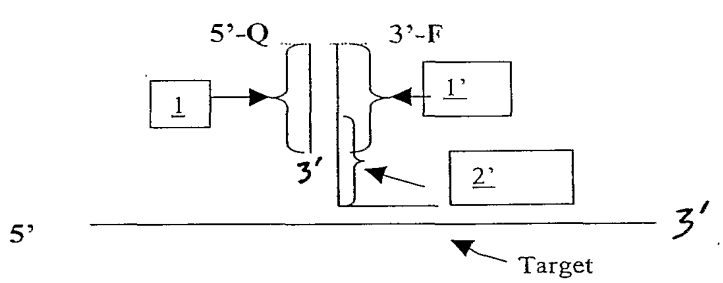


Figure 8

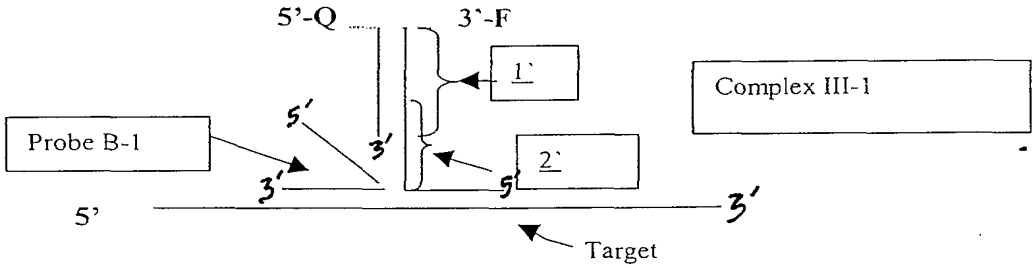


Figure 9

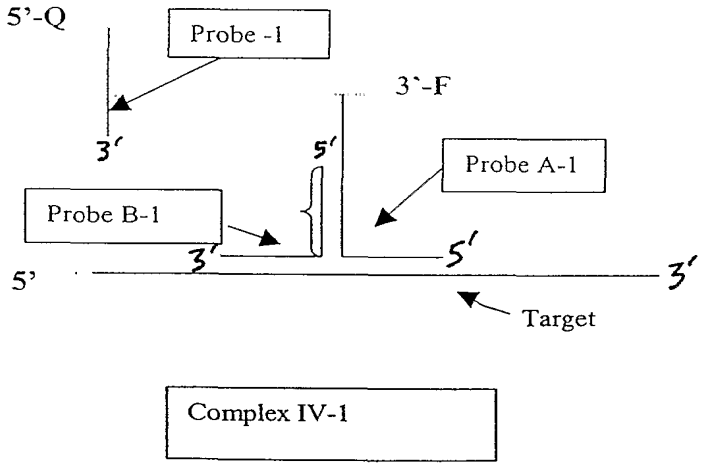


Figure 10

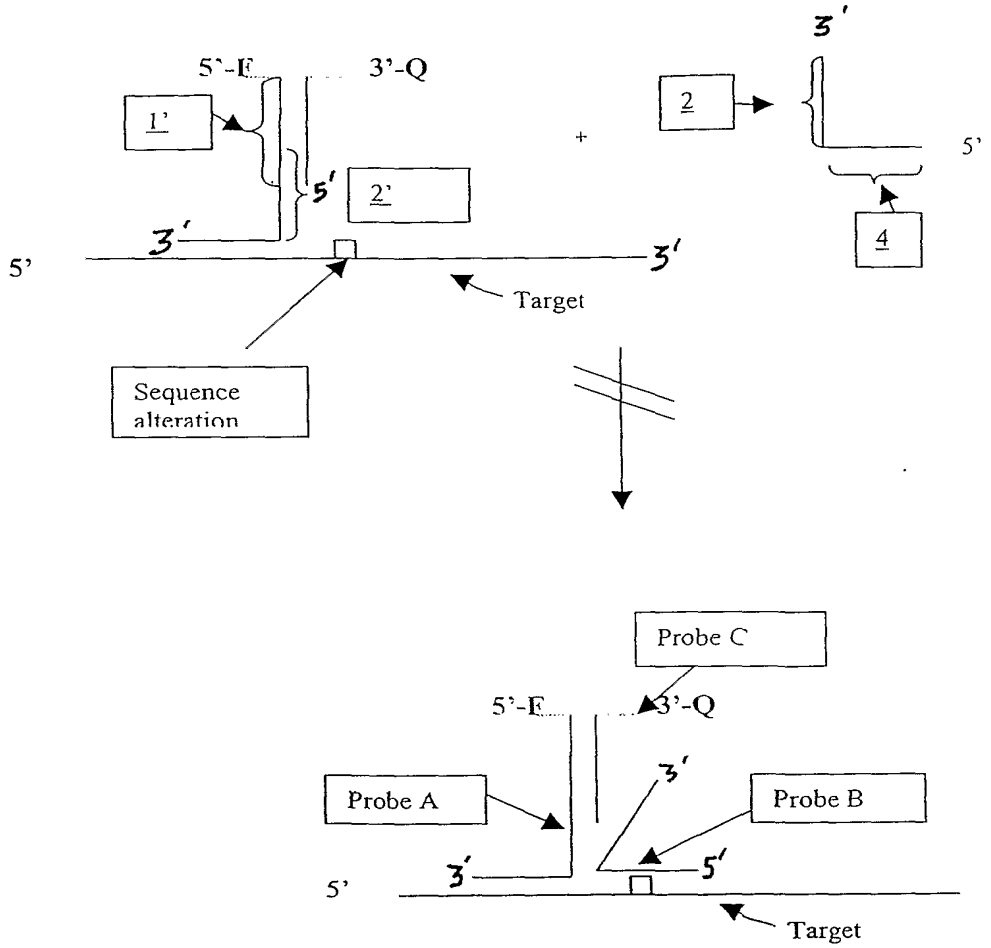

5'-E
3'-Q
1'
2'
5'
3'
5'
3'
Target
Sequence alteration
+
2
3'
5'
4
Probe C
5'-E
3'-Q
Probe A
Probe B
3'
3'
5'
5'
Target

Figure 11

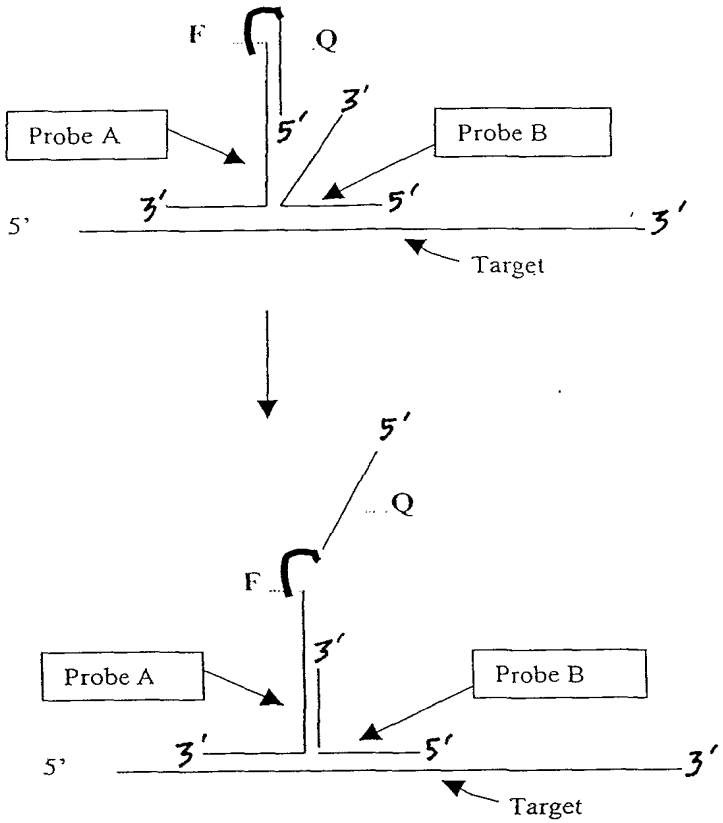

F
Q
Probe A
Probe B
5'
3'
3'
5'
5'
3'
Target
5'
Q
F
Probe A
Probe B
3'
3'
5'
5'
3'
Target

Figure 12

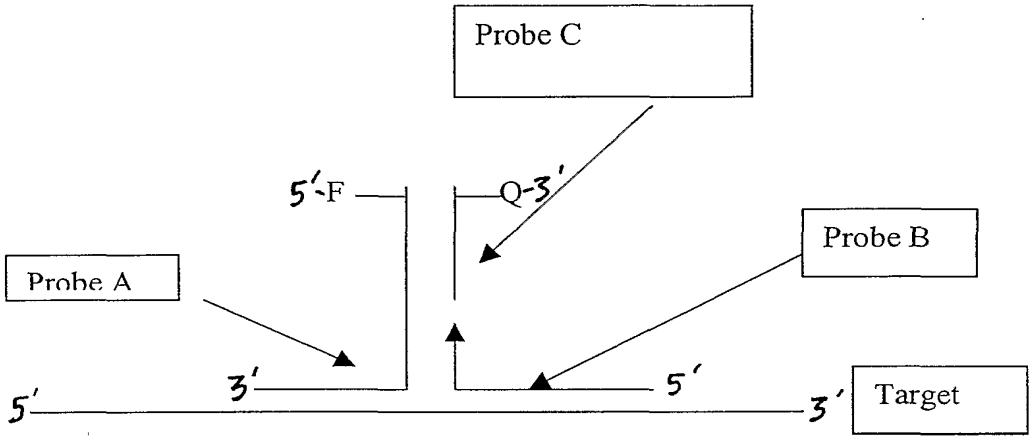


Figure 13

Espacenet

Bibliographic data: EP1071811 (A1) — 2001-01-31

A METHOD FOR THE CHARACTERISATION OF NUCLEIC ACID MOLECULES INVOLVING GENERATION OF EXTENDIBLE UPSTREAM DNA FRAGMENTS RESULTING FROM THE CLEAVAGE OF NUCLEIC ACID AT AN ABASIC SITE

Inventor(s): MCCARTHY THOMAS VALENTINE [IE]; VAUGHAN PATRICK MARTIN [IE] ± (MCCARTHY, THOMAS, VALENTINE, ; VAUGHAN, PATRICK, MARTIN)

Applicant(s): ENTPR IE TRD AS BIORESEARCH IE [IE]; UNIV COLLEGE CORK NAT UNIV IE [IE] ± (ENTERPRISE IRELAND (TRADING AS BIORESEARCH IRELAND), ; UNIVERSITY COLLEGE CORK-NATIONAL UNIVERSITY OF IRELAND, CORK)

Classification: **- international:** ***C12N15/09; C12Q1/68;*** (IPC1-7): C12Q1/68
- cooperative: **C12Q1/6827** more

Application number: EP19980917568 19980422 Global Dossier

Priority number(s): WO1998IE00030 19980422

Also published as: EP1071811 (B1) EP1071811 (B8) AT214104 (T) AU7075098 (A) AU753661 (B2) more

Abstract not available for EP1071811 (A1)
Abstract of corresponding document: WO9954501 (A1)

A method for characterising nucleic acid molecules comprises the steps of: i) introducing a modified base which is a substrate for a DNA glycosylase into a DNA molecule; ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site; iii) cleaving the DNA at the abasic site so as to generate an upstream DNA fragment that can be extended; and iv) incubating the extendible upstream fragment in the presence of an enzyme, for example a polymerase or a ligase, allowing for extension thereof and a template nucleic acid and analysing the resultant fragment(s). The invention provides a novel, versatile and simple method using the above-mentioned extendible upstream DNA fragments which allows characterisation of nucleic acids and which has advantages over existing methods. One of the most important uses (but not the only use) of the method according to the invention is to scan or check a fragment of DNA (target nucleic acid) for the presence or absence of a mutation.

IDT Exhibit 1020.9707

IDT Exhibit 1020.9708

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 071 811 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
06.03.2002 Bulletin 2002/10

(21) Application number: **98917568.2**

(22) Date of filing: **22.04.1998**

(51) Int Cl.[7]: **C12Q 1/68**

(86) International application number:
PCT/IE98/00030

(87) International publication number:
WO 99/54501 (28.10.1999 Gazette 1999/43)

(54) A METHOD FOR THE CHARACTERISATION OF NUCLEIC ACID MOLECULES INVOLVING GENERATION OF EXTENDIBLE UPSTREAM DNA FRAGMENTS RESULTING FROM THE CLEAVAGE OF NUCLEIC ACID AT AN ABASIC SITE

VERFAHREN ZUR CHARAKTERISIERUNG VON NUKLEINSÄUREMOLEKULEN, DAS DIE ERZEUGUNG VON VERLÄNGERBAREN AUFWÄRTSGELEGENEN DNS-FRAGMENTEN, DIE DURCH SPALTUNG DER NUKLEINSÄURE AN EINER ABASISCHEN STELLE ENTSTEHT, UMFASST

METHODE DE CARACTERISATION DE MOLECULES D'ACIDE NUCLEIQUE IMPLIQUANT LA GENERATION DE FRAGMENTS D'ADN D'AMONT EXTENSIBLES A PARTIR DU CLIVAGE D'ACIDE NUCLEIQUE AU NIVEAU D'UN SITE ABASIQUE

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(43) Date of publication of application:
31.01.2001 Bulletin 2001/05

(73) Proprietors:
- **Enterprise Ireland (trading as Bioresearch Ireland)**
  Dublin 9 (IE)
- **University College Cork-National University of Ireland, Cork**
  Cork (IE)

(72) Inventors:
- **McCARTHY, Thomas, Valentine**
  Cork (IE)
- **VAUGHAN, Patrick, Martin**
  Cork (IE)

(74) Representative: **Ryan, Anne Mary et al**
c/o Anne Ryan & Co. 60 Northumberland Road
Ballsbridge Dublin 4 (IE)

(56) References cited:
WO-A-96/30545 **WO-A-97/03210**
WO-A-97/12061

- **VAUGHAN P ET AL: "A novel process for mutation detection using uracil DNA-glycosylase" NUCLEIC ACIDS RESEARCH, vol. 26, no. 12, February 1998, pages 810-15, XP002088826**
- **MCGRATH A ET AL: "A DNA glycolase-based fingerprint for accurate identification of amplified DNA products and its application in the accurate diagnosis of infectiuos organisms" ANALYTICAL BIOCHEMISTRY, vol. 259, no. 2, - June 1998 pages 288-92, XP002088827**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 071 811 B1

Printed by Jouve, 75001 PARIS (FR)

IDT Exhibit 1020.9709

Description

Technical Field

[0001] The present invention relates to a method for characterising nucleic acid molecules which involves generating extendible upstream DNA fragments which result from the cleavage of nucleic acid at an abasic site.

Background Art

[0002] Characterisation of target nucleic acids is highly important for several reasons relating to confirmation of the presence or absence of a gene in a sample, confirmation of part or all of a nucleic acid sequence, and screening for the presence of known and unknown disease causing mutations leading to inherited disease and natural variations in DNA. Although there are many known methods for characterising nucleic acids and for the detection of unknown sequence changes, the increasing amount of new genetic information being generated makes it important to develop new, better and faster methods for characterisation of nucleic acids.

[0003] WO 97/03210 discloses the use of a DNA-glycosylase enzyme, which recognises a modified base, for the direct detection of known and unknown mutations in a target nucleic acid sample. The method typically involves amplifying a target nucleic acid sample using a combination of normal DNA precursor nucleotides and one or more modified precursor nucleotide(s) where the modified precursor nucleotide replaces one of the normal precursor nucleotides which is a substrate for a DNA glycosylase. Following excision of the modified base by the glycosylase, the resulting abasic site is cleaved and the products of the cleavage are analysed. This method allows detection of mutations at candidate loci. However, the method of WO 97/03210 has certain limitations. For example, with this method it is not possible to detect sequence differences between nucleic acid molecules without detecting sequence similarities and thus multiple samples cannot be combined for simultaneous analysis.

Disclosure of Invention

[0004] The invention provides a method for characterising nucleic acid molecules, which comprises the steps of:

i) introducing a modified base which is a substrate for a DNA glycosylase into a DNA molecule;

ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site;

iii) cleaving the DNA at the abasic site so as to generate an upstream DNA fragment that can be extended; and

iv) incubating the extendible upstream fragment in the presence of an enzyme allowing for extension thereof and a template nucleic acid and analysing the resultant fragment(s).

[0005] The invention provides a novel, versatile and simple method using the above-mentioned extendible upstream DNA fragments which allows characterisation of nucleic acids and which has advantages over existing methods as indicated in the following description.

[0006] One of the most important uses (but not the only use) of the method according to the invention is to scan or check a fragment of DNA (target nucleic acid) for the presence or absence of a mutation. The method essentially consists of i) the generation of the extendible upstream DNA fragments and ii) the subsequent use of these fragments in analysing a piece of DNA (e.g. detecting a mutation).

[0007] Preferably, the modified base used is uracil, hypoxanthine or 8-OH guanine.

[0008] Preferably, the modified bases are derived from modified precursor nucleotides which when incorporated into DNA generate said modified bases.

[0009] Thus, the preferred modified precursor nucleotides are dUTP, dITP and 8-OH dGTP which when incorporated into DNA generate the glycosylase substrate bases uracil, hypoxanthine and 8-OH guanine, respectively. Each of the modified precursor nucleotides is a base sugar phosphate comprising said base and a sugar phosphate moiety. Uracil in DNA is recognised specifically by uracil DNA-glycosylase (UDG) and released from DNA. UDG also recognises other uracil related bases when present in DNA. Hypoxanthine is recognised specifically by alkylpurine DNA-glycosylases (ADG) and released from DNA. This enzyme also recognises and releases N3 methyladenine, N3 methylguanine, O^2 methylcytosine and O^2 methylthymine when present in DNA. 8-OH guanine is recognised specifically by formamidopyrimidine DNA-glycosylase (FPG) and released from DNA. This enzyme also recognises and releases ring opened purines when present in DNA. Thymine DNA glycosylase recognises and releases uracil and thymine positioned opposite guanine bases in DNA.

IDT Exhibit 1020.9710

[0010] Modified precursor nucleotide(s) as used herein refers to a modified nucleotide or nucleotides that can be incorporated into a nucleic acid so that a modified base or bases is generated which is/are recognised and can be excised by a DNA glycosylase enzyme.

[0011] Following the introduction of the modified base, the DNA product is treated with a suitable DNA glycosylase enzyme which recognises and releases the glycosylase substrate base present in the target sample and consequently generates an apurinic or apyrimidinic (AP) site, depending on the nature of the modified base(s). AP site is the term given to a site in DNA where the base moiety of a nucleotide has been lost or removed, leaving behind a deoxyribophosphate with the DNA phosphodiester backbone still intact. AP is the abbreviation for either apurinic and/or apyrimidinic depending on whether a purine or pyrimidine base had been attached to the ribose ring. An AP site is also referred to as an abasic site, being the general term for apurinic and apyrimidinic site.

[0012] Release of the glycosylase substrate base from the nucleic acid sample results for example in an apyrimidinic site in the case of uracil and an apurinic site in the case of hypoxanthine and 8-OH dG. Collectively, such sites are referred to as abasic sites.

[0013] Essentially, cleaved fragments must have hydroxyl groups at the 3' termini and the DNA immediately downstream of said 3' termini must not be blocked in a way that prevents extension of the fragment from said 3' termini. Cleaved fragments are generated that have hydroxyl at the 3' termini and downstream blocking groups are removed that prevent extension of the fragment from the said 3' termini, while using as template, that DNA from which the extendible fragment was derived.

[0014] The DNA can be cleaved in a number of ways at the abasic site so as to generate said upstream DNA fragment as hereinafter described in greater detail.

[0015] For example, the phosphate linkages at the abasic sites can be cleaved by a treatment selected from treatment with a basic solution or other chemical treatment, heat treatment and/or treatment with an enzyme.

[0016] According to one embodiment of the invention, the upstream fragment is generated by cleaving the DNA at the 5' side of the abasic site, such that the 3' terminus of the upstream fragment bears a hydroxyl group.

[0017] The terms extendible fragment and upstream DNA fragment are used interchangeably herein.

[0018] As the 3'- OH termini are thereby generated no further processing of the upstream fragments is required prior to step iv).

[0019] Treatment with basic solutions (alkali) at high temperature or with a chemical such as piperidine, or with an enzyme which cuts specifically at abasic sites, such as *E. coli* endonuclease IV results in cleavage of the abasic site on the 5' side.

[0020] Suitably in this embodiment, the cleavage is achieved with a 5' AP endonuclease.

[0021] According to an alternative embodiment the upstream fragment is generated by cleaving at the 5' side of the abasic site so as to leave a phosphate group at the 3' terminus of the upstream fragment and removing the phosphate group so that the upstream fragment bears a hydroxyl group at the 3' terminus.

[0022] Alternatively, cleavage with alkali or the abasic endonuclease activity of FPG is used, followed by removal of the 3' phosphate. For example 3' phosphate groups can be removed by enzymatic means using enzymes with 3' phosphate activity such as T4 polynucleotide kinase.

[0023] According to a still further embodiment, the upstream fragment is generated by cleaving at the 3' side of the abasic site so as to generate a deoxyribose group phosphate at the 3' terminus of the upstream fragment and subsequently removing the deoxyribose group to leave a hydroxyl group at the 3' terminus.

[0024] This can be achieved by using an enzyme with 3' deoxyribophosphodiesterase activity, or using FPG followed by a 3' phosphatase.

[0025] In an alternative embodiment 5' deoxyribose moieties downstream of the 3' terminus of the upstream fragment are removed so that the upstream fragment can be extended on the template.

[0026] Preferably the 5' deoxyribose moieties are removed by a 5' deoxyribophosphodiesterase.

[0027] Treatment with high temperature alone or a 3' AP endonuclease results in cleavage of the abasic site to completion on the 3' side.

[0028] Glycosylase mediated cleavage cuts the extended primer at an abasic site subsequent to release of the modified base by a DNA glycosylase yielding 3' termini with 3'-OH or 3' phosphate groups or deoxyribophosphate groups. Except in cases where 3'-OH termini are generated, all other termini require further processing prior to extension of the upstream fragment.

[0029] Glycosylase mediated cleavage in the method according to the invention refers to both 5' and 3' cleavage, including whatever subsequent treatment that is necessary to generate a 3' OH group at the 3' terminus of the upstream fragment.

[0030] Preferably, the modified base is introduced by enzymatic amplification of the DNA.

[0031] Preferably, the DNA (herein also referred to as target nucleic acid) is amplified using normal DNA precursor nucleotides and at least one modified precursor nucleotide.

[0032] Precursor nucleotides in the case of a DNA amplification process refer to the deoxyribonucleotides dATP,

IDT Exhibit 1020.9711

dCTP, dGTP and dTTP herein referred to as "normal" DNA precursor nucleotides.

[0033] The term "amplifying" as used herein refers to any *in vitro* process for increasing the number of copies of a nucleotide sequence or sequences. Amplification of a target nucleic acid molecule results in the incorporation of precursor nucleotides into the DNA being amplified. Typically, amplification of a target sample is carried out using appropriate primers in the polymerase chain reaction (PCR). During the amplification, primers anneal to the target nucleic acid and are extended using a DNA polymerase in a 5' to 3' direction on the target nucleic acid which acts as a template for synthesis of new DNA. The use of flanking primers, which are referred to herein as initial primers, which anneal to the upper and lower strands of a DNA molecule permit exponential amplification of the DNA segment delimited by the upper and lower primers.

[0034] The amplification will typically involve amplifying a target nucleic acid using a combination of normal DNA precursor nucleotides and one or more modified precursor nucleotide(s) where the modified precursor nucleotide replaces all or a proportion of one of the normal precursor nucleotides. Amplification of a nucleic acid using normal DNA precursor nucleotides results in the incorporation of the four normal bases G,A,T, or C into DNA. Amplification of a nucleic acid using a modified precursor nucleotide in place of one of the normal precursor nucleotides results in the incorporation of a glycosylase substrate base into DNA in place of one of the four normal bases G,A,T, or C.

[0035] The target nucleic acid sample will typically be DNA. However, RNA may also be used following conversion to DNA by reverse transcription.

[0036] When a modified precursor nucleotide replaces a proportion of one of the normal precursor nucleotides, the ratio of the modified precursor nucleotide to the normal precursor nucleotide that it is replacing is such that an optimum of one modified precursor nucleotide is incorporated *per* strand of amplified DNA. This allows subsequent cleavage of the amplified DNA strand into two fragments following contact with a DNA glycosylase and an abasic site cleavage agent as herein described. This method of cleavage is herein referred to as glycosylase mediated cleavage. A higher ratio of the modified precursor nucleotide to normal precursor nucleotide is used to generate more than one cleavage site *per* amplified DNA strand. The incorporation of a modified precursor nucleotide intb the amplified product generates one or more modified bases at one or more positions recognised by a DNA glycosylase enzyme in the amplified product.

[0037] Replacement of all of a normal precursor nucleotide with a modified precursor nucleotide in the amplification reaction if used in step i) permits glycosylase mediated cleavage of a primer extended in an amplification reaction at the first position 3' of the extended primer where a normal base is replaced by a modified base. Thus, if the template sequence immediately 3' of a location where the primer hybridises is CTAG and the modified nucleotide precursor is dUTP replacing dTTP, then the modified base uracil (U) will be incorporated opposite A on the template strand. Thus, in this situation, the primer will have been extended by two nucleotides at the 3' end (primer - GA 3') following amplification and glycosylase mediated cleavage. These extended primers generated following the initial extension and glycosylase mediated cleavage are referred to herein *inter alia* as extendible fragments, as indicated above, whereas the primers prior to extension are referred to as the initial primers herein.

[0038] The extended 3' terminal sequence of the extendible fragment is enzymatically synthesised and is directly related to the nucleic acid being characterised as the nucleic acid acts as the template for its synthesis. Thus the 3 end of the extendible fragment is complementary to the nucleic acid. Accordingly, determination of the nature of the 3' end of the extendible fragment by any means allows characterisation of the nucleic acid from which it was derived. If an initial primer is placed adjacent to a locus where a DNA sequence variation such as a polymorphism or a mutation occurs so that the first modified base incorporated into the extended primer is at the mutation locus, then the initial primer will be extended to a different length depending on whether or not a mutation is present at the mutation locus following glycosylase mediated cleavage. The extendible fragments are subsequently treated (if necessary) so that they can be used as primers for a subsequent extension reaction. Because the sequence at the 3' termini of the extendible fragments differ depending on whether a mutation is present or absent at the mutation locus, analysis of the ability of an extendible fragment to function in a subsequent extension reaction using a template nucleic acid permits the determination of whether a mutation is present or absent at the mutation locus. Any naturally occurring or enzymatically or chemically synthesised template which fully or partially hybridises to the extendible fragment can be designed and/or selected as a template nucleic acid allowing the ability of the extendible fragment to function as a primer to be determined.

[0039] When a proportion of a normal precursor nucleotide is replaced with a modified precursor nucleotide in the amplification reaction, glycosylase mediated cleavage of the primer extended in an amplification reaction will yield a population of extendible fragments of various lengths since different molecules will be cleaved at different points depending on where the modified precursor nucleotide is incorporated. The length of each fragment is determined by the position of incorporation of the modified precursor nucleotide during extension from the 3' end of the initial primer.

[0040] Amplification of a target nucleic acid using the precursor nucleotides dATP, dCTP, dGTP, dTTP and a low amount of the modified precursor nucleotide dUTP results in an amplified DNA where thymine is replaced randomly by uracil. The uracil is incorporated in the newly synthesised DNA strand at positions complementary to adenine residues in the template DNA strand during the amplification process. Amplification of a target nucleic acid using the

IDT Exhibit 1020.9712

precursor nucleotides dATP, dCTP, dGTP, dTTP and a low amount of the modified precursor nucleotide dITP results in an amplified DNA where guanine is preferentially replaced randomly by hypoxanthine. The hypoxanthine is incorporated in the newly synthesised DNA strand at positions complementary to cytosine residues in the template DNA strand during the amplification process when the other precursor nucleotides are not limiting. Amplification of a target nucleic acid using the precursor nucleotides dATP, dCTP, dGTP, dTTP and a low amount of the modified precursor nucleotide 8-OH dGTP results in an amplified DNA where guanine is preferentially replaced randomly by 8-OH guanine. The 8-OH guanine is incorporated in the newly synthesised DNA strand at positions complementary to cytosine residues in the template DNA strand during the amplification process when the other precursor nucleotides are not limiting.

[0041] The amplified DNA strands can be separated for separate analysis of the respective strands. In addition, the separated strands can be immobilised, which can be achieved by several means. A common method for immobilisation and/or separation of DNA strands is by the use of the biotin streptavidin interaction, where normally, the DNA contains the biotin label and the streptavidin is attached to a solid support. However, the method according to the invention in its various steps is amenable to immobilisation formats that allows immobilisation of the upstream fragment, the strand bearing the upstream fragment, the strand complementary to the strand bearing the upstream fragment, the template nucleic acid, the target nucleic acid and the products generated from glycosylase mediated cleavage of amplified or extended nucleic acids bearing modified bases.

[0042] The modified base can be introduced by chemical modification of a nucleic acid, rather than by an amplification technique such as PCR.

[0043] Several methods exist where the treatment of DNA with specific chemicals modifies existing bases so that they are recognised by specific DNA glycosylase enzymes. For example, treatment of DNA with alkylating agents such as methylnitrosourea generates several alkylated bases including N3-methyladenine and N3-methylguanine that are recognised and excised by alkyl purine DNA-glycosylase. Treatment of DNA with sodium bisulfite causes deamination of cytosine residues in DNA to form uracil residues which are recognised and excised by uracil DNA-glycosylase. Treatment of DNA with ferrous sulphate and EDTA causes oxidation of guanine residues in DNA to form 8-OH guanine residues in DNA which are recognised and excised by formamido-pyrimidine DNA glycosylase.

[0044] Thus, bases present in the nucleic acid or indeed, the extendible upstream fragment generated in step iii) can be converted into modified bases by chemical means. A proportion or all of the cytosine residues can be readily converted to an uracil using sodium bisulfite thereby rendering the amplified sample susceptible to uracil DNA-glycosylase cleavage at the sites of cytosine conversion. If the upper or lower primer is synthesised so that it contains 5-methylcytosine rather than cytosine, in such a case the primer will be resistant to uracil DNA-glycosylase mediated cleavage since deamination of 5-methylcytosine occurs at a reduced rate by comparison with cytosine and generates a thymine rather than an uracil residue.

[0045] Prior to treatment with a suitable DNA glycosylase, double stranded DNA may be treated with exonuclease I. This treatment serves to digest the unused primers and any non-specific single stranded DNA amplification products thus improving the signal to noise ratio.

[0046] In the case where the modified precursor nucleotide is dUTP, the modified base uracil will be generated at thymine positions in the amplified target nucleic acid sample. Addition of uracil DNA-glycosylase to the sample releases the uracil from the sample. In the case where the modified precursor nucleotide is dITP, the modified base hypoxanthine will be generated at guanine positions in the amplified target nucleic acid sample. Addition of alkylpurine DNA-glycosylase to the sample releases the hypoxanthine from the sample. In the case where the modified precursor nucleotide is 8-OH dGTP, the modified base 8-OH guanine will be generated at guanine positions in the amplified target nucleic acid sample. Addition of formamido-pyrimidine DNA-glycosylase to the sample releases the 8-OH guanine from the sample.

[0047] Suitably a primer or one or more nucleotide(s) involved in the enzymatic amplification is labelled.

[0048] The initial primer used may be suitably labelled. Labelling of the primers can be performed by a variety of means including addition of a radioactive, fluorescent, or detectable ligand to the primer during or post primer synthesis.

[0049] In one embodiment of the invention, the enzyme used in step iv) is a polymerase which can be incubated with the extendible upstream fragment in the presence of one or more nucleotide(s).

[0050] Also in this embodiment suitably one or more of the nucleotide(s) of step iv) is a dideoxy nucleotide.

[0051] Also one or more of the nucleotides of step iv) can be labelled.

[0052] Various nucleic acid polymerases can be used to extend the 3' terminus of the extendible fragment on a template nucleic acid. Many polymerases are described in the literature that extend 3' termini of primers on a template DNA. For example, DNA polymerases isolated from phage and mesophilic and thermophilic bacteria can be used.

[0053] Several DNA polymerases, including T7 DNA polymerase, incorporate dideoxy terminator nucleotides as a well as normal precursor nucleotides during extension of primers. Thermophilic DNA polymerases are used routinely in amplification of nucleic acids through repeated cyclic extension of primers. The upstream DNA fragments generated in step iii) function as primers for all nucleic acid polymerases capable of extending standard nucleic acid primers.

[0054] The use of a labelled precursor nucleotide or dideoxy terminator nucleotide in any of the extension reactions

IDT Exhibit 1020.9713

facilitates detection of the extended extendible fragment. Direct DNA staining methods such as silver or ethidium bromide staining facilitate detection of all extension products after size separation based on electrophoretic mobility.

[0055] The ability of the extendible fragments to function in a subsequent extension reaction using a template nucleic acid and normal or dideoxy terminator nucleotides (a nucleotide that prohibits further extension of a primer on a template once incorporated) generates a ladder of fragments allowing determination of the location of the total number of positions of the modified precursor nucleotide in one or both strands of the amplified target nucleic acid. The presence of a sequence variation or mutation results in the appearance or disappearance of a cleavage fragment as judged by comparison with the known DNA sequence of the amplified molecule. Size analysis of the fragments allows the precise location and sequence of a mutation in the target nucleic acid sample to be determined. Therefore, if a sequence variation occurs such that an additional site of modified precursor nucleotide incorporation is generated, an additional cleavage fragment will be observed upon analysis of the ladder of cleavage products. If a sequence variation occurs such that a site of modified precursor nucleotide incorporation is lost, the corresponding cleavage fragment will not be observed upon analysis of the ladder of cleavage products.

[0056] The template of choice in this case will be the originally intact or glycosylase cleaved amplified nucleic acid. In cases where glycosylase cleaved amplified nucleic acid is used it can be processed to remove residual moieties downstream of the extendible fragment which prohibit extension of the extendible fragment by a nucleic acid polymerase. Specifically the template DNA can be treated so that a residual 5' deoxyribose moiety is removed. This is achieved by incubation of the template DNA with a 5' deoxyribophosphodiesterase such as *E. coli* RecJ endonuclease or formamidopyrimidine DNA glycosylase. Other naturally occurring or enzymatically or chemically synthesised template nucleic acids which fully or partially hybridise to the extendible fragment can also be designed and/or selected as a template nucleic acid to determine the ability of the extendible fragment to function as a primer.

[0057] The ability of the extendible fragments to function in a subsequent extension reaction using a template nucleic acid and a combination of non-labelled and labelled normal or dideoxy terminator nucleotides permits detection of sequence variations and mutations. Extension of the extendible fragments on the template nucleic acid from which they were derived and which is heterozygous for the sequence variation, using a labelled dideoxy terminator nucleotide having base pairing properties different to those of the modified precursor nucleotide with a non-labelled dideoxy terminator nucleotide having the same base pairing properties as the modified precursor nucleotide allows detection of the variant or mutant loci alone whereas non variant loci are not detected. This aspect of the invention is particularly advantageous as detection of sequence variations alone permits very high throughput mutation scanning and detection, and allows fingerprinting of nucleic acids based on their sequence variations.

[0058] It will be appreciated that amplification of any target DNA, which is heterozygous for a mutation or polymorphism generates four distinct duplex DNAs, i.e. (taking as an example a G to A mutation at position X in a DNA sequence), one quarter will be homoduplex with a GC base pair at position X, one quarter will be homoduplex with an AT base pair at position X, one quarter will be heteroduplex with a GT base pair at position X, and one quarter will be heteroduplex with a AC base pair at position X. Similarly, heteroduplex DNA can be generated readily by denaturing and reannealing two homoduplex DNAs bearing a sequence difference(s).

[0059] Thus, the nature of the sequence of the 3'end of the extendible fragments can be determined by their ability to function as primers in a subsequent extension reaction using a template nucleic acid. Essentially, such determination is based on the ability of the 3'end of the extendible fragment to hybridise to a selected template under selected conditions. Following partial or complete hybridisation the extendible fragment may be extended using a nucleic acid polymerase and nucleic acid precursors or selected combinations of same as herein described. It will be appreciated that multiple possibilities exist for the selection of template molecules. Nonetheless, the extension of the extendible fragment is a measure of its hybridisation to or lack of hybridisation to a selected template molecule and thus the determination of the nature of the sequence of the 3'end of the extendible fragments is made on this basis, since this 3' sequence is indicative of the sequence of the original target nucleic acid.

[0060] Typically, the template molecule is selected so that it bears partial or full sequence complementarity to the upstream fragment. The upstream fragment may be extended one or more nucleotides on the template molecule using a nucleic acid polymerase and nucleic acid precursors or a combination of same or dideoxy terminator nucleotides or a combination of normal nucleic acid precursors and dideoxy terminator nucleotides.

[0061] The extension of step iv) can be achieved by means of an amplification reaction using said extendible DNA fragment.

[0062] Alternatively, the extension of step iv) is achieved by means of an amplification reaction including a primer in addition to using said extendible DNA fragment.

[0063] Repeated extension of an upstream fragment on a template nucleic acid in combination with a second flanking primer which can be extended on the copy of the template permits amplification of the template nucleic acid. Such amplified products can be readily detected by standard nucleic acid staining methods such as ethidium bromide after resolution by electrophoresis.

[0064] Alternatively, the template molecule may be selected so that it can be extended using the upstream fragment

IDT Exhibit 1020.9714

as a template and so that the extension is based on hybridisation to the 3' end of the upstream fragment.

[0065] Also, the upstream fragment may be analysed based on its ability to function in a 5' nuclease assay. During extension of the upstream fragment by a polymerase with 5' to 3' nuclease activity, the 5' to 3' nuclease activity degrades a downstream reporter molecule annealled to the same template strand as said upstream fragment.

[0066] A further possibility is for the upstream fragment to be extended on a synthetic template which contains reporter and quencher labels, then cleavage of the resulting double stranded DNA will release the reporter from the quencher and a signal will be detected. Typically such cleavage will be carried out by an enzyme which recognises the double stranded DNA molecule. Typically such an enzyme will be a restriction enzyme.

[0067] Furthermore, the 3' terminal sequence of the upstream fragment may be determined on the basis of its hybridisation to other nucleic acid molecules.

[0068] A further possibility is that the products extended or amplified using upstream fragments may be detected on the basis of their filtration and/or precipitation properties.

[0069] When analysing extension and incorporation of nucleotides in step iv) where the upstream fragments are used in an extension reaction, it is important to verify that any extension observed is specifically due to extension of the upstream fragments and not due to extension of the initial primers which were unused during the initial amplification, if such is used in step i), or extension of possible non-specific upstream fragments which may be generated by the non-specific breakage or damage of the DNA during previous steps of the procedure. To overcome this 'noise', the DNA, prior to glycosylase mediated treatment, can be treated with a single strand specific DNA nuclease, for example, Exonuclease I, which will degrade the unused primers and non-specific single stranded DNA and can be subsequently heat inactivated. The DNA can be treated with a 3' AP endonuclease/lyase which will cut the DNA and primers at any preexisting AP sites, generated specifically or generated through damage of the DNA. The 3' AP endonuclease/lyase is subsequently removed from the reaction. Because the endonuclease/lyase cuts at the 3' side, the resulting contaminating upstream fragments are not extendible and will not interfere with the extension of the subsequently generated upstream fragments. In addition to these treatments, a control reaction to check for non-specific extensions can be carried out. Thus, in step iii) of the method, the AP sites can be cut with a 3' AP endonuclease/lyase, thereby generating non-extendible upstream fragments. If however, extension and incorporation of nucleotides is observed in the subsequent step iv) then one can measure or determine the level of non-specific extension obtained during the procedure.

[0070] Similarly it is important to ensure that the incorporation of nucleotides, labelled or unlabelled, in step iv) is due to the incorporation of those nucleotides supplied during step iv) and not those from any previous step. This is especially important when the reaction involves incorporation of dideoxynucleotides. Prior to or subsequent to cleaving the DNA at abasic sites and generating upstream fragments, the reaction can then be treated with a phosphatase which digests all unincorporated nucleotides present in the reaction, e.g. shrimp alkaline phosphatase, which can subsequently be heat inactivated.

[0071] In a further embodiment of the invention, the enzyme used in step iv) is a ligase which can be incubated with the extendible upstream fragment in the presence of a reporter oligonucleotide.

[0072] The reporter oligonucleotide may be partially degenerate.

[0073] The method according to the invention can be used *inter alia* to detect a known or unknown mutation and to detect differences and similarities in genomes. These aspects of the invention are illustrated further below.

[0074] The method according to the invention provides in one aspect a means of generating random primers in a simple easy manner. Essentially, introduction of a modified base into an amplified DNA product followed by glycosylase mediated cleavage and subsequent treatment of the cleavage products so that they can be extended by a nucleic acid polymerase provides a rapid means of generating random primers. Subsequent use of such primers, i.e., the extendible upstream fragments, in random amplification of target nucleic acids allows amplification of discrete DNA molecules from the target nucleic acids thus permitting characterisation of the nucleic acid based on similarities and differences of the amplified products. Since these primers are essentially derived from said target nucleic acid, specifically their 3' ends, they are better primers for the subsequent random amplification analysis of said target nucleic acid and the amplification is more specific. Many discrete DNA products are generated during random amplification of nucleic acids. A discrete DNA product can be separated from other products on the basis of size. The method according to the invention permits the generation of primers from all or part of the separated product. Use of an initial primer in the random amplification of a nucleic acid that permits immobilisation of the separated product allows the isolation of the upper and lower primers extended to the first point of glycosylase mediated cleavage. The 3' end of such upstream fragments are derived from the target nucleic acid and thus permit more specific amplification of the target nucleic acid or related nucleic acids. The 3' end of such upstream fragments may be short or long. By short 3' ends herein is meant one to three nucleotides whereas by long 3' ends herein is meant greater than three nucleotides. Longer 3' ends on such upstream fragments are more desirable as they allow highly specific amplification of a target nucleic acid sequence. Upstream fragments generated with longer 3' ends may be selected by sizing methods. Alternatively, the initial amplification primers can be designed so that they promote binding of a protein that protects a section of the region 3' of the initial extended primer. Thus such a region is refractory to glycosylase mediated cleavage due to protection by

IDT Exhibit 1020.9715

the protein and inclusion of such a primer design and protein allows generation of upstream fragments with longer 3' ends.

[0075] This embodiment of the invention provides a rapid and simple method for generation of random and specific primers for nucleic acid amplification without prior knowledge of the nucleic acid sequence.

[0076] It will be appreciated that the random amplification of nucleic acid using arbitrarily chosen primers is a method already known for detecting similarities and differences between genomes. Such random amplification is based on the annealing of arbitrarily chosen primers to target samples followed by multiple rounds of enzymatic amplification whereby the primers are extended using the selected genomic DNA or cDNA as template. Using such primers in such a method results in the amplification of discrete DNA molecules. Analysis of such molecules allows the investigation of similarities and differences between different samples. Typically many different random primers are chosen for investigation of a genome or cDNA and such primers are synthesised chemically and are designed in a random fashion with the assumption that they will hybridise to the target nucleic acid in the amplification process. The method according to the invention allows the easy and rapid generation of primers from a target nucleic acid which can subsequently be used for random amplification of the same or different target nucleic acid.

[0077] In addition to extension by a polymerase reaction, as indicated above, the upstream fragments in step iv) can also be extended by ligation of another single stranded DNA molecule which results in extended upstream fragments greater in size than the initial upstream fragments generated by glycosylase mediated cleavage. The DNA molecule to which the upstream fragment is ligated is termed the reporter oligonucleotide herein and this can vary in length. Ligation of the upstream fragments to the reporter oligonucleotide is dependent on both DNA molecules (the upstream fragment and reporter oligonucleotide) annealing to a template molecule at adjacent sites so that the termini of the upstream fragment and reporter oligonucleotide are juxtaposed. This means that the 3' terminal base of the upstream fragment is juxtaposed to the 5' terminal base of the reporter oligonucleotide. The reporter oligonucleotide is typically a synthetic oligonucleotide, but can also be any other type of DNA molecule or RNA molecule.

[0078] The template molecule is selected so that it bears partial or full sequence complementarity to the upstream fragment and reporter oligonucleotide. The template of choice can be the originally intact or glycosylase cleaved amplified nucleic acid. In addition the template can be a single stranded DNA molecule, e.g. synthetic oligonucleotide, which can vary in length and which will allow complementary annealing of the upstream fragment(s) and reporter oligonucleotide(s). The template DNA can be single stranded or double stranded in nature. Double stranded DNA acting as template consists of the template strand and the complementary strand. Double stranded DNA must first be denatured and then allowed to reanneal in the presence of the upstream fragment and reporter oligonucleotide. The upstream fragment and reporter oligonucleotide then compete with the complementary strand in annealing to the template strand of the double stranded template molecule.

[0079] Various DNA ligases and RNA ligases can be used to extend the 3' terminus of the extendible fragment by ligation of a reporter oligonucleotide on a template nucleic acid. DNA ligases from many sources, including those isolated from phage, e.g. T4 DNA ligase, and mesophilic and thermophilic bacteria, can be used to ligate the reporter oligonucleotide to the extendible fragment. Thermophilic DNA ligases can be used in repeated cyclic ligations of the extendible fragments to reporter oligonucleotides.

[0080] As discussed above, the method according to the invention involves the production of an upstream fragment which has a 3' hydroxyl group. This is an essential requirement for extension of the molecule by addition of nucleotides by a polymerase and also for extension by ligation of a reporter oligonucleotide by a ligase. In addition to a 3' hydroxyl group on the upstream fragment, the reporter oligonucleotide is required to have a 5' phosphate group for ligation to occur. Ligation, in addition, allows detection of the downstream fragment if desired. The downstream fragment is the remainder of the DNA strand from which the upstream fragment was cleaved. Here, the reporter oligonucleotide is required to have a 3' hydroxyl terminus and the downstream fragment to have a 5' terminal phosphate group.

[0081] Since glycosylase mediated cleavage may generate several different upstream fragments in an individual reaction due to the presence of normal and mutant alleles in the target nucleic acid or due to the random introduction of modified bases, the extension by ligation reaction may contain several different upstream fragments in addition to several different reporter oligonucleotides and template nucleic acids. In addition several different nucleic acids may be characterised simultaneously since an individual reporter oligonucleotide and/or template nucleic acid may be used for the characterisation of each individual nucleic acid under investigation. The extended upstream fragment can be detected by any of several means including size analysis, hybridisation and amplification. In addition, the DNA molecule resulting from ligation can be further amplified in a polymerase chain reaction.

[0082] The reporter oligonucleotide, upstream fragment or template nucleic acid may be labelled. For example a fluorescent or radioactive label may be used in addition to a biotin or digoxigenin label. A useful label on the reporter oligonucleotide is a 5' terminal radioactive phosphate, i.e. ^{32}P or ^{33}P. This radioactive phosphate serves as a label with which to detect the DNA and also as the necessary 5' phosphate on the reporter molecule. A biotin label on the reporter oligonucleotide, upstream fragment or template nucleic acid will serve to immobilise the extended upstream fragment directly, or *via* hybridisation, to a solid support. Immobilisation, in combination with multiple different ligation extensions

IDT Exhibit 1020.9716

will serve to produce a very efficient and high throughput system for characterisation of DNA molecules.

[0083] The invention, using extension by ligation, may also be used to identify unknown sequence changes in nucleic acid by its ability to identify a mutant upstream fragment in a mixture of normal upstream fragments. A partially degenerate reporter oligonucleotide may be used in the ligation reaction. Selective ligation of an upstream fragment arising from cleavage of a DNA molecule at a mutation site can also be achieved by using a degenerate reporter oligonucleotide where the 5' terminus is complementary to the normal allele. In addition the invention can be used to investigate all CpG dinucleotide sites within a DNA fragment by using a fully complementary reporter oligonucleotide designed to anneal at each CpG site. The length of any resulting ligation product will indicate the CpG site which has been mutated.

[0084] Suitably, any extended fragments resulting from step iv) are detected by hybridisation.

[0085] The method according to the invention offers significant advantages over existing methods in that

a) it allows detection of the similarities or the differences, or the similarities and differences between nucleic acid samples. In particular, it allows the detection of same at large numbers of multiple different loci in a nucleic acid. While it is possible to use other methods to achieve this end point, this method offers the advantage of a single process which can readily be scaled up allowing rapid and easy characterisation of nucleic acid molecules.

b) It will be appreciated that amplification of nucleic acids is a common method for characterising and detecting nucleic acids. Amplification is dependent on the use of primers which are extended in the amplification process. In addition, the method according to the invention allows the generation of high specificity primers for amplification of nucleic acids without the necessity of prior knowledge of any of the sequence of the nucleic acid. Thus, the present method has a high utility for the characterisation of nucleic acids through analysis of amplification products generated therefrom. The amplification approaches described for characterising nucleic acids according to the invention allow characterisation of nucleic acids in a way that was not possible prior to the present invention.

c) in the field, there is a need for simplified methods for the detection of specific mutations at candidate loci. The method according to the invention offers such a simplified method for the detection of such mutations. In particular, the upstream fragment generated from sequence variation loci in nucleic acids allow the analysis of such loci using many different analytical approaches and allow accurate and simplified detection of sequence variations at such loci.

d) the method according to the invention offers a more reproducible method for characterising nucleic acids by random amplification and is less susceptible to error.

[0086] The method according to the invention can also be used to analyse the CpG content of DNA by detecting C to T transitions in DNA.

[0087] Upon analysis of mutation data and mutation spectra that have been generated over many years of mutation research, a definite hotspot for mutations in all organisms which contain 5-methylcytosine in their genomic DNA e.g. humans, has been identified, namely, mutations at CpG dinucleotides. CpG dinucleotides are a site for cytosine methylation in human cells and have been implicated in many structural and regulatory roles in genome organisation and gene expression, respectively. Cytosine in DNA is normally susceptible to a low but measurable level of deamination to form uracil, an event which is mutagenic if not repaired. However, upon methylation at the 5' site, the cytosine ring now is even more susceptible to spontaneous deamination. Therefore, the 5-methylcytosine residue is deaminated to become thymine. Following deamination, the CpG dinucleotide becomes a TpG dinucleotide. Therefore, since 5-methylcytosine occurs only at a CpG, and the major cause of mutation at this site is due to deamination of the 5-methylcytosine, mutated CpG sequences appear in DNA as TpG dinucleotide sequences which represents a classic C to T transition mutation. Since CpG has been demonstrated to be such a hotspot for mutation in human genetic studies, a rapid screen for enhanced detection of mutations at CpG dinucleotides within the test fragment of DNA is highly advantageous.

[0088] To enhance the detection of mutations at CpG sequences, one of the following procedures can be carried out subsequent to amplification of the target DNA followed by cleavage to generate the upstream fragments.

[0089] Step iv) is carried out in the presence of a polymerase, dideoxyTTP(ddTTP), labelled dideoxy CTP(ddCTP) and wild type DNA as template, the upstream fragments are extended by incorporation of dideoxynucleotides, but only sites where C has mutated to T, including sites where CpG has mutated to TpG will become labelled following the incorporation of the labelled ddCTP. Since the ddCTP is a chain terminator nucleotide, the DNA will not be extended beyond this point. Therefore the DNA will be extended at mutated CpG sequences and will also be labelled and detectable.

[0090] Step iv) is carried out in the presence of a polymerase, wild type DNA as template, dTTP and labelled dideoxyGTP (ddGTP), the upstream fragments are extended by incorporation of dTMP, but only sites where the T is followed

IDT Exhibit 1020.9717

by G, i.e. TpG sequences will become labelled following the incorporation of the labelled ddGTP. Since the ddGTP is a chain terminator nucleotide, the DNA will not be extended beyond this point. Therefore the DNA is now extended at all TpG sequences and is also labelled and detectable. Since TpG dinucleotides occur naturally in DNA as well as arising from CpG deamination, wild type DNA will show a characteristic banding pattern (normally only a few bands will be seen). The label can be e.g. radioactive ^{33}P, fluorescent, or biotin labelled in a manner known *per se.*

[0091] Step iv) is carried out in the presence of polymerase, dUTP, dGTP, dATP, labelled dCTP and wild type DNA as template, the upstream fragments are extended by incorporation of deoxynucleotides and become labelled. The extended upstream fragments are subsequently cleaved by uracil DNA glycosylase and abasic site cleavage agent. Only that upstream fragment which was generated by cleavage at the mutation site will remain extended and labelled. This procedure will detect all mutations generating a T incorporation site including C to T mutations at CpG sites. This procedure has the advantage that step iv) utilises the incorporation of deoxynucleotides which are incorporated more efficiently than dideoxynucleotides by all polymerases.

[0092] In addition, in heterozygous samples, the mutated CpG will give rise to T/G mismatches (or U/G after introduction of dUTP as modified nucleotide). The thymine DNA glycosylase can be used to specifically cleave at these T/G or U/G mismatches at a mutated CpG sequence.

Brief Description of Drawings

[0093]

Fig 1 is a diagram of an AP site generated during one step of the method according to the invention and of a number of ways in which the DNA at the abasic site can be cleaved so as to generate an upstream DNA fragment that can be extended;

Fig 2 is a schematic representation of the method according to the invention as described in Example 1;

Fig 3 is a schematic representation of the method according to the invention as described in Example 2 where a labelled ddTTP was used in a linear amplification reaction following generation of an upstream extendible fragment;

Fig 4 is a schematic representation of the method according to the invention as described in Example 2 where a labelled ddCTP was used in a linear amplification reaction following generation of an upstream extendible fragment;

Fig. 5 is a schematic representation of the method according to the invention as described in Example 3;

Fig. 6A-6D is a schematic representation of the extension products obtained in Example 3 following electrophoresis and autoradiography and analysis of the autoradiographs; and

Fig. 7A-7D is a schematic representation of the ligation reactions carried out in Example 4 on the upstream fragments and the products thereby obtained.

[0094] Fig. 1 depicts a single strand of DNA in 5' to 3' orientation. The position of two bases is shown, i.e. base 1 and base 3. Base 2 has been removed and an AP site exists at this position. The vertical lines denote the ribose ring which is attached to the base. The diagonal lines with 'P' in a circle refer to the phosphodiester linkages linking each ribose. The DNA can be cut at the 5' side or 3' side of an AP site as shown by arrows.

[0095] If the DNA is cleaved at the 3' side of the AP site, i.e. position Z, then the upstream fragment (containing base 1) has a deoxyribose phosphate moiety at its 3' terminus as shown in box C. Cleavage of the DNA at the 3' side of the AP site can be achieved by treatment with a 3' AP endonuclease / lyase or heat treatment as hereinabove described.

[0096] The AP site can also be cut in two different ways on the 5' side of the AP site, i.e. at positions X and Y. Cleavage at position X results in a OH group at the 3' terminus on the upstream fragment, as shown in box A. Cleavage at position X can be achieved by treatment with a 5' AP endonuclease as hereinabove described. Cleavage at position Y results in a phosphate group at the 3' terminus of the upstream fragment, as shown in box B. Cleavage at position Y can be achieved by treatment with heat and alkali as hereinabove described.

[0097] The invention will be further illustrated by the accompanying Examples.

IDT Exhibit 1020.9718

Modes for Carrying Out the Invention

Example 1

[0098] The method according to the invention was used to demonstrate the production of an extended DNA fragment by extension from the 3'OH group of an upstream fragment, which had been generated by cleavage of uracil containing DNA at the site of incorporation of the modified nucleotide. Target nucleic acid was a region of the RYR1 gene (952 to 1044) amplified from human cDNA by using upper (952 to 972) and lower (1024 to 1044) primers to generate a double stranded DNA, 93bp in length. (Nucleotide numbers refer to the sequence of the RYR1 gene)

[0099] Fig. 2 is a schematic diagram of the target nucleic acid and the upper and lower primers (primers contain standard bases, G, A, T and C). Six pmoles of the lower primer was end labelled by incubation with 1 unit of T4 Polynucleotide Kinase (commercially available from New England Biolabs), 70mM Tris-HCl (pH7.6), 10mM $MgCl_2$, 5mM dithiothreitol and 1μCi $\gamma^{32}P$ ATP (3000Ci/mmol) for 30 min at 37°C. The target nucleic acid sample was amplified by PCR in a reaction mix containing target nucleic acid, 0.2mM dATP, dCTP, dGTP and dUTP, 6pmoles of ^{32}P labelled lower primer and non-labelled upper primer in a total volume of 19μl. The reaction mix was then overlaid with an equal volume of mineral oil and a hot start PCR was performed whereby the reaction mix was heated to 94°C for 5min prior to addition of 1 unit of Taq polymerase (available from Promega) (bringing the total volume to 20μl). 30 cycles of denaturation, annealing and extension were carried out in a thermocycler. The reaction mixture bearing the amplified target nucleic acid was then treated with exonuclease I (available from Amersham Life Sciences) to digest the primers not extended in the amplification step and shrimp alkaline phosphatase (SAP) (available from Boehringer Mannheim) to digest the dNTPs not incorporated during the amplification step. This was achieved by incubating 10μl of the PCR reaction with 0.5 units of exonuclease I and 1 unit of SAP at 37°C for 30 min. Exo I and SAP were subsequently heat inactivated by incubating the reaction at 80°C for 15 min.

[0100] Uracil DNA-glycosylase (available from New England Biolabs) (0.5 units) was then added and the incubation continued at 37°C for 30 min. Following treatment with uracil DNA-glycosylase, the abasic sites generated in the amplified product were cleaved to completion by adding NaOH to a final concentration of 0.05M and heating the mixture for 15min at 95°C. The digested DNA was then precipitated by adding 10% volume 3M sodium acetate and 2 volumes of ethanol. The pellet was resuspended in 5μl water. The digested DNA was then treated with 0.5 units of T4 polynucleotide kinase (PNK) which removes the phosphate group from the 3' terminii.

[0101] A linear amplification reaction was then carried out using the products of the above cleavage reaction, the fragment of interest being the labelled upstream extendible fragment, during which, the upstream extendible fragment is extended by a thermostable DNA polymerase in a cycling reaction in a total volume of 10 μl. The template for this reaction is amplified target nucleic acid (952 to 1044 of RYR1 gene) which is free from primers due to pretreatment with ExoI.

[0102] An equal volume of formamide loading dye (90% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol) was added to the sample which was then heated at 85°C for 5 min. The sample was then loaded onto a 20% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60W for size analysis of the extension products. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hrs at -70°C.

[0103] Analysis of the autoradiograph, where the lower primer was labelled, showed a product of 93 nucleotides in length. This product was not observed if the T4 polynucleotide kinase treatment or the linear amplification reaction was not included in the above procedure.

Example 2

[0104] The method according to the invention was used to detect the presence of a G to A mutation at position 1021 in the human RYR1 gene. cDNA from a normal individual and from an individual affected with Malignant Hyperthermia was amplified by using upper (952 to 972) and lower (1204 to 1224) primers to generate a 273 bp double stranded DNA fragment (952 to 1224)(As in the case of Example 1 nucleotide numbers refer to the sequence of the RYR1 gene).

[0105] Figs. 3 and 4 are schematic diagrams of the target nucleic acid and the upper and lower primers (primers contain standard bases, G, A, T and C). The target nucleic acid sample was amplified by PCR in a reaction mix containing target nucleic acid, 0.2mM dATP, dCTP, dGTP and 0.19mM dTTP arid 0.01mM dUTP, and 6pmoles of upper and lower primers in a total volume of 19μl. The reaction mix was then overlaid with an equal volume of mineral oil and a hot start PCR was performed whereby the reaction mix was heated to 94°C for 5min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 20μl). 30 cycles of denaturation, annealing and extension were carried out in a thermocycler. The reaction mixture bearing the amplified target nucleic acid was then treated with exonuclease I to digest the primers not extended in the amplification step and shrimp alkaline phosphatase (SAP) to digest the dNTPs not incorporated during the amplification step. This was achieved by incubating 10μl of the PCR reaction with

IDT Exhibit 1020.9719

0.5 units of exonuclease I and 1 unit of SAP at 37°C for 30 min. Exo I and SAP were subsequently heat inactivated by incubating the reaction at 80°C for 15 min.

[0106] Uracil DNA-glycosylase (0.5 units) was then added and the incubation continued at 37°C for 30 min. Following treatment with uracil DNA-glycosylase, the abasic sites generated in the amplified product were cleaved to completion by adding NaOH to a final concentration of 0.05M and heating the mixture for 15 min at 95°C. The digested DNA was then precipitated by adding 10% volume 3M sodium acetate and 2 volumes of ethanol. The pellet was resuspended in 5μl water. The digested DNA was then treated with 0.5 units of T4 Polynucleotide Kinase which removes the phosphate group from the 3' terminii.

[0107] A linear amplification reaction was then carried out using the products of the above cleavage reaction, i.e. the various extendible upstream fragments, during which, the upstream fragment is extended by a thermostable DNA polymerase (i.e. Thermosequenase (available from Amersham Life Sciences)) in a cycling reaction in a total volume of 10 μl. The template for this reaction is amplified fragment of normal cDNA (952 to 1224 of RYR1 gene) which is free from unlabelled primers due to pretreatment with ExoI. The extension reaction is carried out in the presence of 1mM of three of the dideoxy terminator nucleotides (ddNTP) and 0.02mM of a ^{33}P-labelled ddNTP.

[0108] An equal volume of formamide loading dye (90% formamide, 0.025% bromophenol blue, 0.025% xylene cylanol) was added to the sample which was then heated at 85°C for 5 min. The sample was then loaded onto a 6% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60W for size analysis of the extension products. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hrs at -70°C.

[0109] Analysis of the autoradiograph, where ddTTP was the labelled ddNTP (Fig. 3), showed a ladder of labelled fragments corresponding to the distance from the 5' end of the primers to the site of a dUMP incorporation (Only extension of the lower strand is shown). These sites corresponded to the wild type T pattern of the DNA. When mutant target nucleic acid was used to generate the extension primers, an additional band was observed in the T pattern of bands (203 nucleotides). Analysis of the autoradiograph, where ddCTP was labelled (Fig. 4) showed no labelled bands when amplified normal target nucleic acid was the source of upstream fragments, however, when amplified mutant target nucleic acid was used, analysis showed just one band. The size of the labelled fragment corresponds to the distance between the 5' end of the lower primer and the mutation site (i.e. 203 nucleotides) thereby demonstrating the presence of the G to A mutation at position 1021 of the RYR1 gene in that individual.

Example 3

[0110] The method according to the invention was used to detect the presence of a G to A mutation at position 6411 (codon 12) in the human Ki-ras gene. Genomic DNA from normal tissue and tumour tissue of an individual with colon cancer was amplified by using upper (6390 to 6409) and lower (6417 to 6443) primers to generate a 54 bp double stranded DNA fragment (6390 to 6443)(Nucleotide numbers refer to the genomic sequence of the Ki-ras gene including introns).

[0111] Fig. 5 is a schematic diagram of the target nucleic acid and the upper and lower primers (primers contain standard bases, G, A, T and C). The target nucleic acid sample was amplified by PCR in a reaction mix containing target nucleic acid, 0.2mM dATP, dCTP, dGTP and dUTP and 6pmoles of upper and lower primers in a total volume of 19μl. The reaction mix was then overlaid with an equal volume of mineral oil and a hot start PCR was performed whereby the reaction mix was heated to 94°C for 5min prior to addition of 1 unit of Taq polymerase (bringing the total volume to 20μl). 30 cycles of denaturation, annealing and extension were carried out in a thermocycler. The reaction mixture bearing the amplified target nucleic acid was then treated with Exonuclease I (Exo I) to digest the primers not extended in the amplification step and Shrimp Alkaline Phosphatase (SAP) to digest the dNTPs not incorporated during the amplification step. This was achieved by incubating 20μl of the PCR reaction with 0.5 units of Exo I and 1 unit of SAP at 37°C for 30 min. Exo I and SAP were subsequently heat inactivated by incubating the reaction at 80°C for 15 min. Uracil DNA-glycosylase (0.5 units) and Endonuclease IV (1 unit) were then added and the incubation continued at 37°C for 30 min to allow complete excision of all the uracils present in the amplified DNA and cleavage of the resulting abasic sites to completion. This cleavage resulted in the upstream fragments having a 3' hydroxyl group at their 3' terminus which can be extended by action of a DNA polymerase.

[0112] Extension reactions were then carried out in a reaction mixture (10μl) containing the products of the above cleavage reaction, i.e. the extendible upstream fragments (2μl of cleavage reaction, approx. 1 pmol of extendible fragments), and 100fmol of a synthetic template oligonucleotide, 0.2mM dATP, dTTP, and dGTP, 0.02mM dCTP, 1μCi α^{32}PdCTP, 6pmol reverse primer and 1 unit Taq DNA polymerase. The reaction was carried out for 40 cycles of denaturation, annealing and extension.

[0113] An equal volume of formamide loading dye (90% formamide, 0.025% Bromophenol blue, 0.025% Xylene cylanol) was added to the sample which was then heated at 85°C for 5 min. The sample was then loaded onto a denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60W for size analysis of

IDT Exhibit 1020.9720

the extension products. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 12 hrs at -20°C.

[0114] Analysis of DNA from normal tissue results in the generation of a 37 nucleotide extendible fragment following glycosylase mediated cleavage using the above mentioned upper and lower primers (Fig. 5). Similar analysis of DNA from tumour tissue results in the generation of a 32 nucleotide extendible fragment (Fig. 5). Analysis of the autoradiograph showed a 66 nucleotide band following analysis of DNA from normal tissue and using template oligonucleotide 1 (Fig. 6A). This band was not observed when template oligonucleotide 2 was used in the above analysis (Fig. 6B). Analysis of the autoradiograph showed a 66 nucleotide band following analysis of DNA from tumour tissue and using template oligonucleotide 2 (Fig. 6C). This band was not observed when template oligonucleotide 1 was used in the above analysis (Fig. 6D). Therefore the presence of a mutation at codon 12 of the Ki-ras gene was determined by the ability of the upstream fragment to be extended on a mutant template oligonucleotide whereas the upstream fragment was not extended on a normal template oligonucleotide and *vice versa* for the absence of the mutation.

Example 4

[0115] Example 3 was repeated to the stage where the abasic sites had been cleaved to completion. This cleavage resulted in the upstream fragments having a 3' hydroxyl group at their 3' terminus as before.

[0116] Reporter oligonucleotides 1 and 2 (Fig. 7)(complementary to nucleotides 6397 to 6406 (5' TCCAACTACC3' No. 1 (SEQ ID NO. 28)) and nucleotides 6402 to 6411 (5'CCAGCTCCAA3', No. 2 (SEQ ID NO. 30)) of the Ki-ras gene respectively) were 5' endlabelled using γ^{32}PATP and T4 polynucleotide kinase. This served to label the resulting ligated fragment and also provided a 5' terminus phosphate on the reporter oligonucleotide which is required for any possible ligation.

[0117] Ligation reactions (16°C for 60 min) were then carried out using the products of the above cleavage reaction, i.e. the extendible upstream fragments (5μl of cleavage reaction, approx. 2pmol of extendible fragments), 4pmol of labelled reporter oligonucleotide and 2pmol of a template oligonucleotide 1 or 2 (5'GGTAGTTGGAGCTGGTGGCG3' (SEQ ID No. 27) (nuc 6397 to 6416) or 5'TTGGAGCTGGTGGCGTA GGC3' (SEQ ID No. 31) (nuc 6402 to 6421) respectively) (Fig. 7) and 1 unit T4 DNA ligase, during which, the upstream fragment is extended in length by the ligation of a reporter oligonucleotide in a 20μl reaction.

[0118] An equal volume of formamide loading dye (90% formamide, 0.025% Bromophenol blue, 0.025% Xylene cylanol) was added to the sample which was then heated at 85°C for 5 min. The sample was then loaded onto a denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60W for size analysis of the extension products. Following electrophoresis, autoradiography was carried out by exposing the gel directly to X-ray photographic film for 3 hrs at -20°C.

[0119] Analysis of DNA from normal tissue results in the generation of a 37 nucleotide extendible fragment following glycosylase mediated cleavage using the above mentioned upper and lower primers (Fig. 5). Similar analysis of DNA from tumour tissue results in the generation of a 32 nucleotide extendible fragment (Fig. 5). Analysis of the autoradiograph showed a 47 nucleotide band following analysis of DNA from normal tissue and using reporter oligonucleotide 1 and template oligonucleotide 1 (Fig. 7A). This band was not observed when reporter oligonucleotide 2 was used in the above analysis (Fig. 7B). Analysis of the autoradiograph showed a 42 nucleotide band following analysis of DNA from tumour tissue and using reporter oligonucleotide 2 and template oligonucleotide 2 (Fig. 7C). This band was not observed when reporter oligonucleotide 1 was used in the above analysis (Fig. 7D). Therefore the presence of a mutation at codon 12 of the Ki-ras gene was determined by the presence of a 42 nucleotide band whereas the presence of the normal allele was determined by the presence of a 47 nucleotide band. Samples containing normal and tumour DNA yielded both 42 and 47 nucleotide bands.

[0120] The above analysis was also carried out where the lower primer was 5' endlabelled with ^{32}P. This required that the reporter oligonucleotides be phosphorylated using unlabelled ATP as phosphate donor. The results were similar to those observed above in that a 42n band demonstrated the presence of the mutant Ki-ras gene (codon 12) whereas a 47n band demonstrated the absence of the mutation at codon 12 of the Ki-ras gene. In addition the above analysis was also carried out using the initially amplified Ki-ras gene fragment from normal or mutant sample as template during the ligation reaction. Normal amplified product was used instead of template oligonucleotide 1, whereas mutant amplified product was used instead of template oligonucleotide 2. Again the same results were observed as described above.

[0121] As indicated above, the method according to the invention has numerous advantages over known methods, especially the method of WO 97/03210. After the glycosylase mediated cleavage in the case of WO 97/03210 (which can be carried out in a number of ways and which yields a number of different 3' termini), the resulting DNA fragments are not further processed and are analysed directly. In the present invention, the glycosylase mediated cleavage step is followed by a step that permits extension of the 3' termini generated by glycosylase mediated cleavage.

[0122] A major advantage of the present invention, which is not possible in the case of the prior art methods, is that

IDT Exhibit 1020.9721

the present invention allows detection of sequence differences such as mutations and polymorphisms between nucleic acid molecules without detecting sequence similarities. As indicated above, it is not possible with the method of WO 97/03210 to detect sequence differences between nucleic acid molecules without detecting sequence similarities. This is a limitation of the method of WO 97/03210 as multiple samples cannot be combined for simultaneous analysis. The present invention also allows for the analysis of multiple genes or gene segments simultaneous analysis. Furthermore, the present invention allows for the generation of specific primers for amplification of nucleic acids without the necessity of having prior knowledge of the sequence of the nucleic acid. This is not possible with the known methods. In addition, the present invention permits the generation of specific primers from nucleic acids in a unique way and such primers can be subsequently assayed by polymerase extension to determine the nature of the sequence at the 3' termini of said primers.

SEQUENCE LISTING

[0123]

(1) GENERAL INFORMATION :

(i) APPLICANT:

(A) NAME : FORBAIRT (trading as BioResearch Ireland)
(B) STREET: Galsnevin
(C) CITY : Dublin 9
(E) COUNTRY : Ireland
(F) POSTAL CODE (ZIP) : none

(A) NAME : UNIVERSITY COLLEGE CORK
(B) STREET: College Road
(C) CITY : Cork
(D) COUNTRY : Ireland
(F) POSTAL CODE (ZIP) : none

(A) NAME : McCARTHY, Thomas Valentine
(B) STREET: Vista Villa, Montenotte
(C) CITY : Cork
(E) COUNTRY : Ireland
(F) POSTAL CODE (ZIP) : none

(A) NAME : VAUGHAN, Patrick Martin
(B) STREET: 175 West Avenue, Parkgate, Frankfield
(C) CITY : Cork
(E) COUNTRY : Ireland
(F) POSTAL CODE (ZIP) : none

(ii) TITLE OF INVENTION : A method for the characterisation of nucleic acid molecules involving generation of extendible upstream DNA fragments resulting from the cleavage of nucleic acid at an abasic site

(iii) NUMBER OF SEQUENCES : 32

(iv) COMPUTER READABLE FORM :

(A) MEDIUM TYPE : Floppy disk
(B) COMPUTER : IBM PC compatible
(C) OPERATING SYSTEM : PC-DOS/MS-DOS
(D) SOFTWARE : PatentIn Realease #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO : 1

(i) SEQUENCE CHARACTERISTICS :

IDT Exhibit 1020.9722

(A) LENGTH : 93 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : cDNA

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(vi) ORIGINAL SOURCE :

(F) TISSUE TYPE : Skeletal muscle

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 1 :

```
TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGATGTGG AGGGCATGGG CCCCCCTGAG    60

ATCAAGTACG GGGAGTCACT GTGCTTCGTG CAG                                 93
```

(2) INFORMATION FOR SEQ ID NO : 2 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 93 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by PCR amplification"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 2 :

```
TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGAUGUGG AGGGCAUGGG CCCCCCUGAG    60

AUCAAGUACG GGGAGUCACU GUGCUUCGUG CAG                                 93
```

(2) INFORMATION FOR SEQ ID NO : 3 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 93 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

IDT Exhibit 1020.9723

(A) DESCRIPTION : /desc = "DNA generated by PCR amplification"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 3 :

CTGCACGAAG CACAGTGACT CCCCGUACUU GAUCUCAGGG GGGCCCAUGC CCUCCACAUC 60

CCGCUUGGGG GCCACAUCCA GCUUCUCCUU GGA 93

(2) INFORMATION FOR SEQ ID NO : 4 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 25 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage and has a 3' phosphate group"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 4 :

CTGCACGAAG CACAGTGACT CCCCG 25

(2) INFORMATION FOR SEQ ID NO : 5 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 25 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage and has a 3' hydroxyl group"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

IDT Exhibit 1020.9724

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 5 :

```
CTGCACGAAG CACAGTGACT CCCCG                                              25
```

(2) INFORMATION FOR SEQ ID NO : 6 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 93 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage followed by extension of upstream fragment"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 6 :

```
CTGCACGAAG CACAGTGACT CCCCGTACTT GATCTCAGGG GGGCCCATGC CCTCCACATC        60
CCGCTTGGGG GCCACATCCA GCTTCTCCTT GGA                                     93
```

(2) INFORMATION FOR SEQ ID NO : 7 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 273 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : cDNA

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(vi) ORIGINAL SOURCE :

(F) TISSUE TYPE : Skeletal muscle

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 7 :

IDT Exhibit 1020.9725

```
TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGATGTGG AGGGCATGGG CCCCCCTGAG      60

ATCAAGTACG GGGAGTCACT GTGCTTCGTG CAGCATGTGG CCTCAGGACT GTGGCTCACC     120

TATGCCGCTC CAGACCCCAA GGCCCTGCGG CTCGGCGTGC TCAAGAAGAA GGCCATGCTG     180

CACCAGGAGG GCCACATGGA CGACGCACTG TCGCTGACCC GCTGCCAGCA GGAGGAGTCC     240

CAGGCCGCCC GCATGATCCA CAGCACCAAT GGC                                  273
```

(2) INFORMATION FOR SEQ ID NO : 8 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 273 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : cDNA

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(vi) ORIGINAL SOURCE :

(F) TISSUE TYPE : Skeletal muscle

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 8 :

```
TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGATGTGG AGGGCATGGG CCCCCCTGAG      60

ATCAAGTACA GGGAGTCACT GTGCTTCGTG CAGCATGTGG CCTCAGGACT GTGGCTCACC     120

TATGCCGCTC CAGACCCCAA GGCCCTGCGG CTCGGCGTGC TCAAGAAGAA GGCCATGCTG     180

CACCAGGAGG GCCACATGGA CGACGCACTG TCGCTGACCC GCTGCCAGCA GGAGGAGTCC     240

CAGGCCGCCC GCATGATCCA CAGCACCAAT GGC                                  273
```

(2) INFORMATION FOR SEQ ID NO : 9 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 196 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage and upstream fragment extension, and has a 3' hydrogen atom"

IDT Exhibit 1020.9726

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 196
(D) OTHER INFORMATION :/mod_base= OTHER /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 9 :

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGT                                                     196
```

(2) INFORMATION FOR SEQ ID NO : 10 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 200 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 200
(D) OTHER INFORMATION : /mod_base= OTHER /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 10 :

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGTGACT                                                 200
```

(2) INFORMATION FOR SEQ ID NO : 11 :

(i) SEQUENCE CHARACTERISTICS :

IDT Exhibit 1020.9727

(A) LENGTH : 204 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 204
(D) OTHER INFORMATION : /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 11 :

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGTGACT CCCT                                            204
```

(2) INFORMATION FOR SEQ ID NO : 12 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 206 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage and followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 206
(D) OTHER INFORMATION : /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 12 :

IDT Exhibit 1020.9728

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGTGACT CCCCGT                                          206
```

(2) INFORMATION FOR SEQ ID NO : 13 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 209 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 209
(D) OTHER INFORMATION : /mod_base= OTHER /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 13 :

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGTGACT CCCCGTACT                                       209
```

(2) INFORMATION FOR SEQ ID NO : 14 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 204 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

IDT Exhibit 1020.9729

(iii) HYPOTHETICAL : NO

(iv) ANTI-SENSE : NO

(ix) FEATURE :

(A) NAME/KEY : modified_base
(B) LOCATION : 204
(D) OTHER INFORMATION : /mod_base= OTHER /note= "Dideoxy C"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 14 :

```
GCCATTGGTG CTGTGGATCA TGCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG      60

CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC     120

GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG     180

CTGCACGAAG CACAGTGACT CCCC                                            204
```

(2) INFORMATION FOR SEQ ID NO : 15 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 54 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : double
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : DNA (genomic)

(vi) ORIGINAL SOURCE :

(A) ORGANISM : Homo sapiens

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 15 :

```
AACTTGTGGT AGTTGGAGCT GGTGGCGTAG GCAAGAGTGC CTTGACGATA CAGC          54
```

(2) INFORMATION FOR SEQ ID NO : 16 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 54 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by PCR amplification of genomic DNA"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 16 :

IDT Exhibit 1020.9730

AACTTGTGGT AGTTGGAGCT GGUGGCGUAG GCAAGAGUGC CUUGACGAUA CAGC 54

(2) INFORMATION FOR SEQ ID NO : 17 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 54 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by PCR amplification of genomic DNA"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 17 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCUCC AACUACCACA AGUU 54

(2) INFORMATION FOR SEQ ID NO : 18 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 54 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by PCR amplification of genomic DNA"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 18 :

AACTTGTGGT AGTTGGAGCT GAUGGCGUAG GCAAGAGUGC CUUGACGAUA CAGC 54

(2) INFORMATION FOR SEQ ID NO : 19 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 54 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by PCR amplification"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 19 :

IDT Exhibit 1020.9731

GCTGTATCGT CAAGGCACTC TTGCCTACGC CAUCAGCUCC AACUACCACA AGUU 54

(2) INFORMATION FOR SEQ ID NO : 20 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 37 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by glycosylase mediated cleavage of PCR amplified DNA"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 20 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGC 37

(2) INFORMATION FOR SEQ ID NO : 21 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 32 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by glycosylase mediated cleavage of PCR amplified DNA"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 21 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CA 32

(2) INFORMATION FOR SEQ ID NO : 22 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 66 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 22 :

IDT Exhibit 1020.9732

GCTGTAAACG ACGGCCAGTT TCATGCAGGG CTGGAGTCGT AGGCAAGAGT GCCTTGACGA 60

TACAGC 66

(2) INFORMATION FOR SEQ ID NO : 23 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 24 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 23 :

GCTGTAAACG ACGGCCAGTT TCAT 24

(2) INFORMATION FOR SEQ ID NO : 24 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 66 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by primer extension"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 24 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCCCT GCATGAAACT GGCCGTCGTT 60

TACAGC 66

(2) INFORMATION FOR SEQ ID NO : 25 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 66 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 25 :

IDT Exhibit 1020.9733

GCTGTAAACG ACGGCCAGTT TCATGCAGGA TCCATGGCGT AGGCAAGAGT GCCTTGACGA 60

TACAGC 66

(2) INFORMATION FOR SEQ ID NO : 26 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 66 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by primer extension"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 26 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CATGGATCCT GCATGAAACT GGCCGTCGTT 60

TACAGC 66

(2) INFORMATION FOR SEQ ID NO : 27 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 20 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 27 :

GGTAGTTGGA GCTGGTGGCG 20

(2) INFORMATION FOR SEQ ID NO : 28 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 10 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

IDT Exhibit 1020.9734

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 28 :

TCCAACTACC 10

(2) INFORMATION FOR SEQ ID NO : 29 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 47 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by ligation of two DNA molecules"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 29 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCTCC AACTACC 47

(2) INFORMATION FOR SEQ ID NO : 30 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 10 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 30 :

CCAGCTCCAA 10

(2) INFORMATION FOR SEQ ID NO : 31 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 20 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 31 :

IDT Exhibit 1020.9735

TTGGAGCTGG TGGCGTAGGC 20

(2) INFORMATION FOR SEQ ID NO : 32 :

(i) SEQUENCE CHARACTERISTICS :

(A) LENGTH : 42 base pairs
(B) TYPE : nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) MOLECULE TYPE : other nucleic acid

(A) DESCRIPTION : /desc = "Generated by ligation of two DNA molecules"

(xi) SEQUENCE DESCRIPTION : SEQ ID NO : 32 :

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCTCC AA 42

Claims

1. A method for characterising nucleic acid molecules, which comprises the steps of:

   i) introducing a modified base which is a substrate for a DNA glycosylase into a DNA molecule;

   ii) excising the modified base by means of said DNA glycosylase so as to generate an abasic site;

   iii) cleaving the DNA at the abasic site so as to generate an upstream DNA fragment that can be extended; and

   iv) incubating the extendible upstream fragment in the presence of an enzyme allowing for extension thereof and a template nucleic acid and analysing the resultant fragment(s).

2. A method according to Claim 1, wherein the upstream fragment is generated by cleaving the DNA at the 5' side of the abasic site, such that the 3' terminus of the upstream fragment bears a hydroxyl group.

3. A method according to Claim 2, wherein the cleavage is achieved with a 5' AP endonuclease.

4. A method according to Claim 1, wherein the upstream fragment is generated by cleaving at the 5' side of the abasic site so as to leave a phosphate group at the 3' terminus of the upstream fragment and removing the phosphate group so that the upstream fragment bears a hydroxyl group at the 3' terminus.

5. A method according to Claim 1, wherein the upstream fragment is generated by cleaving at the 3' side of the abasic site so as to generate a deoxyribose phosphate group at the 3' terminus of the upstream fragment and subsequently removing the deoxyribose group to leave a hydroxyl group at the 3' terminus.

6. A method according to any preceding claim, wherein 5' deoxyribose moieties downstream of the 3' terminus of the upstream fragment are removed so that the upstream fragment can be extended on the template.

7. A method according to Claim 6, wherein the 5' deoxyribose moieties are removed by a 5' deoxyribophosphodiesterase.

8. A method according to any preceding claim, wherein the modified base is introduced by enzymatic amplification of the DNA.

IDT Exhibit 1020.9736

9. A method according to Claim 8, wherein the amplified strands are separated for a separate analysis of the respective strands.

10. A method according to Claim 8 or 9, wherein a primer or one or more nucleotide(s) involved in the enzymatic amplification is labelled.

11. A method according to any preceding claim, wherein the enzyme is a polymerase.

12. A method according to Claim 11, wherein the extendible upstream fragment is incubated with the polymerase in the presence of one or more nucleotide(s).

13. A method according to Claim 12, wherein one or more of the nucloeotide(s) of step iv) is a dideoxy nucleotide.

14. A method according to Claim 12 or 13, wherein one or more of the nucleotide(s) of step iv) is labelled.

15. A method according to any one of Claims 11-14, wherein the extension of step iv) is achieved by means of an amplification reaction using said extendible DNA fragment.

16. A method according to any one of Claims 11-15, wherein the extension of step iv) is achieved by means of an amplification reaction including a primer in addition to using said extendible DNA fragment.

17. A method according to any one of Claims 1-10, wherein the enzyme is a ligase.

18. A method according to Claim 17, wherein the extendible upstream fragment is incubated with the ligase in the presence of a reporter oligonucleotide.

19. A method according to Claim 18, wherein the reporter oligonucleotide is partially degenerate.

20. A method according to any preceding claim, wherein any extended fragments resulting from step iv) are detected by hybridisation.

21. A method according to any preceding claim, which is used to detect a known or unknown mutation.

22. A method according to any one of Claims 1-20, wherein the method is used to analyse the CpG content of DNA by detecting C to T transitions in DNA.

Patentansprüche

1. Verfahren zur Charakterisierung von Nukleinsäuremolekülen, das die folgenden Schritte umfasst:

   i) Einführen einer modifizierten Base, die ein Substrat für eine DNAGlycosylase ist, in ein DNA-Molekül;

   ii) Herausschneiden der modifizierten Base mittels der DNAGlycosylase, um eine abasische Stelle zu erzeugen;

   iii) Spalten der DNA an der abasischen Stelle, um ein stromaufwärts liegendes DNA-Fragment, das verlängert werden kann, zu erzeugen; und

   iv) Inkubieren des verlängerbaren stromaufwärts liegenden Fragments in Gegenwart eines Enzyms, das eine Verlängerung desselben erlaubt, und einer Matrizen-Nukleinsäure und Analysieren des resultierenden Fragments oder der resultierenden Fragmente.

2. Verfahren nach Anspruch 1, wobei das stromaufwärts liegende Fragment durch Spalten der DNA an der 5'-Seite der abasischen Stelle erzeugt wird, so dass der 3'-Terminus des stromaufwärts liegenden Fragments eine Hydroxylgruppe trägt.

3. Verfahren nach Anspruch 2, wobei die Spaltung mit einer 5'-AP-Endonuklease erzielt wird.

IDT Exhibit 1020.9737

4. Verfahren nach Anspruch 1, wobei das stromaufwärts liegende Fragment erzeugt wird, indem an der 5'-Seite der abasischen Stelle gespalten wird, so dass eine Phosphatgruppe an dem 3'-Terminus des stromaufwärts liegenden Fragments zurückgelassen wird, und die Phosphatgruppe entfernt wird, so dass das stromaufwärts liegende Fragment eine Hydroxylgruppe an dem 3'-Terminus trägt.

5. Verfahren nach Anspruch 1, wobei das stromaufwärts liegende Fragment erzeugt wird, indem an der 3'-Seite der abasischen Stelle gespalten wird, so dass eine Desoxyribosephosphatgruppe an dem 3'-Terminus des stromaufwärts liegenden Fragments erzeugt wird, und nachfolgend die Desoxyribosegruppe entfernt wird, wodurch eine Hydroxylgruppe an dem 3'-Terminus zurückgelassen wird.

6. Verfahren nach einem der vorangegangenen Ansprüche, wobei 5'-Desoxyribosegruppierungen stromabwärts des 3'-Terminus des stromaufwärts liegenden Fragments entfernt werden, so dass das stromaufwärts liegende Fragment an der Matrize verlängert werden kann.

7. Verfahren nach Anspruch 6, wobei die 5'-Desoxyribosegruppierungen durch eine 5'-Desoxyribophosphodiesterase entfernt werden.

8. Verfahren nach einem der vorangegangenen Ansprüche, wobei die modifizierte Base durch eine enzymatische Amplifizierung der DNA eingeführt wird.

9. Verfahren nach Anspruch 8, wobei die amplifizierten Stränge für eine getrennte Analyse der jeweiligen Stränge getrennt werden.

10. Verfahren nach Anspruch 8 oder 9, wobei ein Primer oder ein oder mehrere Nukleotid(e), die an der enzymatischen Amplifizierung beteiligt sind, markiert ist.

11. Verfahren nach einem der vorangegangenen Ansprüche, wobei das Enzym eine Polymerase ist.

12. Verfahren nach Anspruch 11, wobei das verlängerbare stromaufwärts liegende Fragment mit der Polymerase in Gegenwart von einem oder mehreren Nukleotid(en) inkubiert wird.

13. Verfahren nach Anspruch 12, wobei ein oder mehrere der Nukleotid(e) von Schritt iv) ein Didesoxynukleotid ist.

14. Verfahren nach Anspruch 12 oder 13, wobei ein oder mehrere der Nukleotid(e) von Schritt iv) markiert ist.

15. Verfahren nach einem der Ansprüche 11-14, wobei die Verlängerung von Schritt iv) mittels einer Amplifizierungsreaktion unter Verwendung des verlängerbaren DNA-Fragments erzielt wird.

16. Verfahren nach einem der Ansprüche 11-15, wobei die Verlängerung von Schritt iv) mittels einer Amplifizierungsreaktion erzielt wird, die einen Primer zusätzlich zu einer Verwendung des verlängerbaren DNA-Fragments einbezieht.

17. Verfahren nach einem der Ansprüche 1-10, wobei das Enzym eine Ligase ist.

18. Verfahren nach Anspruch 17, wobei das verlängerbare stromaufwärts liegende Fragment mit der Ligase in Gegenwart eines Reporter-Oligonukleotids inkubiert wird.

19. Verfahren nach Anspruch 18, wobei das Reporter-Oligonukleotid partiell degeneriert ist.

20. Verfahren nach einem der vorangegangenen Ansprüche, wobei jegliche verlängerte Fragmente, die aus Schritt iv) resultieren, durch Hybridisierung nachgewiesen werden.

21. Verfahren nach einem der vorangegangenen Ansprüche, das verwendet wird, um eine bekannte oder unbekannte Mutation nachzuweisen.

22. Verfahren nach einem der Ansprüche 1-20, wobei das Verfahren verwendet wird, um den CpG-Gehalt von DNA zu analysieren, indem C-zu-T-Transitionen in DNA nachgewiesen werden.

IDT Exhibit 1020.9738

Revendications

1. Méthode pour caractériser des molécules d'acide nucléique, qui comprend les étapes consistant :

   i) à introduire dans une molécule d'ADN une base modifiée qui est un substrat pour une ADN glycosylase ;

   ii) à exciser la base modifiée au moyen de ladite ADN glycosylase de façon à générer un site abasique ;

   iii) à cliver l'ADN au niveau du site abasique de façon à générer en amont un fragment d'ADN qui peut être allongé ; et

   iv) à incuber le fragment en amont qui peut être allongé en présence d'une enzyme permettant son allongement et d'un acide nucléique matrice et à analyser le ou les fragment(s) résultant(s).

2. Méthode selon la revendication 1, dans laquelle le fragment en amont est généré par clivage de l'ADN du côté 5' du site abasique, de telle façon que la terminaison 3' du fragment en amont porte un groupe hydroxyle.

3. Méthode selon la revendication 2, dans laquelle le clivage est effectué à l'aide d'une 5'AP endonucléase.

4. Méthode selon la revendication 1, dans laquelle le fragment en amont est généré par clivage du côté 5' du site abasique de façon à laisser un groupe phosphate à la terminaison 3' du fragment en amont et à éliminer le groupe phosphate de façon à ce que le fragment en amont porte un groupe hydroxyle à sa terminaison 3'.

5. Méthode selon la revendication 1, dans laquelle le fragment en amont est généré par clivage du côté 3' du site abasique de façon à générer un groupe désoxyribose phosphate à la terminaison 3' du fragment en amont et à éliminer ensuite le groupe désoxyribose pour laisser un groupe hydroxyle à la terminaison 3'.

6. Méthode selon l'une quelconque des revendications précédentes, dans laquelle des fragments 5' désoxyribose en aval de la terminaison 3' du fragment en amont sont éliminés de telle façon que le fragment en amont puisse subir un allongement sur la matrice.

7. Méthode selon la revendication 6, dans laquelle les fragments 5' désoxyribose sont éliminés par une 5' désoxyribophosphodiestérase.

8. Méthode selon l'une quelconque des revendications précédentes, dans laquelle la base modifiée est introduite par amplification enzymatique de l'ADN.

9. Méthode selon la revendication 8, dans laquelle les brins amplifiés sont séparés pour une analyse distincte des brins respectifs.

10. Méthode selon la revendication 8 ou 9, dans laquelle une amorce ou un ou plusieurs nucléotide(s) impliqués dans l'amplification enzymatique sont marqués.

11. Méthode selon l'une quelconque des revendications précédentes, dans laquelle l'enzyme est une polymérase.

12. Méthode selon la revendication 11, dans laquelle le fragment en amont qui peut être allongé est incubé avec la polymérase en présence d'un ou plusieurs nucléotide(s).

13. Méthode selon la revendication 12, dans laquelle un ou plusieurs des nucléotides de l'étape iv) est un didéoxynucléotide.

14. Méthode selon la revendication 12 ou 13, dans laquelle un ou plusieurs des nucléotides de l'étape iv) est marqué.

15. Méthode selon l'une quelconque des revendications 11 à 14, dans laquelle l'allongement de l'étape iv) est réalisé au moyen d'une réaction d'amplification utilisant ledit fragment d'ADN qui peut être allongé.

16. Méthode selon l'une quelconque des revendications 11 à 15, dans laquelle l'allongement de l'étape iv) est réalisé au moyen d'une réaction d'amplification comprenant une amorce en plus de l'utilisation dudit fragment d'ADN qui

IDT Exhibit 1020.9739

peut être allongé.

17. Méthode selon l'une quelconque des revendications 1 à 10, dans laquelle l'enzyme est une ligase.

18. Méthode selon la revendication 17, dans laquelle le fragment en amont qui peut être allongé est incubé avec la ligase en présence d'un oligonucléotide reporter.

19. Méthode selon la revendication 18, dans laquelle l'oligonucléotide reporter est partiellement dégénéré.

20. Méthode selon l'une quelconque des revendications précédentes, dans laquelle tout fragment allongé résultant de l'étape iv) est détecté par hybridation.

21. Méthode selon l'une quelconque des revendications précédentes, qui est utilisée pour détecter une mutation connue ou inconnue.

22. Méthode selon l'une quelconque des revendications 1 à 20, dans laquelle la méthode est utilisée pour analyser le contenu CpG de l'ADN par détection des transitions de C vers T dans l'ADN.

IDT Exhibit 1020.9740

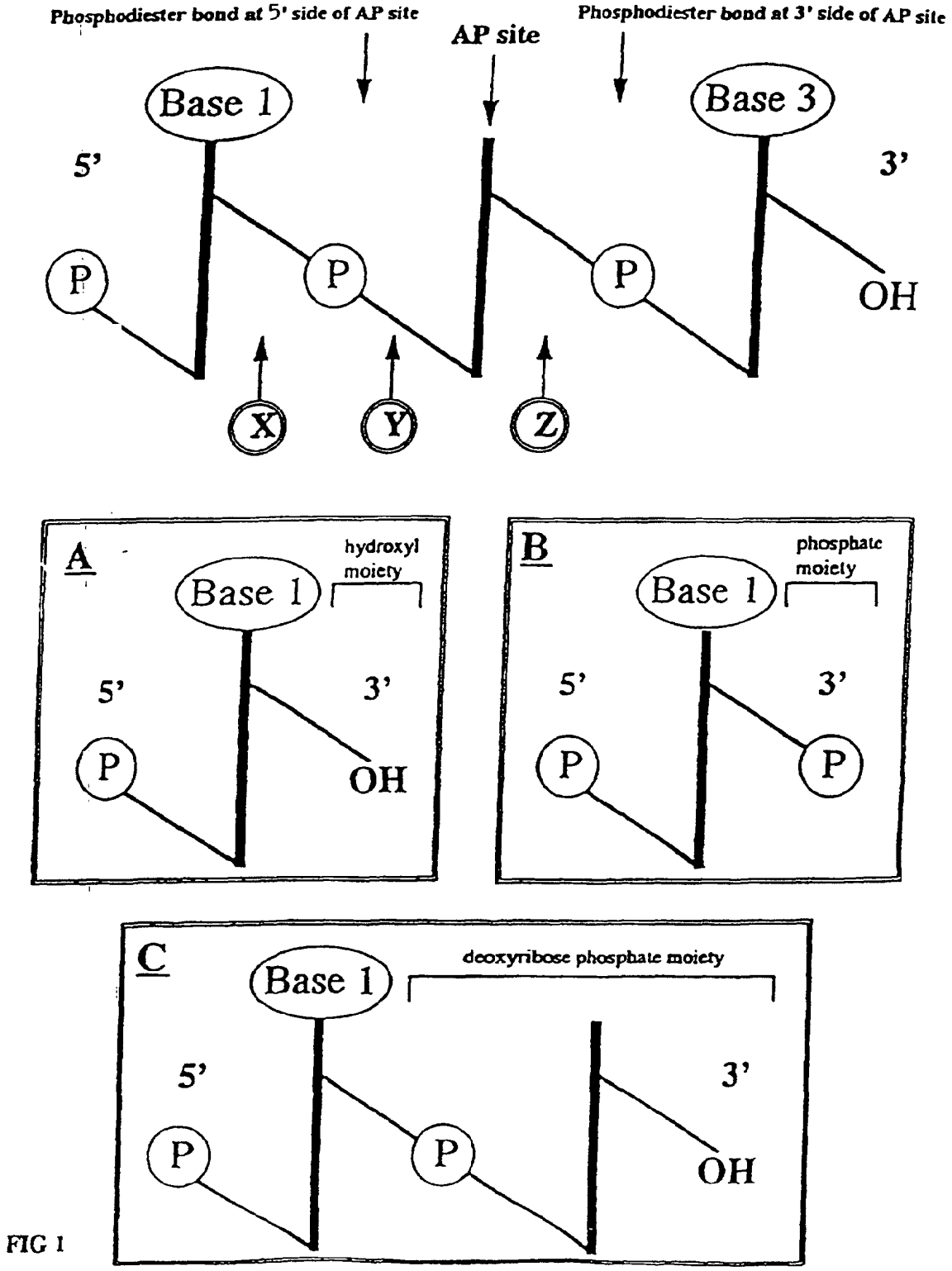


FIG 1

IDT Exhibit 1020.9741

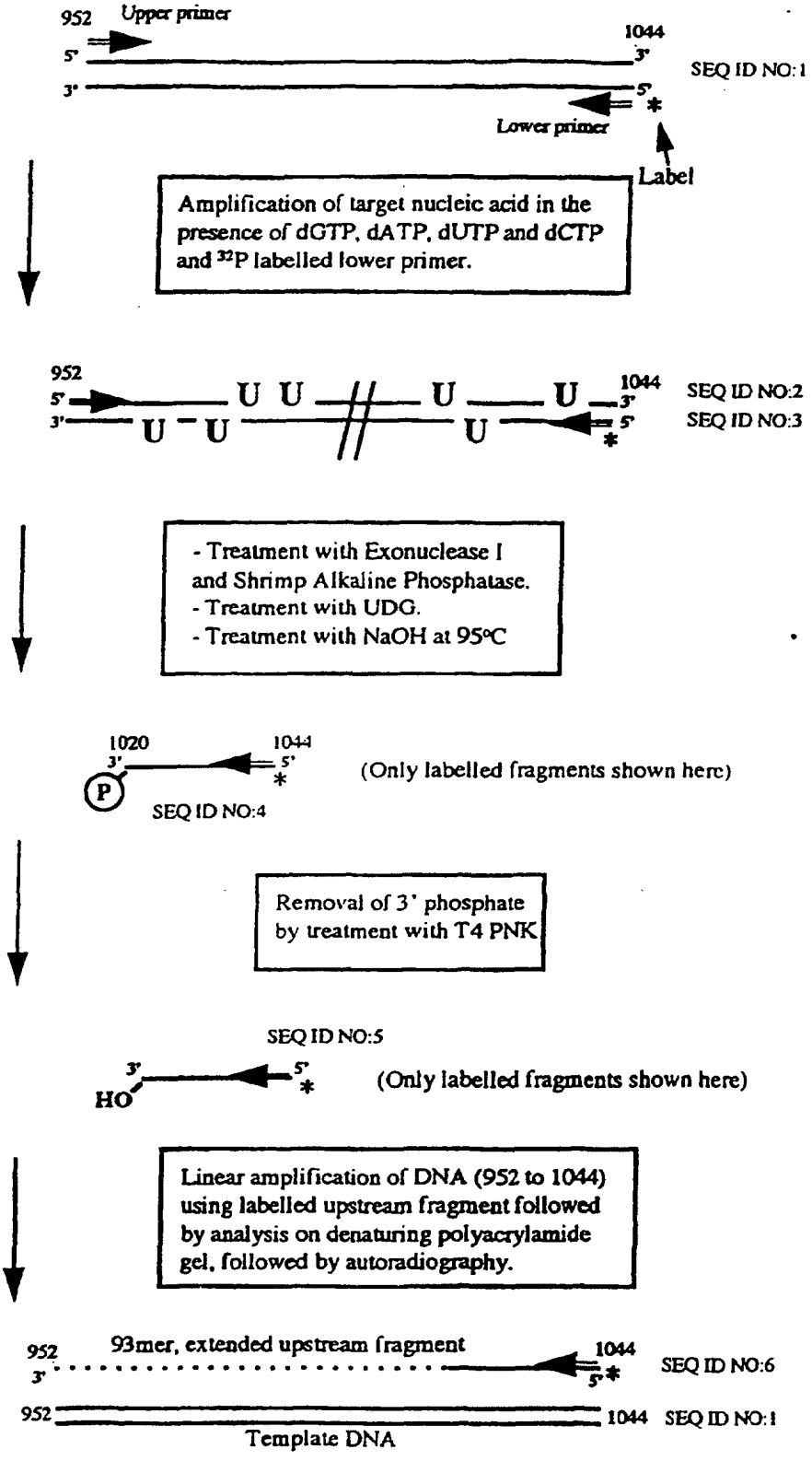
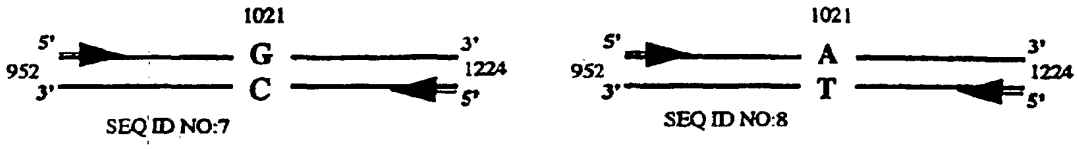


FIG 2

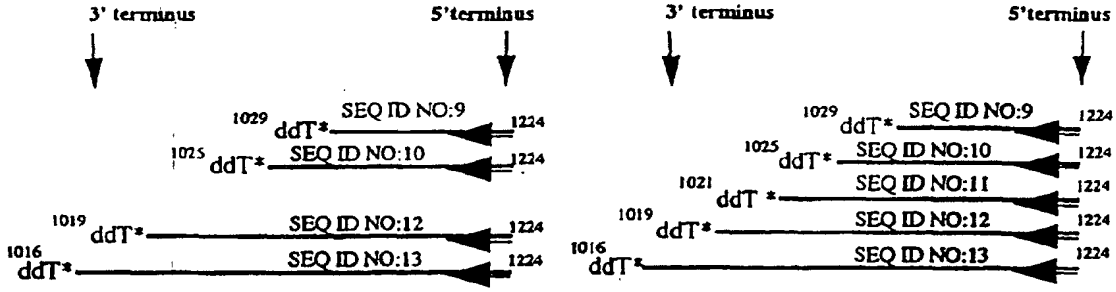
IDT Exhibit 1020.9742

- Amplification of normal and mutanttarget nucleic acid in the presence of dGTP, dATP, dCTP and1/20 ratio of dUTP to dTTP.



- Treatment with Exonuclease I and Shrimp Alkaline Phosphatase.
- Treatment with UDG.
- Treatment with NaOH at 95°C
- DNA is precipitated.
- Treatment with T4 PNK

Extension of the upstream fragments generated above in the presence of ^{33}P-labelled ddTTP* and unlabelled ddGTP, ddATP and ddCTP.



Only some fragments corresponding to cleavage at U incorporation sites surrounding the mutation site are shown here.

Detection of extended labelled fragments by PAGE and autoradiography

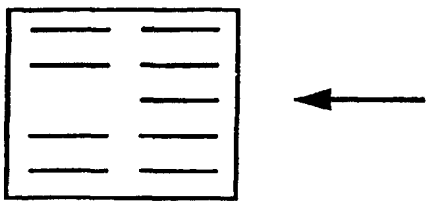


FIG 3

IDT Exhibit 1020.9743

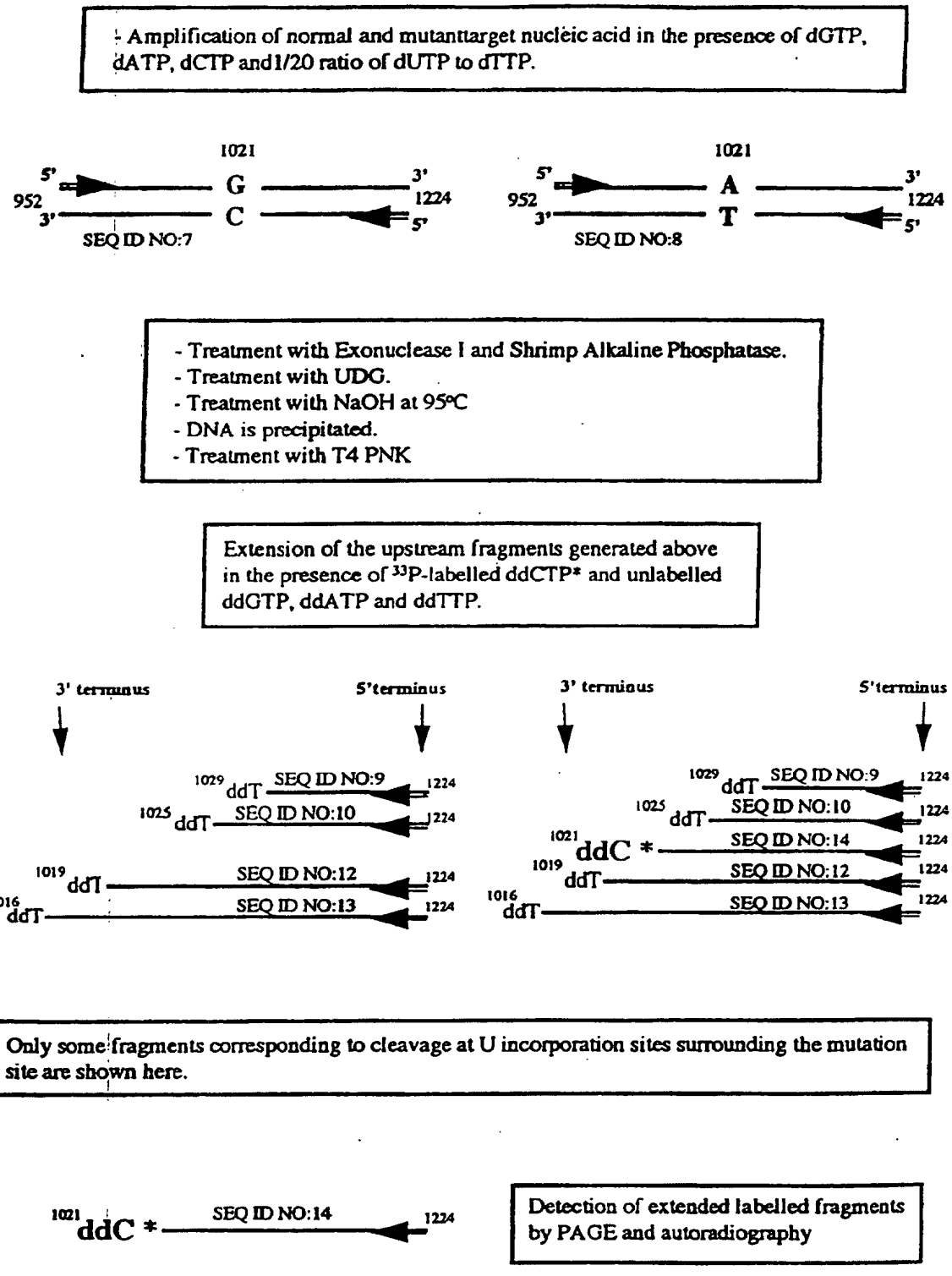
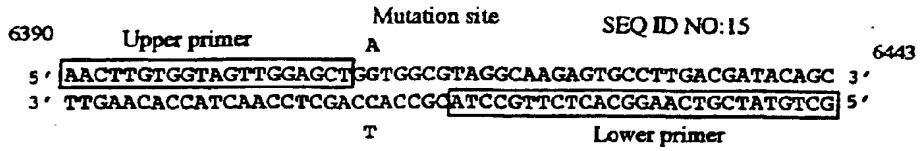
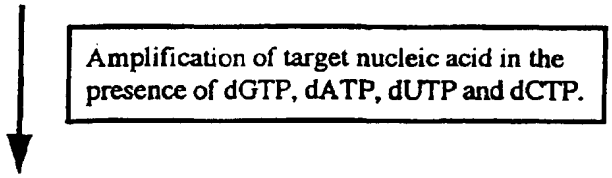


FIG 4

IDT Exhibit 1020.9744





SEQ ID NO:16

Amplified normal allele

5' AACTTGTGGTAGTTGGAGCTGGUGGCGUAGGCAAGAGUGCCUUGACGAUACAGC 3'
3' UUGAACACCAUCAACCUCGACCACCGCATCCGTTCTCACGGAACTGCTATGTCG 5'

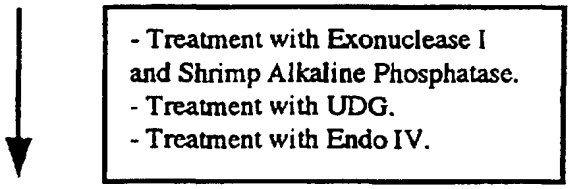
SEQ ID NO:17

SEQ ID NO:18

Amplified mutant allele

5' AACTTGTGGTAGTTGGAGCTGAUGGCGUAGGCAAGAGUGCCUUGACGAUACAGC 3'
3' UUGAACACCAUCAACCUCGACUACCGCATCCGTTCTCACGGAACTGCTATGTCG 5'

SEQ ID NO:19



SEQ ID NO:20 Normal upstream fragment

3' CGACCACCGCATCCGTTCTCACGGAACTGCTATGTCG 5'

SEQ ID NO:21 Mutant upstream fragment

3' ACCGCATCCGTTCTCACGGAACTGCTATGTCG 5'

FIG 5

IDT Exhibit 1020.9745

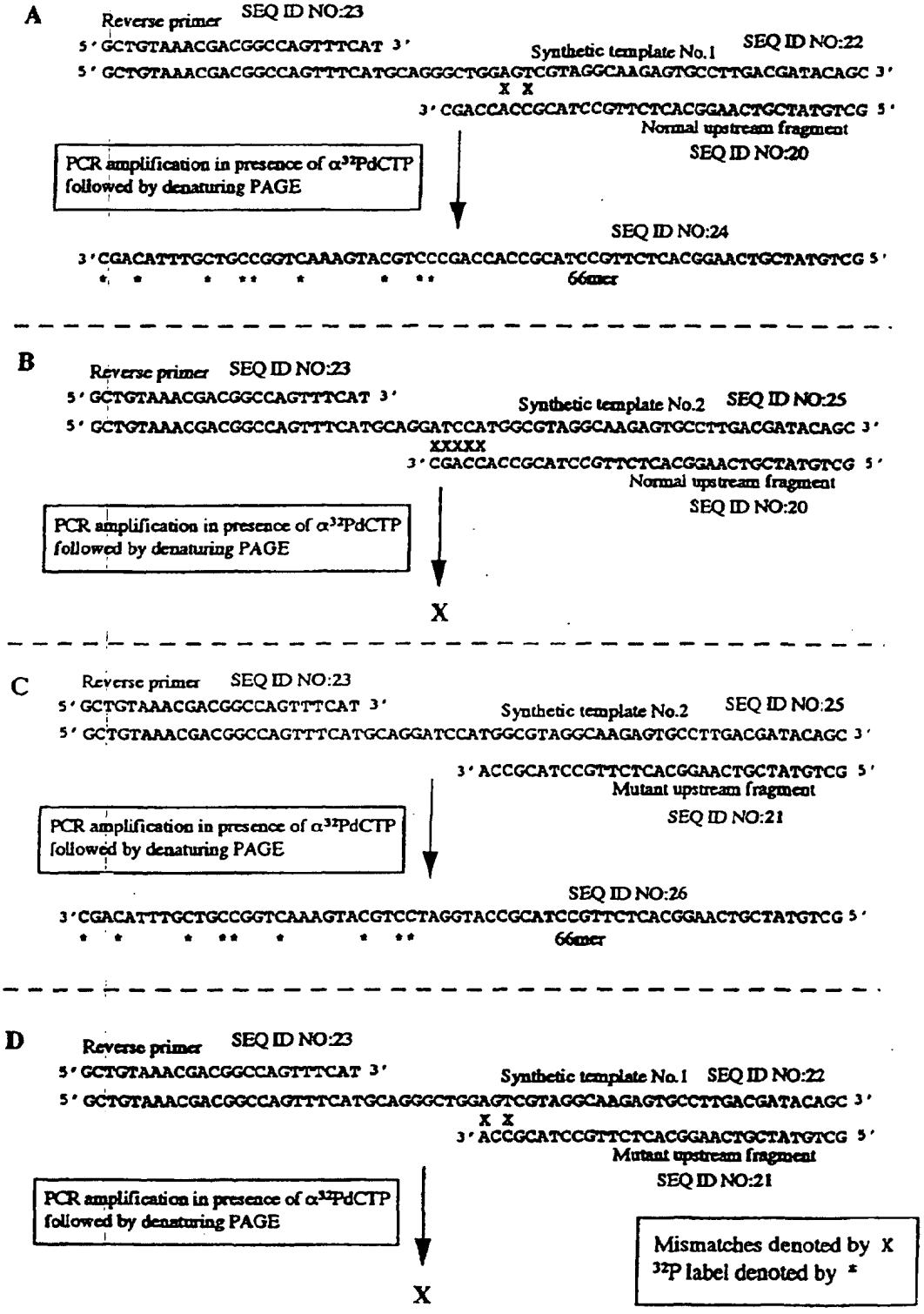


FIG 6

IDT Exhibit 1020.9746

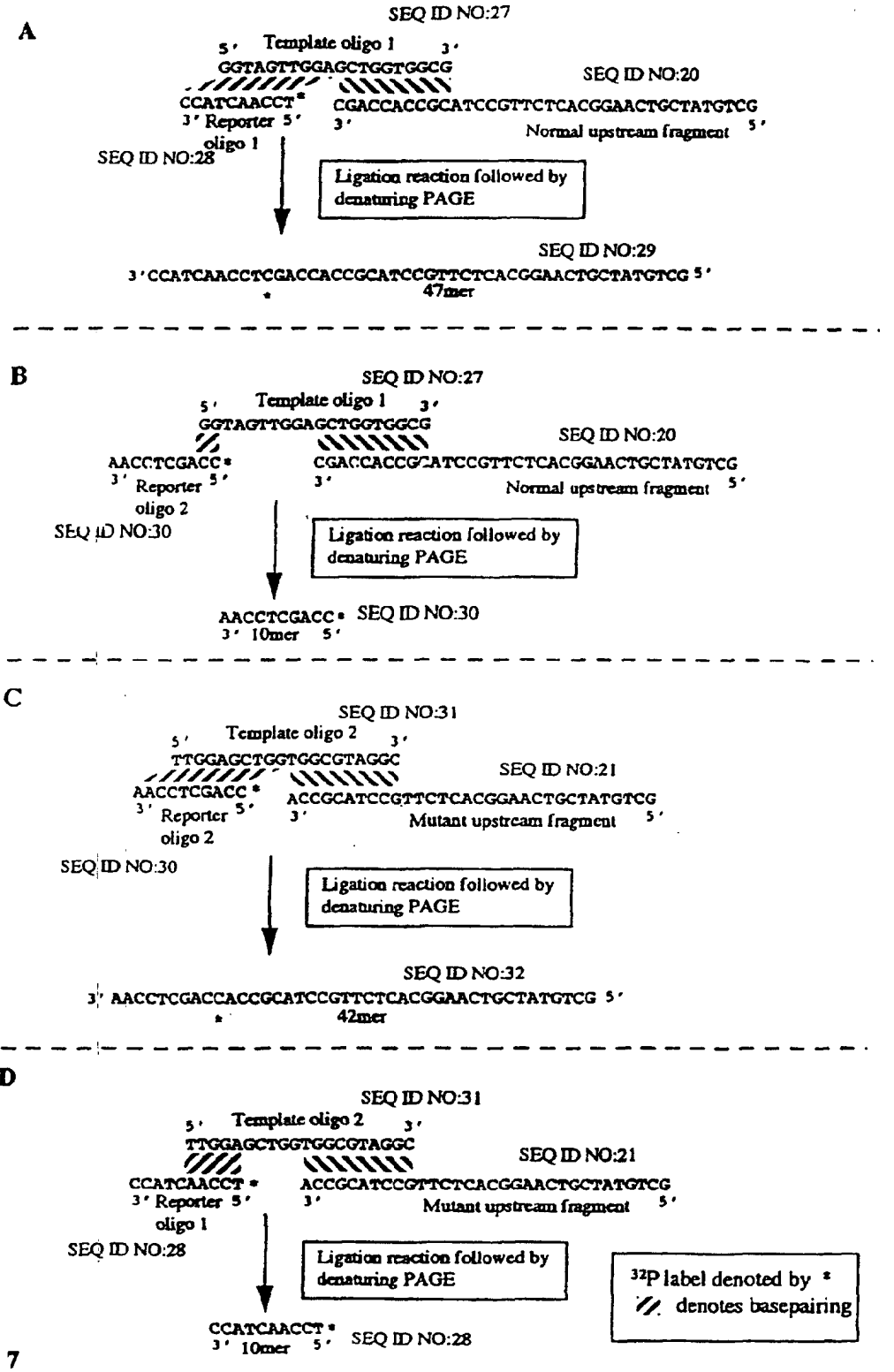


FIG 7

IDT Exhibit 1020.9747

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 365 627 B1**

(12) # EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **22.12.93** (51) Int. Cl.[5]: **C12Q 1/68**

(21) Application number: **89903936.6**

(22) Date of filing: **13.03.89**

(86) International application number:
PCT/US89/01025

(87) International publication number:
WO 89/09284 (05.10.89 89/24)

(54) **CATALYTIC HYBRIDIZATION SYSTEMS FOR THE DETECTION OF NUCLEIC ACID SEQUENCES BASED ON THEIR ACTIVITY AS COFACTORS IN CATALYTIC REACTIONS IN WHICH A COMPLEMENTARY LABELED NUCLEIC ACID PROBE IS CLEAVED.**

(30) Priority: **24.03.88 US 173127**

(43) Date of publication of application:
02.05.90 Bulletin 90/18

(45) Publication of the grant of the patent:
22.12.93 Bulletin 93/51

(84) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(56) References cited:
EP-A- 0 185 494
EP-A- 0 227 976
EP-A- 0 246 864

(73) Proprietor: **UNIVERSITY OF IOWA RESEARCH FOUNDATION**
Room 203,
Gilmore Hall
Iowa City, Iowa 52242(US)

(72) Inventor: **WALDER, Joseph, A.**
2107 Slagle Circle
Iowa City, IA 52240(US)
Inventor: **WALDER, Roxanne, Y.**
2107 Slagle Circle
Iowa City, IA 52240(US)

(74) Representative: **Thomson, Paul Anthony et al**
Potts, Kerr & Co.
15, Hamilton Square
Birkenhead Merseyside L41 6BR (GB)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

Rank Xerox (UK) Business Services
(3.10/3.6/3.3.1)

EP 0 365 627 B1

IDT Exhibit 1020.9748

Description

1. Field of the Invention

This invention provides a diagnostic assay methodology for the sensitive detection of specific nucleic acid (DNA and RNA) sequences in biological and clinical samples. The invention particularly addresses the problem of detecting low copy numbers of a target sequence with high specificity.

All current methods for the detection of DNA and RNA are based on the stoichiometric (1:1) hybridization of a complementary, labeled nucleic acid probe to the target sequence (see Figure 1). The present invention is a marked departure from this approach: the target sequence serves as a cofactor for a reaction in which the probe is cleaved catalytically. After cleavage of the probe, the target sequence is released intact and can be repeatedly recycled through the reaction leading to a large amplification of the response (see Figure 2). We have termed such reaction schemes catalytic hybridization amplification reactions. In such assays, cleavage of the labeled probe signals the presence of the target sequence. The sensitivity of the method is far greater than existing procedures, all of which are based on stoichiometric or unitary hybridization schemes in which the target sequence captures one, and only one, molecule of the probe. Several embodiments of the catalytic hybridization amplification method are described below. All operations required for such assays are simple and can be readily carried out in a clinical laboratory. The invention will have applications in the diagnosis of a variety of diseases in man and other species including genetic disorders, infectious diseases, and cancer.

2. Description of Prior Art

It is well known that nucleic acids, i.e., deoxyribonucleic acids (DNA) and ribonucleic acids (RNA) are essential building blocks of all organisms. These are high molecular weight polymers that are made up of many nucleotide units, each such nucleotide unit being composed of a base (a purine or a pyrimidine), a sugar (which is either ribose or deoxyribose) and a molecule of phosphoric acid. DNA contains deoxyribose as the sugar moiety and the bases adenine, guanine, cytosine, and thymine (which may be represented as A, G, C and T, respectively). RNA contains ribose instead of deoxyribose and uracil (U) instead of thymine.

The nucleotide units in DNA and RNA are assembled in definite linear sequences which determine specific biological functions. In bacterial cells, and in all higher species, DNA directs its own replication, and also serves as the template for the synthesis of RNA molecules whose nucleotide sequences carry the information encoded by the DNA. The process of RNA synthesis is termed transcription. RNA molecules serve several different functions within the cell. Messenger RNAs (mRNAs) direct protein synthesis. Ribosomal RNAs (rRNAs) are important constituents of ribosomes, the organelles within the cell at which mRNAs are read, or translated, and proteins are made.

Collectively, the genetic information of an organism is termed the genome. The genome of bacteria and all higher species is composed of DNA. The genome of viruses may be either DNA or RNA. In any case, the genome of any particular species, whether a virus, bacteria or a higher organism has a characteristic nucleotide sequence which, stated simply, can be viewed as the "fingerprint" of that species. Sequences within a genome that code for proteins, or that are transcribed to form RNAs with specific functions such as ribosomal RNA represent individual genes. Small viruses such as the AIDS virus have as few as 10 genes. The human genome contains approximately 50,000 genes.

According to the well known Watson-Crick model, DNA molecules consist of two polynucleotide strands coiled about a common axis. The resulting double helix is held together by hydrogen bonds between complementary base pairs in each strand. Hydrogen bonds are formed only between adenine (A) and thymine (T) and between guanine (G) and cytosine (C). Hence within the double helix, adenine (A) and thymine (T) are viewed as complementary bases which bind to each other as A-T. The same is true for guanine (G) and cytosine (C) as G-C.

In a single polynucleotide chain, any sequence of nucleotides is possible. However, once the order of bases within one strand of a DNA molecule is specified, the exact sequence of the other strand is simultaneously determined due to the indicated rules of base pairing. Accordingly, each strand of DNA molecule is the complement of the other. The process by which two complementary DNA strands associate with one another is termed hybridization. The process of strand separation, which is generally accomplished by heating the sample, is termed melting or denaturation of the duplex. Duplexes have a characteristic melting temperature (Tm) depending primarily on length, composition, and salt concentration. The following equation provides a useful approximation of the Tm:

IDT Exhibit 1020.9749

$$Tm = 16.6 \times \log Cs + 0.41 \times (\%G + C) + 81.5 - 820/L \quad (1)$$

where Cs is the salt concentration, (%G + C) is the percent G-C content, and L is the length of the duplex (Schildkraut C. and Lifson, S. (1965) Biopolymers 3, 195-208; and Thomas, C.A. Jr. and Dancis, B.M. (1973) Journal of Molecular Biology 77, 43-55). Mismatches between the two strands of a duplex (e.g. a C-T base pair) can occur, but decrease the stability of the helix significantly and lower the melting temperature.

In addition to DNA:DNA duplexes, DNA:RNA duplexes and RNA:RNA duplexes can also form and occur naturally. DNA:RNA duplexes are formed transiently in the process of transcription in which DNA serves as a template for RNA synthesis. RNA:RNA duplexes occur as the genetic material of certain viruses and are also formed in so-called hairpin loops of ribosomal RNA and transfer RNA. In duplexes involving an RNA strand, uracil (U) pairs with adenine (A). Within a single polynucleotide strand, sequences of both RNA and DNA can also occur. Such RNA-DNA copolymers are formed as intermediates in the replication of DNA. The melting temperature of a DNA:RNA or an RNA:RNA duplex may differ somewhat, depending on solution conditions, from that of a DNA:RNA duplex of the same sequence, but the equation given above remains a useful approximation of the Tm. For a further discussion of these various polynucleotide structures and their biological functions see "Genes III" by Benjamin Lewin, John Wiley and Sons, 1987.

Recent advances in molecular biology employing recombinant DNA techniques have led to new diagnostic and therapeutic strategies. In the area of DNA probe diagnostics, DNA probes are used to detect the presence of a complementary target nucleic acid sequence in the sample. Applications include the diagnosis of infectious diseases, cancer and genetic disorders. In the case of infectious diseases, the target can be a DNA or RNA sequence unique to a particular bacteria or virus.

In certain cancer cells there are specific gene rearrangements which can be detected, e.g. the translocation of the c-abl oncogene in chronic myelogenous leukemia. Genetic defects can involve large deletions or insertions of DNA, or a single point mutation as in the case of sickle cell disease. In the latter case, it is necessary to detect the alteration of a single base pair in the context of flanking sequences that are identical to those found in the normal gene.

All nucleic acid diagnostic assays involve three fundamental steps:

(1) Isolation and further preparation of DNA or RNA from the sample. When DNA is isolated it is generally sheared by mechanical forces. With gentle manipulation, the largest fragment size may be up to about 100,000 base pairs. DNA may also be purposefully cut with sequence specific endonucleases termed restriction enzymes to give rise to fragments of discrete sizes. Such is the case in so-called Southern blot analysis in which the fragments are subsequently separated according to size by electrophoresis and then immobilized on a filter such as nitrocellulose or a nylon membrane (Southern, E.M. (1975) Journal of Molecular Biology 98, 503-517).

(2) Hybridization of a complementary, labeled nucleic acid probe to the target sequence. In this process, the sample is denatured, generally by heating, to ensure that the target sequence is single stranded. The labeled probe is then allowed to hybridize to the target. Molecules of the probe that are weakly bound to other sequences within the sample, or the filter on which it is immobilized, are then washed away. To achieve adequate specificity, the melting temperature of the probe:target duplex must be at least 5 to 10 degrees above that of duplexes that may be formed between the probe and other sequences present within the sample with which there is one or more mismatches.

(3) Detection of the amount of the labeled probe hybridized to the DNA or RNA within the sample. Suitable reporter groups with which the probe may be labeled include radioisotopes such as ^{3}H, ^{32}P or ^{121}I, fluorescent dyes such as fluorescene, Texas Red or the phycobiliproteins, or enzyme markers such as alkaline phosphatase, β-galactosidase, glucose oxidase or peroxidases. An example of a diagnostic assay in which the sample nucleic acid is immobilized on a filter is shown in Figure 1 (see Falkow et al. (1982) U.S. Patent No. 4,358,535).

In addition to the filter format, a variety of other strategies for diagnostic hybridization assays have been described, see for example Heller et al. (1983) EPA 0 070 685 and EPA 0 070 687; Rentz, M. and Kurz, C. (1984) Nucleic Acids Research, 12, 3435-3444; Ruth, J.L. (1984) P.C.T. W084/03285; and Kohne, D. (1984) P.C.T. W084/02721. In each of these systems the reporter group that is to be detected is attached directly to the probe as is shown in Figure 1. Alternatively, an indirect labelling scheme may be employed. In the method described by P. Kourlisky et al. (1979) GB 2 019 408 the probe is labeled with biotin. After the probe is hybridized to the target sequence, a complex of the enzyme β-galactosidase covalently linked to the protein avidin is added to the sample. Avidin binds specifically and with very high affinity to the biotin label attached to the probe. The probe:target duplex is then detected by providing an appropriate substrate for β-galactosidase which is hydrolyzed to give a fluorescent or colored product. Biotin-avidin probe systems have also been described by Ward et al. (1982) EPA 0 063 879, and by Engelhardt et al. (1984)

IDT Exhibit 1020.9750

EPA 0 097 373. More recently, Carrico (1987) EPA 0 209 702 have described an indirect labelling strategy in which a probe:target duplex formed between an RNA strand and a DNA strand are detected using an antibody which binds specifically to RNA:DNA duplexes.

All of the above prior art methods involve a stoichiometric or unitary hybridization reaction in which the target sequence binds to or captures one and only one molecule of the probe. The present invention provides a novel and improved hybridization assay method in which the target sequence is able to capture many molecules of the probe in a repeating series of reactions. This is accomplished by cleavage of the probe within the probe:target duplex such that the target sequence is released intact and can repeatedly recycle through the reaction pathway. We term such methods catalytic hybridization amplification (CHA) reactions. A generalized version of the CHA method is shown in Figure 2. In such assays, the signal that is detected results from the cleavage of the probe. The sensitivity of the method is far greater than that of earlier hybridization schemes, all of which are based on a stoichiometric, i.e. 1:1, association between the target sequence and the probe.

Several methods described previously involve cleavage of the probe used in the hybridization assay, but preclude recycling of the target sequence, the essential element of the present invention. Thus, Ashihara et al. (1985) EPA 0 142 299 describe a method in which the DNA:DNA duplex formed between the probe and the target sequence is cleaved by a restriction enzyme. The restriction enzyme recognizes a specific sequence, generally 4 to 6 base pairs in length, within the duplex and cleaves both the probe and the target strand. Because the target is cleaved, it is impossible for it to recycle through the reaction pathway. A procedure for detecting single base changes, i.e. point mutations, within DNA has been described by Myers et al. (1985) Science 230, 1242-1246 that is based on the cleavage of an RNA probe. A labeled RNA probe is first hybridized to the target DNA. The RNA probe hybridized to the DNA is then subjected to cleavage by the enzyme RNaseA. Excess, unbound molecules of the probe are fully degraded. If the RNA forms a perfect duplex with no base pair mismatches with the DNA, the hybridized RNA strand will not be cut. If there are one or more base pair mismatches, the RNA probe will be cut by the enzyme. Cleavage of the RNA probe at the site of the mismatch(es), which may be assayed by a number of different means, is thus used to detect the altered DNA sequence. Since RNaseA will cleave free, unhybridized RNA molecules, it is essential for the cleaved fragments to remain hybridized to the target sequence. This makes it impossible for the target DNA sequence to react repeatedly with multiple copies of the probe. An assay having some similar characteristics has recently been described by Duck et al. (1987) EPA 0 227 976. In this case, excess, unhybridized probe is first digested away with an enzyme that will not cleave the probe when bound in the probe:target duplex. An assay is then carried out to detect the remaining molecules of the probe hybridized to the target sequence. Because it is essential for excess, unhybridized probe molecules to be first degraded, the target sequence cannot turnover and react repeatedly in the reaction as in the present invention.

Hull Vary et al. (1986) EPA 0 200 057 has described a system in which hybridization of the target sequence to the probe displaces a third polynucleotide that is bound to the probe. At the end of the hybridization reaction, the displaced polynucleotide is degraded into mononucleotides which are detected. Again in this system, each molecule of the target sequence hybridizes with one and only one copy of the probe. Recycling of the target sequence cannot occur. A method to enhance the rate of nucleic acid hybridization has been described by Zapolski et al. (1985) P.C.T. WO 85/05685. In this procedure, the enzyme RecA and a single stranded DNA binding protein are used to promote the hybridization of the probe to the target sequence. This increases the rate of hybridization, but the reaction is still stoichiometric involving the binding of a single molecule of the probe by each molecule of the target. Turnover of the target sequence enabling it to capture multiple copies of the probe does not occur as in the present invention.

The prior art presented above represents a number of different strategies for labeling probes, carrying out hybridization reactions, and enhancing signal detection. There are as yet, however, no nucleic acid hybridization assay systems in widespread use. Existing methods lack adequate sensitivity for many clinical applications and often involve complex assay procedures difficult to carry out routinely in a clinical laboratory. All of the prior systems involve a stoichiometric or unitary hybridization of the probe to the target sequence, such that each target molecule can capture one an only one molecule of the probe. The present invention makes a fundamental departure from this concept.

The primary objective of the present invention is to provide an improved hybridization assay method in which the target sequence serves as a catalytic cofactor for the cleavage of a complementary, labeled nucleic acid probe. Recycling of the target through the reaction pathway enables it to capture many molecules of the probe, thereby leading to a large increase in the sensitivity of the assay.

IDT Exhibit 1020.9751

Yet a further objective of the present invention is to provide formats, reagents, and assay conditions for such catalytic hybridization amplification (CHA) reactions useful in diagnostic tests.

According to the present invention there is provided a method of detecting nucleic acid sequences in which said sequences serve as a cofactor for a catalytic reaction in which a complementary labeled nucleic acid probe is cleaved and detected, said method comprising:

hybridizing of a target sequence to a labeled nucleic acid probe to provide a probe:target sequence duplex;

cleaving of the labeled probe within the probe:target sequence duplex to release the target sequence intact;

recycling of the target sequence repeatedly through the reaction pathway; and

detecting the extent of cleavage of the labeled probe.

The invention also provides a probe useful for carrying out the method of the invention made of an oligonucleotide 15 to 50 nucleotides in length containing an RNA sequence and one or more additional sequences not cleavable by RNaseH, or a polynucleotide from 50 to several thousand nucleotides in length composed of RNA plus one or more interspersed sequences that are not cleavable by RNaseH.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic illustration of a standard hybridization assay for the detection of a DNA target sequence based on the formation of a stoichiometric (1:1) complex between the target sequence and a complementary, labeled nucleic acid probe. In this case the sample DNA is immobilized on a filter support.

Figure 2 depicts the use of the CHA method for the detection of a specific DNA sequence. The target DNA sequence serves as a catalytic cofactor for the cleavage of a complementary, labeled nucleic acid probe. Repeated recycling of the target sequence through the reaction (dashed arrow) leads to a large amplification of the signal and enhanced sensitivity compared to standard assay systems based on a stoichiometric or unitary hybridization of the probe to the target sequence.

Figure 3 illustrates the CHA method in a solid support format in which cleavage of the complementary labeled probe attached to the support is mediated by RNaseH.

Figure 4 illustrates a hybridization detection system in which two CHA reactions are coupled to produce an exponential amplification of signal.

Figure 5 demonstrates the use of a competition assay to detect a specific RNA sequence with the CHA method.

SUMMARY OF THE INVENTION

The present invention of catalytic hybridization amplification (CHA) provides a novel and improved nucleic acid hybridization method useful in diagnostic assays. The target sequence is not detected on the basis of stoichiometric (1:1) hybridization to the probe as in all previous assay systems. The invention concerns the target nucleic acid sequence serving as a highly specific cofactor for a catalytic reaction which leads to the cleavage of a complementary, labeled probe hybridized to the target. Molecules of the probe not hybridized to the target sequence are not cut. Upon cleavage of the labeled probe in the labeled probe:target duplex, the target sequence is released intact and can repeatedly cycle through the reaction pathway. In such assays, cleavage of the probe generates the signal that is detected. The ability of each molecule of the target to capture multiple copies of the probe leads to a large enhancement in the sensitivity of the method. Several embodiments of the CHA method are provided employing simple assay procedures suitable for use in clinical diagnostic tests. Applications of these methods include, but are not limited to, the diagnosis of infectious diseases, genetic defects and cancer. Both single point mutations as well as longer unique sequences can be detected.

DETAILED DESCRIPTION OF THE INVENTION

The invention of catalytic hybridization amplification is unique in that it represents the first example of a non-stoichiometric hybridization process. In such systems the target sequence (either DNA or RNA) serves as a cofactor in a catalytic reaction in which a complementary, labeled nucleic acid probe is cleaved. A generalized version of the CHA method is shown in Figure 2.

In the first step of the process, the target sequence hybridizes to the probe by the usual rules of base pairing. Once hybridized to the target sequence, the probe is then cleaved into two or more smaller fragments. The cleavage reaction can be catalyzed by an enzyme or occur by other chemical methods.

IDT Exhibit 1020.9752

Only molecules of the probe hybridized to the target sequence are cleaved. Excess molecules of the probe not hybridized to the target are not cut. The smaller fragments of the probe generated by the cleavage reaction form less stable duplexes with the target than does the parent molecule, and under conditions of the CHA reaction dissociate from the target. For dissociation to occur rapidly, the melting temperature of the duplexes formed with the smaller fragments should generally be 10 °C or more below the reaction temperature. For short oligonucleotide probes from about 15 to about 35 nucleotides in length, this will occur following a single cleavage near the middle of the probe sequence. For longer probes, which may be up to several thousand nucleotides in length, multiple cleavages must occur, generating fragments shorter than about 35 nucleotides in length, preferably between about 10 and 20 nucleotides long. The target sequence must not be cleaved in the reaction. It is released intact from the probe:target duplex and recycles repeatedly through the reaction pathway. At the completion of the CHA reaction the extent of cleavage of the probe is measured. This generally involves separation of the probe fragment bearing the reporter group label form the intact molecule. It is the cleavage of the probe that signals the presence of the target sequence, Turnover of the target sequence, enabling each molecule to capture and promote the cleavage of many copies of the probe, greatly enhances the sensitivity of the method. Several embodiments of the CHA method will now be described.

In the principal embodiment of this invention, as applied to the detection of a DNA sequence, the cleavage reaction is catalyzed by the enzyme RNaseH, and takes place only when the target DNA sequence is hybridized to a complementary RNA probe, forming an RNA:DNA duplex. RNaseH is highly specific: it will not cut free molecules of the probe that are not hybridized to the target DNA sequence, nor does it cleave the target DNA sequence. Hence, upon cleavage of the probe, the target sequence is released intact and can repeatedly cycle through the reaction pathway as outlined in Figure 2. A detailed description of the components (probes, labels, etc.), procedures, and conditions for this embodiment of the CHA method follows.

Probes that can be used in the invention range from oligonucleotides of 15 to 50 nucleotides in length to longer polynucleotides which may be up to several thousand nucleotides long. A probe of about 15 residues is the shortest sequence that can be used to selectively hybridize to a complementary target sequence (e.g. a unique viral sequence) against the background of human DNA (see Thomas, C.A., Jr. (1966) Progress in Nucleic Research and Molecular Biology 5, 315). More often, sequences of at least 20 to 25 nucleotides have to be used to assure a high level of specificity. Short oligonucleotides up to about 25 residues are generally labeled with a single reporter group, either near the 5'- or 3'-end of the strand. Longer sequences may be labeled up to a density of about one reporter group per 15 nucleotide residues. The composition of the probe can be entirely RNA. Alternatively, the probe may contain one or more interspersed sequences not cleavable by RNaseH. Such sequences not cleaved by RNaseH can be composed of DNA, or of RNA in which either the phosphate groups or the sugar moieties are modified so as to prevent cleavage by the enzyme. Appropriate modifications of the phosphate groups include, but are not limited to, alkyl or aryl phosphotriesters, hydrogen phosphonates, alkyl or aryl phosphonates, alkyl or aryl phosphoramidates, phosphorothioates, or phosphoroselenates. The preferred modification of the ribose sugar to prevent cleavage by RNaseH is the conversion of the 2'-OH group to an alkyl or aromatic ether. To permit recognition by RNaseH, the cleavable RNA sequence(s) must be at least 3 residues in length.

Probes containing interspersed sequences not cleavable by RNaseH focus the activity of the enzyme to specified regions and are, therefore, preferred If an oligonucleotide 26 resi-dues in length composed entirely of RNA was used as a probe, cleavage at many sites would be nonproductive, that is, would not lead to release of the target sequence. For example, cleavage of the strand between residues 23 and 24 would generate one fragment 23 residues in length and a second fragment of only 3 residues. The trinucleotide would dissociate from the target sequence very readily, but the 23-mer would form a duplex with the target sequence almost as stable as that formed with the parent molecule, and, hence, would not be released. Alternatively, if the oligonucleotide had the sequence $DNA_{11}RNA_4DNA_{11}$ cleavage would be limited to the central four RNA residues. The longest fragment so generated would be 15 residues. The Tm of this fragment would be about 20 °C to 25 °C below that of the parent 26-mer (see Equation 1) and would readily dissociate from the target sequence under the CHA reaction conditions. Thus, each time the probe was cleaved, the target sequence would be released and able to recycle through the reaction pathway. In the case of the pure RNA oligonucleotide probe two or more cleavages of the probe may be necessary to release the target. For longer probes, the length of the interspersed noncleavable sequences should be less than about 35 nucleotides long and preferable between 10 and 20 nucleotides in length, to permit rapid release of the fragments generated by RNaseH cleavage from the target sequence.

A mixed DNA-RNA-DNA oligonucleotide probe is particularly useful for the detection of single point mutations, as is required for the diagnosis of certain genetic diseases such as sickle cell disease. Consider

IDT Exhibit 1020.9753

a probe that forms a perfect duplex with the normal target sequence, and a duplex with the mutant gene in which there is a single base pair mismatch. The entire probe, including the flanking DNA sequences, provides the specificity for the target sequence normally achieved through hybridization. The greatest difference in the stability of the two helices is obtained with a relatively short oligonucleotide probe in which the mismatch is near the middle of the sequence. Even in this case, however, the melting temperature of the perfect duplex formed with the normal sequence is only about 5 °C to 10 °C greater than that for the duplex formed with the mutant sequence having the mismatch. Although it is possible to distinguish the normal form the mutant sequence based on this small difference using a standard hybridization format such as that shown in Figure 1, the reliability of the method is not sufficient for routine diagnostic use. The CHA reaction scheme, however, affords a further level of specificity that enables these two sequences to be readily differentiated. As noted above, cleave of an DNA-RNA-DNA probe by RNaseH is limited to the central RNA sequence. If the RNA sequence is four or five residues, a single base pair mismatch disrupts the structure of the RNA:DNA duplex sufficiently to prevent cleavage of the RNA strand by RNaseH. Thus, the probe would be cleaved when hybridized to the normal target sequence but not when hybridized to the mutant sequence. Only the normal target sequence would serve as a catalytic cofactor for the cleavage of the probe. If the central RNA sequence is between six and eight residues, two mismatches within the RNA:DNA portion of the duplex may be necessary to disrupt the structure sufficiently to prevent cleavage by RNaseH. This can be accomplished using a probe in which there is a single mismatch with the normal sequence and two mismatches with respect to the mutant sequence, e.g.

```
.....dAdCrArUrCrCrGrArUdAdG... probe
                *
.....dTdGdTdAdTdGdCdTdAdTdC... normal target sequence

.....dAdCrArUrCrCrGrArUdAdG.... probe
                * *
.....dTdGdTdAdTdGdAdTdAdTdC.... mutant sequence
```

where dN (N = A,C,T or G) is a deoxyribonucleotide (DNA), rN a ribonucleotide (RNA), and * a mismatched base pair.

Probes themselves can be made and modified by both molecular biology and synthetic techniques. Long probes greater than about 60 nucleotides in length are generally synthesized and labeled by enzymatic methods. Long RNA fragments can be synthesized from a DNA template using either SP6 or T7 RNA polymerase. Oligonucleotides are generally synthesized by chemical methods, see "Oligonucleotide Synthesis: A Practical Approach" (Gait, M.J., ed.) IRL Press, Oxford (1984). The phosphoramidite method is the preferred chemistry at present. The synthesis can be carried out with the growing oligonucleotide chain attached to a solid support, such as polystyrene or glass beads, or free in solution. DNA, RNA and DNA-RNA-DNA oligonucleotides can be prepared by these approaches. DNA, RNA and mixed DNA-RNA molecules can also be further ligated together in any combination using the enzyme T_4 RNA ligase. Any of the above techniques can be employed for the synthesis of probes useful with this invention.

Radioisotopes, fluorophores, lumiphores (chemilumenescent or biolumeniscent substrates), and enzymes can all be used as reporter groups in catalytic hybridization amplification systems. The fact that the system produces a large amplification response before the actual detection procedure is carried out, in most cases, significantly reduces the demand on the sensitivity of the reporter group being used. Reporters can be chosen for convenience in the assay system, or because they are better adapted to the assay format (instrumentation versus manual procedures), as well as for greatest sensitivity.

Radioisotope such as ^{32}P, ^{14}C, ^{125}I, and tritium can be easily incorporated into DNA and RNA probes with molecular biology techniques, either during synthesis or after completion. Probes synthesized chemically can also be labeled with radioisotopes. While radiolabeling has certain advantages for experimental and research purposes, it does not appear to be the appropriate method for new tests and assays in the clinical laboratory, and is therefore not preferred.

Nonisotopic labels of choice include: (1) fluorophores such as fluorescein, Texas Red, Lucifer Yellow, pyrenes, chelated lanthanides, and phycobiliproteins to name a few; (2) lumiphores such as luminol and derivatives; and (3) enzymes such as alkaline phosphatase, peroxidase, β-galacto-sidase, and luciferases to name a few. The various enzymes can be used to produce color reactions (alkaline phosphatase, peroxidase), fluorescent products (alkaline phosphatase), and luminescence (luciferase).

IDT Exhibit 1020.9754

Fluorophores and lumiphores can be covalently attached to RNA and DNA sequences by a variety of techniques. In the case of a nucleic acid probe in which the 3'-terminal nucleotide is a ribose residue, the ribose diol at the 3'-end of the sequence can be oxidized with periodate to form a reactive dialdehyde. A reductive alkylation reaction can now be carried out where the dialdehyde can form a Schiff's base with a variety of fluorescent derivatives containing a primary amino group (dansylethylenediamine, N-(1-pyrenesulfonyl)ethylenediamine, 5-((2-amino-ethyl)thiouridyl)-fluorescein, etc.). Primary amine derivatives of luminol and other lumiphores can also be coupled. The Schiff's base adducts can be reduced to a stable secondary amine linkage using sodium borohydride or sodium cyanoborohydride. Fluorescent derivatives containing hydrazine groups (fluorescein thiosemicarbazide, Texas Red hydrazide, coumarin hydrazide, Lucifer Yellow CH, etc.) can also be coupled to terminal aldehydes produced by periodate oxidation (see Odom O.W., Jr., et al. (1980) Biochemistry 19, 5947-5954).

In the case of synthetic oligonucleotides, an aliphatic primary amino group can be readily incorporated at the 5'-terminal end of the molecule in several ways.[1] A specially protected phosphoramidite derivative of 5'-amino-5'-deoxythymidine can be incorporated as the final residue (Smith, L.M. et al. (1985) Nucleic Acids Research 13, 2399-2412). Primary amines can also be incorporated into oligonucleotides via special phosphoramidite derivatives, which can also be added at the last step in the synthetic procedure. Two such derivatives are available commercially, one from Applied Biosystems, Inc. (Aminolink 1, Part Number 400498) and one from ChemGenes, Inc: in which the amino group is protected with the monomethoxytrityl group. Once a primary amine is incorporated into the oligonucleotide, a number of amine selective fluorescent derivatives (fluorescein-5-isothiocyanate, Texas Red, succinimdyl-1-pyrene butyrate, etc.) can be easily coupled to the primary amine group under mild conditions. Fluorescent lanthanides (Europium, Terbium, etc.) can be incorporated into oligonucleotides by first reacting the primary amine with diethylenetriaminepentacetic acid anhydride to produce a chelator for binding the lanthanides.

Enzyme labels of particular interest include: alkaline phosphatase, peroxidase, β-galactosidase, glucose oxidase, bacterial (FMN/NADPH) luciferases, and insect (ATP) luciferases to mention a few. Other labels include microperoxidases, functionalized heme derivatives, and other metal chelates with catalytic activity. Enzymes can be incorporated into RNA, DNA and mixed RNA-DNA probes using procedures similar to those discussed above for fluorescent derivatives. As an example, alkaline phosphatase can be incorporated into an oligonucleo-tide probe con taining a 5'-terminal primary amine in the following manner. The primary amine containing oligonucleotide is reacted with an excess of the bifunctional reagent disuccinimidyl suberate (DSS is a highly specific reagent for coupling primary amines under mild conditions) to produce the mono adduct with the oligonucleotide. The mono adduct DSS-oligonucleotide can be easily purified. The mono adduct DSS-oligonucleotide is now reacted with alkaline phosphatase under mild conditions. The final product "alkaline phosphatase-DSS-5'-oligonucleotide" is purified by gel filtration and ion-exchange chromatography. The specifics of this coupling procedure are given in Jablonski, E. et al. (1986) Nucleic Acid Research 14, 6115-6128.

In this first embodiment of the CHA method, RNaseH is responsible for catalyzing the cleavage of the probe once it hybridizes to the target sequence. The enzyme is highly specific; it is a ribonuclease that cleaves only the RNA strand within an RNA:DNA duplex. The enzyme can be obtained from both eukaryotic and bacterial sources. The E. coli RNaseH enzyme, used in the examples, is a single polypeptide chain of about 17,500 molecular weight. The enzyme is very stable, the gene for the protein has been cloned and sequenced, and overproducing strains have been constructed which produce the enzyme in large amounts (Kanaya, S. and Crouch, R.J. (1983) Journal of Biological Chemistry 258, 1276-1281. For these reasons, its use is generally preferred. For CHA reactions carried out above a temperature of about 50°C, a more thermostable form of the enzyme is desirable. Such variants of the enzyme may be isolated from thermophilic organisms, or produced by mutagenesis of E. coli RNaseH using recombinant DNA techniques (e.g. by the introduction of a suitably positioned disulfide bridge).

The optimal temperature for carrying out the CHA reaction is generally from about 5 °C to about 25 °C below the melting temperature of the probe:target duplex. This provides for a rapid rate of hybridization and high degree of specificity for the target sequence. Consider, for example, a DNA-RNA-DNA probe 26 residues in length in which the G plus C fraction is 50%. At a salt concentration of 0.1 M, the predicted melting temperature of the probe:target duplex, according to Equation 1, is about 55 °C. The upper limit for the CHA reaction temperature is, therefore, about 50 °C. Cleavage of the probe near the middle of the sequence generates two fragments each of which have melting temperatures of about 25 °C. As noted above, for these fragments to dissociate rapidly from the target sequence, the CHA reaction temperature should be at least 10 °C greater than this value, i.e. equal to or greater than 35 °C. Hence, the optimal range

[1] Oligonucleotides are generally synthesized from the 3'-to the 5'-end of the chain.

IDT Exhibit 1020.9755

for the CHA reaction temperature with this probe is from about 35 ° C to about 50 ° C.

The CHA reaction is usually carried out for about 15 minutes to about 1 hour. Generally, each molecule of the target sequence will turnover between 100 and 1000 times in this period, depending on the length and sequence of the probe, and the specific reaction conditions (see Example 1). For each copy of the target sequence present in the test sample 100 to 1000 molecules of the labeled probe will be cleaved by RNaseH. This level of amplification leads to an enormous increase in sensitivity compared to assays based on a stoichiometric hybridization format. Even higher levels of amplification can be obtained by allowing the CHA reaction to proceed longer.

During the CHA reaction, it is necessary to suppress cleavage of the probe, as well as the target sequence, by nonspecific nucleases. This is important for assays based on standard hybridization formats as well. Such nucleases are generally removed from the sample during the isolation of the DNA by heating or extraction procedures. A number of inhibitors of single-stranded ribonucleases such as vanadate, Inhibit-ACE (5 Prime - 3 Prime, Inc.) and RNAsin, a placental protein, do not affect the activity of RNaseH. In order to further protect the probe from nonspecific degradation, it is desirable to include such inhibitors during the CHA reaction.

Following the CHA reaction, it is necessary to determine the extent of cleavage of the probe. This is usually accomplished by separating the cleavage products bearing the label from the remaining uncleaved molecules of the probe. Methods for such separation include, but are not limited to:

(1) Electrophoresis of the CHA reaction products to separate the cleaved fragments from molecules of the intact probe on the basis of size.

(2) The use of strong acids, like trichloroacetic acid, to selectively precipitate the relatively large uncleaved labeled probes from the smaller cleaved fragments.

(3) The incorporation of a high affinity label into the probe, in addition to the reporter group, to provide for the separation of the uncleaved probes after the CHA reaction, using an affinity support.

(4) The use of a support material containing a complementary polynucleotide sequence to immobilize the uncleaved labeled probe after the CHA reaction is carried out, leaving the cleaved fragments in solution.

In the first separation method, electrophoresis of the CHA reaction products through a gel matrix is used to separate the cleaved fragments bearing the label from the intact probe. The smaller fragments migrate through the gel more rapidly than the larger intact molecule. For probes up to about 100 nucleotides in length a polyacrylamide gel matrix is generally used. For longer probes, polyacrylamide, agarose or a composite gel made up of these two components is generally preferred. After the separation has been achieved, fragments bearing the reporter group label and the original probe are resolved into separate bands which can be readily localized and quantitated.

The second separation method involving strong acid precipitation is used mainly in those situations in which a long probe (greater than about 50 nucleotides in length) containing either radioisotope or fluorescent labels is cleaved via the CHA process into a number of relatively small labeled fragments (<10 nucleotides in length). Introduction of a strong acid (trichloroacetic, etc.) into the solution causes the intact probe to precipitate, while the smaller cleaved fragments remain in solution. The solution can be centrifuged or filtered to remove the precipitate. The supernatant containing the cleaved labeled fragments can now be quantitated. A more detailed discussion of this method is given in Example 1, below.

The third separation method involves a probe labeled with both a high affinity group and a reporter group, and an affinity support is used to carry out the separation procedure. The probes used in this case can be either an RNA sequence or contain interspersed sequences not cleaved by RNaseH such as a mixed DNA-RNA-DNA sequence. The probe is designed with an affinity label (such as biotin, avidin, lectins, haptens, or antibodies) incorporated at or near one of the terminal positions of the probe. The reporter group label (enzyme, fluorophore, lumiphore or radioisotope) is incorporated at or near the opposite terminal position. (With a long polynucleotide probe, many or all of the interspersed sequences not cleavable by RNaseH can be labeled with the reporter group). The cleavable portion of the probe sequence, thus, lies between the affinity group and the reporter group. A minimum of about 15 nucleotides is necessary between the two labels, with 20 to 50 nucleotides being more ideal. The same methods as discussed above for incorporating reporter group labels into probes are also used for the attachment of affinity labels. In this separation procedure it is also necessary to have support material which contains a corresponding affinity group which binds specifically and strongly to the affinity label incorporated into the probe. For example, if the affinity label biotin is incorporated into the probe, then the appropriate affinity support material should contain avidin or streptavidin. Many support materials such as glass beads, latex beads, cross-linked dextran (Sepharose®) beads, etc. containing covalently linked avidin are commercially available.

After the CHA reaction is carried out, the support material is added to the sample. Molecules of the intact probe bind tightly to the support. Cleaved fragments bearing the reporter group but not the affinity

IDT Exhibit 1020.9756

label remain in solution and can be quantitated.

The fourth separation procedure is in principle very similar to that just described, but employs a nucleic acid sequence within the probe not cleaved during the CHA reaction as the affinity label. A complementary nucleic acid sequence is attached to a solid support and serves as a catcher sequence. When the cleavage of the probe during the CHA reaction is mediated by RNaseH, the sequence playing the role as the affinity label can be DNA, or an RNA sequence in which the sugar or phosphate groups are modified to prevent cleavage by RNaseH. The affinity label sequence is separated from the reporter group by a portion of the probe cleaved by RNaseH during the CHA reaction, for example

5'-DNA-RNA-DNA-(reporter group)

in which case the DNA sequence at the 5'-end of the probe, or a portion thereof, serves as the affinity label. An affinity label sequence of 20 to 50 nucleotides in length is most ideal, but longer sequences can also be used. Sequences less than 15 nucleotides are least useful. A portion of the affinity label sequence can be complementary to the target sequence or it can be a separate appendage attached to the probe. The latter is preferred because a generalized catcher sequence can then be used. In this case one set of catcher beads can be used with many different target specific probes.

A number of different chemical methods may be used to attach the catcher sequence to a solid support, see for example, Gilham, P.T. (1971) Methods of Enzymology, Vol. 21 (Grossman, L. and Moldave, K., eds.) pp. 191-197, Academic Press, New York; Potuzak, H. and Dean, P.D.G. (1978) Nucleic Acids Research 5, 297-303; Robertson, D.L. and Davidson, N. (1972) Biochemistry 11, 533-537; and Moss, L.G. et al. (1981) Journal of Biological Chemistry 256, 12655-12658.

It is also possible to synthesize the catcher sequence directly on the solid support which later will be used in the assay. As noted above, oligonucleotides are generally synthesized from the 3'- to the 5'-end of the chain. If the synthesis is carried out on a solid support, the first nucleotide from the 3'-terminal end is generally attached to the support via its 3'-hydroxyl group through an ester linkage. Conditions required to remove protecting groups on the nucleotide bases (treatment with concentrated ammonia at 55 °C for 6 to 10 hours) cleave this linkage readily. If the oligonucleotide is to remain bound to the support the linkage to the 3'-hydroxyl group must be changed to one not cleaved under these conditions. Suitable functional groups through which the 3'-OH may be attached include ether linkages, phosphate triesters, phosphate diesters and aromatic carbamates. The functional group preferred for simplicity of synthesis is a β-cyanoethyl phosphtriester. The attachment is accomplished by the reaction of a 3'-β-cyanoethyl-N,N-diisopropyl phosphoramidite derivative of the first nucleotide (available commercially for all four bases) to a solid support containing free OH groups. Such supports can be prepared from a variety of materials including glass, polystyrene, latex or cross-linked dextran (Sepharose®) beads, cellulose, or nylon or tetrafluoroethylenehomopolymer (teflon®) membranes. After the first nucleotide is attached to the support, the remaining sequence is synthesized by the normal phosphoramidite method. At the culmination of the synthesis, the deblocking procedures convert the β-cyanolethyl phosphotriester to a phosphodiester through which the oligonucleotide remains linked to the support. After a support material is prepared by any of the above procedures, the attached oligonucleotide can be extended, if desired, by a number of different enzymatic methods, thereby yielding longer nucleotide sequences bound to the support. Suitable procedures include, but are not limited to, addition of nucleotide residues to the fragment with DNA polymerase or terminal transferase, and ligation of polynucleotides to the initial sequence attached to the support using DNA or RNA ligase.

In each of the four separation methods just described, the CHA reaction is first carried out in solution. Alternatively, the labeled probe may be attached to a solid support, in which case cleavage of the probe during the CHA reaction releases fragments bearing the reporter group into solution (see Figure 3). The major advantage of this approach is the relative ease of the separation step after the CHA reaction is completed: the solid support material containing remaining molecules of the intact probe are simply removed from the sample by physical means such as filtration or centrifugation; the cleaved fragments bearing the reporter group are left in solution and can be quantitated. Suitable solid support materials include glass, polystyrene, latex or cross-linked dextran (Sepharose®) beads, cellulose, or nylon or teflon® membranes. Magnetized beads can also be used, in which case the beads are separated from the solution by the application of a magnetic field. The probe can be attached to the solid support either through a covalent or a noncovalent (e.g. biotin-avidin) linkage. A covalent linkage is preferred because of its greater stability. Suitable procedures for covalently attaching nucleotide sequences to solid supports by both chemical and enzymatic methods are described in the preceding paragraph.

When using the solid support CHA method, it is desirable to first cut the sample DNA into smaller fragments so that the target sequence can readily diffuse to the surface of the bead to hybridize with the bound probe. This can be accomplished by shearing the DNA through mechanical forces (e.g. sonication) or

IDT Exhibit 1020.9757

by cleaving the DNA at specific sites using a restriction enzyme. Alternatively, a small fragment containing the target sequence can be synthesized with DNA polymerase using the sample DNA as a template. In this procedure an oligonucleotide used as a primer is first hybridized to the sample DNA upstream (i.e. 5') from the target sequence. The primer is then extended across the target sequence (copying the complementary strand) using DNA polymerase. The product:

5'-Primer-Target Sequence

is then brought into the CHA reaction. It is also possible to repeat the reaction with DNA polymerase a number of times to increase the number of copies of the target sequence. Such a procedure is carried out in the so-called polymerase chain reaction (Mullis, K.B. et al. (1986) EPA O 200 362; and Mullis, K.B. (1986) EPA 0 201 184). Since target amplification and the CHA method involve entirely separate steps in the assay procedure the two can be used in combination. This can be useful to detect a nucleic acid sequence present originally in the sample in extremely low copy number. For most diagnostic applications, however, the high degree of sensitivity of the CHA method alone is sufficient.

All of the above discussions on the use of the CHA method have centered around the detection of DNA target sequences. In some cases, it may be advantageous to detect RNA target sequences, such as ribosomal RNA sequences. This can be readily accomplished with any of the above CHA reaction procedures utilizing RNaseH by first synthesizing a complementary DNA (cDNA) copy of the original RNA sequence with the enzyme reverse transcriptase. In this procedure, an oligodeoxynucleotide primer is first hybridized to the sample RNA upstream (i.e. 5') of the target sequence. With the addition of reverse transcriptase and the four deoxynucleotide triphosphates (A, G, T and C), the primer is extended producing a cDNA copy of the target RNA. The product:

5'-Primer-cDNA Target Sequence

now forms the substrate for the CHA reaction. For a more detailed description of protocols for the use of reverse transcriptase see Maniatis, T., Fritsch, E.F., and Sambrook, J., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, pp. 128-132, 1982.

A second embodiment of the CHA method will now be described in which a phosphorothioate derivative of the target sequence is used as the substrate and a complementary labeled DNA probe is cleaved by a restriction enzyme.

Restriction enzymes recognize a specific DNA sequence, generally 4 to 6 nucleotides in length, and cut both strands of the DNA duplex at that site. Several hundred of these enzymes are now known with many different recognition sequences. These enzymes generally cannot be used with the CHA method since the target sequence, as well as the probe, would be cut, and therefore would be unable to recycle through the reaction pathway. However, if the phosphate groups of one of the two strands of a DNA duplex are phosphorothioates, rather than normal phosphate diesters, many restriction enzymes (e.g. Ava I, Ava II, Ban II, Hind II, Nci I, Pst I, and Pvu I) will cleave only the unmodified strand (Taylor, J.W. et al. (1985) Nucleic Acids Research 13, 8749-8764). The phosphorothioate sequence is not cut and hence can serve as a catalytic cofactor for cleavage of a complementary, labeled probe by a restriction enzyme. The target sequence to be used in such an assay must of course contain the recognition site for the restriction enzyme. The modified DNA sequence is prepared from a DNA template using DNA polymerase. If the sample is RNA, reverse transcriptase is used. These reactions are carried out as described above, except that deoxynucleotide α-thiotriphosphates are used as substrates instead of the normal deoxynucleotide triphosphates. All four of the deoxynucleotide α-thiotriphosphates are available commercially. Once the modified target sequence is prepared, the CHA reaction can be carried out using any of the assay formats described above. The same methods for synthesis and labeling of the probe are also applicable. Modified DNA or RNA sequences can be incorporated within the probe flanking the restriction enzyme site, but the recognition site itself must be composed of DNA.

All of the various embodiments and formats of the CHA method described thus far provide a linear amplification of the signal. Since the concentration of the target sequence remains constant, the velocity of the reaction remains unchanged, and the number of molecules of the labeled probe that are cleaved increases linearly with time (see Example 1). As mentioned above, the turnover number, i.e. the number of molecules of the probe cleaved per target sequence per hour, generally ranges between 100 and 1000.

CHA reactions can also be carried out in novel manners that produce an exponential amplification, and therefore, offer an even greater sensitivity. In general, an exponential response is achieved by coupling two or more CHA reactions so as to produce additional copies of the target sequence. A number of different exponential CHA formats can be designed. One version of an exponential CHA format is shown in Figure 4. In this case, two sets of probe sequences are immobilized on solid supports. The first probe sequence on support I contains a cleavable sequence (T') complementary to the original target sequence, and a non-cleavable sequence (X). The sequence (X) is complementary to a cleavable sequence (X') immobilized on

IDT Exhibit 1020.9758

support II. In the first step of the process, target DNA from the sample hybridizes to the complementary cleavable sequence (T') on support I. Once hybridized to the target sequence, (T') is cut, releasing the sequence (X) from the support. (X) then diffuses from the support and hybridizes with the cleavable sequence (X') on support II. This cannot occur until (X) is released from the support. Sequences (X) and (X') while immobilized on two separate macroscopic supports cannot hybridize with one another. Once the duplex between (X) and (X') is formed, (X') is cleaved. This releases a second copy of the target sequence which can in turn react with support I and further propagate the reaction. Thus, one cycle of the CHA reaction yields two copies of the target sequence; the second cycle yields four copies of the target sequence; the third cycle yields eight copies; etc. In the configuration shown, one reporter group label is released from support II for each molecule of target that is generated in the reaction. It can be shown that the amount of the free label released from support II is given by the following formula:

$$L = (A/2) \times (e^{kt} + e^{-kt}) - A \qquad (2)$$

where A is the initial number of copies of the target present in the sample, k is the square root of the product of the turnover numbers for the two separate CHA reactions, and t is time. The first term of the expression ($A/2 \times e^{kt}$) rapidly becomes dominant. The amount of free label increases exponentially with time, conferring an extremely high degree of sensitivity to the assay. If the turnover number for each of the CHA reactions is only 1 per minute, a single copy of the target sequence will cause 2.4×10^8 labels to be released into solution within 20 minutes. This number can be easily detected with simple fluorescent or enzyme labeling groups. Turnover numbers much greater than 1 per minute can be achieved (see Example 1). Coupled CHA assay systems, thus, provide the capability for single copy target detection with a reaction time of only a few minutes.

Finally, we describe a competition assay based on the CHA method to detect an RNA target sequence. In this case, both a DNA molecule and a labeled probe are added to the sample to be analyzed (Figure 5). The probe sequence (A') may be composed entirely of RNA or contain interspersed sequences that cannot be cleaved by RNaseH. The DNA fragment contains within it a sequence (A) complementary to the labeled probe. In the example shown, the central portion of the probe is RNA, and it is flanked on either side by DNA sequences. The RNA:DNA duplex formed when (A) and (A') hybridize together forms a substrate for RNaseH. RNaseH can cleave the labeled probe, releasing the DNA fragment intact which serves as a catalyst to further propagate the reaction. This would occur if the target RNA sequence were not present within the sample (Figure 5C). If the target RNA sequence were present, this reaction would be inhibited (Figure 5B). The target sequence contains the sequence (A'), as well as an adjacent sequence (B') further complementary to the DNA fragment. The DNA fragment thus forms a more stable duplex with the target sequence than with the probe. This enables the target molecule to compete very effectively for the binding of the DNA fragment, and thereby blocks the cleavage of the labeled probe by RNaseH (Figure 5B). This competitive inhibition of the cleavage of the labeled probe is the basis for the assay. The sensitivity of the method may be further enhanced by preventing the RNA-DNA duplex formed with the target sequence from being degraded by RNaseH. This may be accomplished by suitably modifying the DNA fragment. We have found that a variety of modifications of the phosphate backbone of the DNA strand will prevent cleavage of the complementary RNA strand by RNaseH. Such modifications include, but are not limited to, substitution of the normal phosphate diesters with alkyl or aryl phosphotriesters, or with alkyl or aryl phosphonates. These modifications would be incorporated into the DNA molecule used in the assay throughout region B, and in all of region A except for that portion which hybridizes to the central RNA sequence of the probe. If this unmodified sequence within region A is precisely complementary to the RNA sequence within the probe, but contains one or more mismatches with respect to the target sequence, only the probe will be cleaved by RNaseH. The mismatch(es) within the duplex formed between the DNA strand and the target molecule disrupt the helix, thereby, preventing cleavage of the target strand within this region by RNaseH. Cleavage of the target strand elsewhere is prevented by the modifications introduced into the phosphate backbone of the DNA molecule The mismatch(es) would decrease the stability of the duplex formed between the target sequence and the DNA molecule somewhat, but, due to the additional sequence complementary in region B, the stability of the duplex would remain much greater than that formed with the labeled probe.

The following examples are offered to further illustrate, but not limit, the process, product and techniques of the invention and serve to establish the unique features of the invention which form the basis of its utility.

IDT Exhibit 1020.9759

EXAMPLE 1

Demonstration of Catalytic Hybridization Amplification Mediated by RNaseH

The following example demonstrates the principle of the catalytic hybridization amplification method. The target sequence in this case was oligodeoxythymidylic acid of 12 to 18 nucleotide residues in length (oligo dT_{12-18}, purchased from Pharmacia). The molecule serving as the probe was polyriboadenosine labeled with tritium (Amersham). The specific activity of the poly rA used in the study was 574 mCi/nucleotide residue; the chain length varied from about 40 to 140 residues. In the presence of the target sequence, RNaseH catalyzes the cleavage of the labeled poly rA probe. Oligo dT is then released from the duplex intact and recycles repeatedly through the reaction as outlined in Figure 2.

Cleavage of the poly rA probe catalyzed by RNaseH gas compared under three conditions: in the absence of oligo dT; with an equivalent concentration of oligo dT (based on nucleotide residues); and in the presence of 1/1000 the concentration of oligo dT. Each reaction was conducted in a final volume of 100 ul containing 44 picamoles of poly rA (by nucleotide units), 50 mM Tris-HCl buffer pH 8.0, 20 mM KCl, 9 mM $MgCl_2$, 1 mM 2-mercaptoethanol and 25 μg of bovine serum albumin. The reaction was initiated by the addition of 2.3 units of E. coli RNaseH (Bethesda Research Laboratories) and was maintained at 30 °C. Aliquots of the reaction mixture were removed at 30 minutes and added to 0.2 ml of a carrier DNA solution (0.5 mg/ml of salmon sperm DNA, 0.1 M sodium pyrophosphate and 1 mM ethylenediamine tetraacetic acid) and 0.3 ml of 10% trichloroacetic acid. Trichloroacetic acid causes the precipitation of higher molecular weight nucleic acids over about 10 nucleotides in length. Smaller cleavage products, i.e. mononucleotides and short oligonucleotides, remain in solution. After incubation of the samples on ice for 20 minutes, the precipitated nucleic acids were pelleted by centrifugation at 16,000 X g for 15 minutes. The supernatants Containing the cleaved fragments were carefully removed, mixed with scintillation cocktail and counted for radioactivity.

In the control sample without oligo dT, less than 0.5% of the poly rA was cleaved in the course of 90 minutes. In the presence of an equivalent concentration of oligo dT nearly 100% of the poly rA was cleaved to trichloroacetic acid soluble fragments within 30 minutes. Cleavage of poly rA in the presence of 1/1000 the amount of oligo dT was linear over 90 minutes, at which time 21% of the poly rA was cleaved to trichloroacetic acid soluble fragments. Since oligo dT was present in a concentration sufficient to hybridise to only 0.1% of the poly rA sequence, each molecule of the oligo dT must have recycled through the reaction, on average, at least 210 times (2.3 times/minute). Since cleavage of the poly rA strand is in some cases nonproductive, i.e. does not lead to the release of labeled fragments small enough to remain in solution in the presence of trichloroacetic acid, the true turnover number must be even higher. In diagnostic tests based on the CHA method, turnover numbers exceeding 100/hour provide a great enhancement in sensitivity compared to assays based on a stoichiometric hybridization format.

EXAMPLE 2

Site-Specific Cleavage of β-Globin mRNA by RNaseH in the Presence of a Complementary DNA Target Sequence

The following is an important example in which the specificity of strand cleavage by RNaseH is clearly demonstrated. RNaseH was shown to cleave β-Globin mRNA at one unique site when a mixture of both α- and β-globin mRNAs was hybridized to a short 25-residue DNA target sequence complementary to a portion of the β-globin mRNA.

The oligodeoxynucleotide 5'-TGTCCAAGTGATTCAGGCCATCGTT (CODMB-25) was synthesized by the phosphoramidite method using a Beckman automated DNA synthesizer. It is complementary to the nucleotide sequence of mouse β-globin mRNA from residue 269 to 293. Mouse globin mRNA was isolated from reticulocytes obtained from mice rendered anemic by treatment with phenylhydrazine as described by Goosens and Kan (Methods in Enzymology (Antonini, E., Rossi-Bernardi, L. and Chiancone, E., eds.) Vol. 76, pp. 805-817, 1981). Two micrograms of the RNA was added to 1.0 nanomoles of CODMB-25 in a final volume of 25 microliters containing 10 mM Tris-HCl buffer at pH 7.5, 5 mM $MgCl_2$, 25 mM NaCl and 1 mM dithiothreitol. The concentration of the oligonucleotide is approximately 100-fold greater than the mRNA on a molar basis. The reaction was initiated by the addition of 2 units of E. coli RNaseH, and allowed to proceed for 30 minutes at 37 °C. The reaction was stopped by extraction of the sample with a 1:1 mixture (volume: volume) of chloroform and phenol. The RNA species remaining after digestion with RNaseH were then isolated by precipitation with ethanol and analyzed on Northern blots.

IDT Exhibit 1020.9760

The RNA samples were reacted with 1 M glyoxal in a 1:1 volume/volume mixture of 10 mM sodium phosphate buffer at pH 6.5 and dimethylsulfoxide for 1 hour at 50 ° C, and then electrophoresed on a 1.5% agarose gel. The RNAs were transferred from the gel to a Gene Screen Plus filter (DuPont-NEN) by capillary blotting according to the method described by the manufacturer. Prehybridization of the filter was done in 1% sodium dodecylsulfate, 1 M NaCl, 10% dextran sulfate, 50 mM sodium phosphate pH 6.5 for 2 hours at 35 ° C. The filter was then hybridized with either an α-or β-globin specific oligonucleotide probe. The probes were 25 residues in length and hybridize to the very 5'-end of the respective RNAs. Hybridization was conducted in 5 ml of the prehybridization buffer containing 200 microgram/ml denatured salmon sperm DNA plus 2 X 10^6 counts/min of the probe, radiolabeled with ^{32}P at the 5'-end of the molecule, for 24 hours at 35 ° C. After the hybridization, the filter was washed in the following order: once in 2X SSC (1X SSC = 0.15 M NaCl, 0.015 M sodium citrate) for 5 minutes at room temperature, twice in 2X SSC plus 1% sodium dodecylsulfate for 30 minutes at 35 ° C, and once in 0.1X SSC for 5 minutes at room temperature. The filter was blotted dry and exposed to Kodak® XAR-5 film for autoradiography.

In the absence of oligonucleotide, no cleavage of either α- or β-globin mRNA occurred during the reaction with RNaseH. The α- or β-globin specific probes each hybridized to a single band on the Northern blots corresponding to the full length mRNA. The reaction in the presence of CODMB-25 resulted in essentially complete cleavage of the β-globin mRNA. The size of the cleavage product detected with the 5'-β-globin specific probe was exactly that predicted for scission of the mRNA at the site of the RNA:DNA duplex, i.e. 280 +/- 20 nucleotides. No other cleavage products were observed. In addition, no cleavage of the α-globin mRNA was evident using the α-globin specific oligonucleotide probe. RNaseH, thus, has the requisite specificity for use in diagnostic tests based on the CHA method. An RNA sequence used as probe will be cleaved by the enzyme if, and only if, it encounters the complementary target DNA sequence within the sample; and free single stranded molecules of the probe not hybridised to the target sequence will not be cut.

EXAMPLE 3

Detection of Cytomegalovirus Using a CHA Method

Cytomegalovirus (CMV) is a large double-stranded DNA virus. Infection with CMV is clinically very important in immunocompromised patients. It is the most prevalent serious infectious disease complication following renal transplantation.

The following illustrates the use of the CHA method to detect CMV. The probe used is a CMV specific sequence 25 nucleotides in length derived from the major immediate early gene:

5'-TCTTGGCAGAGGACTCCATCGTGTC

The central six residues are RNA; the flanking sequences are composed of DNA. The probe is linked to latex beads at the 3'-end of the sequence. Alkaline phosphatase is attached at the 5'-terminal position. DNA is isolated from blood, urine and sputum samples obtained from the patient. The assay for CMV in each of these samples is carried out by the CHA method shown in Figure 3. Cleavage of the labeled probe is mediated by E. coli RNaseH (5 units/100 microliters). The CHA reaction is carried out for a period of 45 minutes at a temperature of about 40 ° C. RNasin (50 units/100 microliters) is added to the reaction mixture to suppress nonspecific cleavage of the probe. If CMV is present within the sample, the target sequence will hybridise to the probe, and the probe will be cleaved by RNaseH. Repeated cycling of the target sequence through the reaction pathway enables each copy of the target sequence to capture and promote the cleavage of many molecules of the labeled probe. Following the CHA reaction, the extent of cleavage of the probe is quantitated by measuring the amount of the reporter group, i.e. alkaline phosphatase, released into solution. The beads are first removed from the sample by filtration or centrifugation. The alkaline phosphatase label on the probe fragments in solution is then measured using either a colorimetric or fluorescent producing reagent detection system. The enzyme can be assayed spectrophotometrically at 405 nm by following the hydrolysis of the substrate p-nitrophenyl phosphate to the colored product p-nitrophenol at pH 8.5. Alternatively, the enzyme can be measured fluorometrically using the substrate 4-methylumbelliferyl phosphate. The fluorescent excitation and emission wavelengths used for the assay are 363 nm and 447 nm, respectively. The final color or fluorescent signal provides a measure of the amount of CMV present in the original sample.

In summary, the results of the studies presented demonstrate the utility of the CHA method as a means for the detection of specific nucleic acid sequences. In such assay systems, the target sequence serves as a cofactor for a catalytic reaction in which a complementary, labeled nucleic acid probe is cleaved. Each molecule of the target sequence is thus able to capture and promote the cleavage of multiple copies of the

IDT Exhibit 1020.9761

probe. This results in a large increase in the level of sensitivity of the method compared to current diagnostic tests, all of which are based on a stiochiometric hybridization reaction in which the target sequence is able to bind one, and only one, molecule of the probe. As in the examples described, cleavage of the probe must occur only when hybridized to the target sequence. The target sequence is then released from the duplex intact and repeatedly recycles through the reaction pathway. Formats are provided for use of the CHA method in both direct and competition assays. In general, a detection system based on a single CHA reaction provides a linear amplification of signal with time. By coupling two or more CHA reactions together in an appropriate manner, an exponential amplification of signal can be achieved, giving rise to an even greater level of sensitivity.

It therefore can be seen that the invention accomplishes all of the objectives heretofore stated.

Claims

1. A method of detecting nucleic acid sequences in which said sequences serve as a cofactor for a catalytic reaction in which a complementary labeled nucleic acid probe is cleaved and detected, said method comprising:

   hybridizing of a target sequence to a labeled nucleic acid probe to provide a probe: target sequence duplex;

   cleaving of the labeled probe within the probe:target sequence duplex to release the target sequence intact;

   recycling of the target sequence repeatedly through the reaction pathway; and

   detecting the extent of cleavage of the labeled probe.

2. A method as claimed in claim 1, wherein said detection method is used for the diagnosis of a disease or genetic trait in man or other species.

3. A method as claimed in claim 1, wherein said detection method is used to identify a nucleic acid sequence in an industrial or agricultural process.

4. A method as claimed in claim 1,2 or 3, wherein said nucleic acid probe is labeled with a radioactive, fluorescent, electrochemical, luminescent or enzyme marker, or other reporter group to permit its detection.

5. A method as claimed in any preceding claim, wherein the extent of cleavage of the probe is determined by electrophoretic separation of the smaller cleavage fragments carrying the labeling group from the remaining molecules of the intact probe.

6. A method as claimed in any one of claims 1 to 4, wherein the extent of cleavage of the probe is determined by using strong acids such as trichloroacetic acid to selectively precipitate the larger uncleaved molecules of the probe, leaving the smaller cleaved fragments bearing the label in solution.

7. A method as claimed in any preceding claim, wherein the probe is modified with both a high affinity label and a reporter group, and in which a solid support to which a corresponding affinity group is attached is used to adsorb molecules of the intact probe, leaving the cleaved fragments bearing the label in solution, said method comprising the additional steps of:

   adding of solid support material to the sample following the hybridizing and cleaving reaction;

   associating of the high affinity label attached to the probe with the corresponding affinity group on the solid support;

   separating of the solid support from the solution containing the cleaved fragments of the probe bearing the reporter group; and

   detection of the reporter groups in solution.

8. A method as claimed in claim 7, in which said solid support is composed of latex, polystyrene, cross-linked dextran or glass beads, cellulose, or a tetrafluoroethylene homoploymer or nylon membrane.

9. A method as claimed in claim 7 or 8 in which said solid support is magnetized and the solid support is separated from the bulk solution by the application of a magnetic field.

IDT Exhibit 1020.9762

10. A method as claimed in claim 7, 8 or 9, wherein said affinity label is a nucleic acid sequence not cleaved in the hybridization and cleavage reaction and the corresponding affinity group attached to the solid support is a complementary nucleic acid sequence.

11. A method as claimed in any preceding claim wherein said labeled probe is attached to a solid support during the hybridization and cleavage reaction and cleaved fragments bearing the reporter group are released into solution.

12. A method as claimed in any preceding claim for detection of a DNA sequence, in which cleavage of the labeled probe is mediated by an enzyme with RNaseH activity.

13. A method as claimed in claim 12 in which nonspecific cleavage of the labeled probe is suppressed by the use of single-stranded ribonuclease inhibitors.

14. A method as claimed in claim 12 or 13, wherein the labeled probe is cleaved by a heat stable form of RNaseH.

15. A method as claimed in claim 12 wherein the labeled probe is an oligonucleotide from 15 to 50 nucleotides in length.

16. A method as claimed in claim 15, wherein said oligonucleotide is composed entirely of RNA.

17. A method as claimed in claim 15, wherein said oligonucleotide contains one or more interspersed sequences that are not cleavable by RNaseH.

18. A method as claimed in claim 15, wherein an oligonucleotide composed of a DNA-RNA-DNA sequence is used as the probe.

19. A method as claimed in claim 18, wherein said DNA-RNA-DNA probe is used to detect a point mutation such as occurs in sickle cell disease.

20. A method as claimed in claim 12, wherein the labeled probe is a polynucleotide from 50 to several thousand nucleotides in length.

21. A method as claimed in claim 20, wherein said polynucleotide is composed entirely of RNA.

22. A method as claimed in claim 20, wherein said polynucleotide contains one or more interspersed sequences that are not cleavable by RNaseH.

23. A method as claimed in claim 1 for detection of an RNA sequence in which the sample RNA is first used as template for the synthesis of a complementary DNA copy of the target sequence with reverse transcriptase, said method comprising:

hybridizing of a DNA primer to the sample RNA 5' to the target sequence;

extending of the primer with reverse transcriptase across the target sequence to produce a complementary DNA copy of the target sequence; and detecting of the complementary DNA copy.

24. A method as claimed in claim 23, in which the hybridization and cleavage method is based on cleavage of a labeled probe by an enzyme with RNaseH activity.

25. A method as claimed in claim 23 or 24, in which a phosphorothioate derivative of the complementary DNA copy of the target sequence is synthesized with reverse transcriptase using deoxynucleotide α-thiotriphosphates as substrates for the reaction such that the phosphate groups of the DNA strand synthesized are phosphorotioates rather than the naturally occurring phosphodiesters.

26. A method as claimed in claim 25, in which the hybridization and cleavage method is based on cleavage of a labeled probe by an enzyme with RNaseH activity.

IDT Exhibit 1020.9763

27. A method as claimed in claim 25, in which the hybridization and cleavage method is based on cleavage of a labeled DNA probe by a restriction enzyme, said method comprising:
hybridizing of the phosphorothioate derivative synthesized to a complementary labeled DNA probe;
cleaving of the labeled probe within the probe:phosphorotioate DNA duplex by a restriction enzyme at its recognition site to release the phosphorothioate strand intact;
recycling of the phosphorothioate strand repeatedly through the reaction pathway; and
detecting of the extent of cleavage of the probe following the reaction.

28. A method as claimed in claim 1 for detecting of a DNA sequence in which a copy of the target sequence is first synthesized with DNA polymerase using the sample DNA as a template, said method comprising:
denaturing the sample DNA;
hybridizing of a DNA primer to the sample DNA 5' to the target sequence;
extension of the primer with DNA polymerase across the target sequence to produce a copy of the complementary DNA strand;
detecting of the DNA copy.

29. A method as claimed in claim 28, in which a phosphorothioate derivative of the copy of the target sequence is synthesized with DNA polymerase using deoxynucleotide α-thiotriphosphates as substrates for the reaction such that the phosphate groups of the DNA strand synthesized are phosphorothioates rather than the naturally occurring phosphodiesters.

30. A method as claimed in claim 29, in which the hybridization and cleavage method is based on cleavage of a labeled probe by an enzyme with RNaseH activity.

31. A method as claimed in claim 29, in which the hybridization and cleavage method is based on cleavage of a labeled DNA probe by a restriction enzyme, said method comprising:
hybridization of the phosphorothioate derivative synthesized to a complementary labeled DNA probe;
cleavage of the labeled probe within the probe:phosphorothioate DNA duplex by a restriction enzyme at its recognition site;
release of the phosphorothioate strand intact;
recycling of the phosphorothioate strand repeatedly through the reaction pathway; and
detection of the extent of cleavage of the probe following the reaction.

32. A method as claimed in claim 1, in which two or more hybridization and cleavage reactions are coupled to achieve an exponential amplification of signal.

33. A method for detection of an RNA sequence using a competition assay, said method comprising:
hybridization of the sample RNA to a DNA molecule complementary to the target sequence;
addition of a labeled RNA probe complementary to said DNA molecule plus an enzyme with RNaseH activity to the test sample;
incubation of the test sample to allow the hybridizing and cleaving reaction to proceed; and
measuring of the extent of cleavage of the RNA probe.

34. A method as claimed in claim 33, wherein said DNA molecule contains one segment complementary to both the target RNA sequence and the labeled probe and a second segment complementary to just the target sequence.

35. A method as claimed in claim 34, wherein the DNA molecule contains a sequence at least 4 nucleotides in length in which the phosphate groups are unmodified, and the remaining phosphate groups are modified so as to prevent cleavage of the complementary target and probe RNA sequences by RNaseH.

36. A method as claimed in claim 33, in which said DNA molecule is complementary to the labeled probe but forms a duplex with the target RNA sequence with one or more mismatches, such that cleavage of the target RNA sequence by RNaseH is inhibited.

IDT Exhibit 1020.9764

37. A probe useful for carrying out the method as claimed in any preceding claim, made of an oligonucleotide 15 to 50 nucleotides in length containing an RNA sequence and one or more additional sequences not cleavable by RNaseH.

38. A probe as claimed in claim 37, in which said RNA sequence is from 3 to 15 nucleotides in length, and in which the oligonucleotide is labeled at or near the 5'- or 3'-end of the chain with a radioactive, fluorescent, electrochemical, luminescent or enzyme marker, or other reporter group.

39. A probe as claimed in claim 38, in which the oligonucleotide is further modified with a high affinity group at or near the opposite end of the chain from which said reporter group is attached.

40. A probe useful for carrying out the method as claimed in any one of claims 1 to 35, made of a polynucleotide from 50 to several thousand nucleotides in length composed of RNA plus one or more interspersed sequences that are not cleavable by RNaseH.

41. A probe as claimed in claim 40, in which the stretches of RNA sequence are 3 nucleotides in length or greater, and in which one or more of the interspersed sequences not cleavable by RNaseH is labeled with a radioactive, fluorescent, electrochemical, luminescent or enzyme marker, or other reporter group.

Patentansprüche

1. Verfahren zum Nachweis von Nukleinsäuresequenzen, bei dem besagte Sequenzen als ein Cofaktor für eine katalytische Reaktion dienen, in der eine komplementäre markierte Nukleinsäuresonde gespalten und nachgewiesen wird, wobei besagtes Verfahren umfaßt:
   Hybridisieren einer Zielsequenz an eine markierte Nukleinsäuresonde, um eine Sonden/Zielsequenz-Duplex bereitzustellen;
   Spalten der markierten Sonde innerhalb der Sonden/Zielsequenz-Duplex, um die Zielsequenz intakt freizusetzen;
   wiederholtes Führen der Zielsequenz im Kreislauf durch den Reaktionsweg; und
   Nachweis des Umfangs der Spaltung der markierten Sonde.

2. Verfahren wie beansprucht in Anspruch 1, wobei besagtes Nachweisverfahren für die Diagnose einer Erkrankung oder eines genetischen Merkmals bei Menschen oder anderen Spezies verwendet wird.

3. Verfahren wie beansprucht in Anspruch 1, wobei besagtes Nachweisverfahren verwendet wird, um eine Nukleinsäuresequenz in einem industriellen oder landwirtschaftlichen Verfahren zu identifizieren.

4. Verfahren wie beansprucht in Anspruch 1, 2 oder 3, wobei besagte Nukleinsäuresonde mit einer anderen radioaktiven, fluoreszierenden, elektrochemischen, lumineszierenden oder Enzym-Markierung oder einer Reporter-Gruppe markiert wird, um ihren Nachweis zu ermöglichen.

5. Verfahren wie beansprucht in irgendeinem vorangehenden Anspruch, wobei der Umfang der Spaltung der Sonde durch elektrophoretische Trennung der kleineren Spaltungsfragmente, die die Markierungsgruppe tragen, von den restlichen Molekülen der intakten Sonde bestimmt wird.

6. Verfahren wie beansprucht in einem der Ansprüche 1 bis 4, wobei der Umfang der Spaltung der Sonde bestimmt wird, indem starke Säuren, wie etwa Trichloressigsäure, verwendet werden, um die größeren, nicht-gespaltenen Moleküle der Sonde selektiv auszufällen, was die kleineren gespaltenen Fragmente, die die Markierung tragen, in Lösung zurückläßt.

7. Verfahren wie beansprucht in irgendeinem vorangehenden Anspruch, wobei die Sonde sowohl mit einer Markierung mit hoher Affinität als auch mit einer Reporter-Gruppe modifiziert wird und bei dem ein fester Träger, an den eine entsprechende Affinitätsgruppe gebunden ist, verwendet wird, um Moleküle der intakten Sonde zu adsorbieren, was die gespaltenen Fragmente, die die Markierung tragen, in Lösung zurückläßt, wobei besagtes Verfahren die zusätzlichen Schritte umfaßt:
   Zugeben eines festen Trägermaterials zur Probe im Anschluß an die Hybridisierungs- und Spaltungsreaktion;
   Assoziieren der Markierung mit hoher Affinität, die an die Sonde mit der entsprechenden Affinitätsgrup-

IDT Exhibit 1020.9765

pe auf dem festen Träger gebunden ist;
Trennen des festen Trägers von der Lösung, die die gespaltenen Fragmente der Sonde enthält, die die Reporter-Gruppe tragen; und
Nachweis der Reporter-Gruppen in Lösung.

8. Verfahren wie beansprucht in Anspruch 7, bei dem besagter fester Träger aus Latex, Polystyrol, vernetztem Dextran oder Glasperlen, Cellulose oder einer Tetrafluorethylen-Homopolymer- oder Nylonmembran besteht.

9. Verfahren wie beansprucht in Anspruch 7 oder 8, bei dem besagter fester Träger magnetisiert ist und der feste Träger von der Hauptlösung durch das Anlegen eines Magnetfeldes getrennt wird.

10. Verfahren wie beansprucht nach Anspruch 7, 8 oder 9, wobei besagte Affinitätsmarkierung eine Nukleinsäuresequenz ist, die in der Hybridisierungs- und Spaltungsreaktion nicht gespalten wird, und die entsprechende Affinitätsgruppe, die an den festen Träger gebunden ist, eine komplementäre Nukleinsäuresequenz ist.

11. Verfahren wie beansprucht in irgendeinem vorangehenden Anspruch, wobei besagte markierte Sonde während der Hybridisierungs- und Spaltungsreaktion an einen festen Träger gebunden wird und gespaltene Fragmente, die die Reporter-Gruppe tragen, in Lösung hinein freigesetzt werden.

12. Verfahren wie beansprucht in irgendeinem vorangehenden Anspruch, zum Nachweis einer DNA-Sequenz, bei dem Spaltung der markierten Sonde indirekt durch ein Enzym mit RNaseH-Aktivität bewirkt wird.

13. Verfahren wie beansprucht in Anspruch 12, bei dem nicht-spezifische Spaltung der markierten Sonde durch die Verwendung einzelsträngiger Ribonuklease-Inhibitoren unterdrückt wird.

14. Verfahren wie beansprucht in Anspruch 12 oder 13, wobei die markierte Sonde durch eine wärmestabile Form von RNaseH gespalten wird.

15. Verfahren wie beansprucht in Anspruch 12, wobei die markierte Sonde ein mit von 15 bis 50 Nukleotide langes Oligonukleotid ist.

16. Verfahren wie beansprucht in Anspruch 15, wobei besagtes Oligonukleotid vollständig aus RNA besteht.

17. Verfahren wie beansprucht in Anspruch 15, wobei besagtes Oligonukleotid eine oder mehrere eingestreute Sequenzen enthält, die nicht durch RNaseH spaltbar sind.

18. Verfahren wie beansprucht in Anspruch 15, wobei ein Oligonukleotid, das aus einer DNA-RNA-DNA-Sequenz besteht, als die Sonde verwendet wird.

19. Verfahren wie beansprucht in Anspruch 18, wobei besagte DNA-RNA-DNA-Sonde verwendet wird, um eine Punktmutation nachzuweisen, wie sie bei Sichelzellenanämie auftritt.

20. Verfahren wie beansprucht in Anspruch 12, wobei die markierte Sonde ein von 5 bis mehrere Tausend Nukleotiden langes Polynukleotid ist.

21. Verfahren wie beansprucht in Anspruch 20, wobei besagtes Polynukleotid vollständig aus RNA besteht.

22. Verfahren wie beansprucht in Anspruch 20, wobei besagtes Polynukleotid eine oder mehrere eingestreute Sequenzen enthält, die nicht von RNaseH spaltbar sind.

23. Verfahren wie beansprucht in Anspruch 1, zum Nachweis einer RNA-Sequenz, bei dem die Proben-RNA zunächst als Matrize für die Synthese einer komplementären DNA-Kopie der Zielsequenz mit Reverser Transkriptase verwendet wird, wobei besagtes Verfahren umfaßt:
Hybridisieren eines DNA-Primers an die Proben-RNA 5' zur Zielsequenz;

IDT Exhibit 1020.9766

Extendieren des Primers mit Reverser Transkriptase über die Zielsequenz, um eine komplementäre DNA-Kopie der Zielsequenz herzustellen; und
Nachweis der komplementären DNA-Kopie.

24. Verfahren wie beanspruCht in Anspruch 23, bei dem das Hybridisierungs- und Spaltungsverfahren auf der Spaltung einer markierten Probe durch ein Enzym mit RNaseH-Aktivität beruht.

25. Verfahren wie beansprucht in Anspruch 23 oder 24, bei dem ein Phosphorothioat-Derivat der komplementären DNA-Kopie der Zielsequenz mit Reverser Transkriptase unter Verwendung von Desoxynukleotid-α-Thiotriphosphaten als Substraten für die Reaktion synthetisiert wird, so daß die Phosphatgruppen des synthetisierten DNA-Stranges Phosphorothioate statt der natürlich auftretenden Phosphodiester sind.

26. Verfahren wie beansprucht in Anspruch 25, bei dem das Hybridisierungs- und Spaltungsverfahren auf der Spaltung einer markierten Sonde durch ein Enzym mit RNaseH-Aktivität beruht.

27. Verfahren wie beansprucht in Anspruch 25, bei dem das Hybridisierungs- und Spaltungsverfahren auf der Spaltung einer markierten DNA-Sonde durch ein Restriktionsenzym beruht, wobei besagtes Verfahren umfaßt:
Hybridisieren des synthetisierten Phosphorothioat-Derivats an eine komplementäre markierte DNA-Sonde;
Spalten der markierten Sonde innerhalb der Sonden/Phosphorothioat-DNA-Duplex durch ein Restriktionsenzym an seiner Erkennungsstelle, um den Phosphorothioat-Strang intakt freizusetzen;
wiederholtes Führen des Phosphorothioat-Stranges im Kreislauf durch den Reaktionsweg; und
Nachweis des Umfangs der Spaltung der Sonde im Anschluß an die Reaktion.

28. Verfahren wie beansprucht in Anspruch 1, zum Nachweis einer DNA-Sequenz, bei dem eine Kopie der Zielsequenz zunächst mit DNA-Polymerase unter Verwendung der Proben-DNA als einer Matrize synthetisiert wird, wobei besagtes Verfahren umfaßt:
Denaturieren der Proben-DNA;
Hybridisieren eines DNA-Primers an die Proben-DNA 5' zur Zielsequenz;
Extension des Primers mit DNA-Polymerase über die Zielsequenz, um eine Kopie des komplementären DNA-Stranges herzustellen;
Nachweis der DNA-Kopie.

29. Verfahren wie beansprucht in Anspruch 28, bei dem ein Phosphorothioat-Derivat der Kopie der Zielsequenz mit DNA-Polymerase unter Verwendung von Desoxynukleotid-α-Thiotriphosphaten als Substraten für die Reaktion synthetisiert wird, so daß die Phosphatgruppen des synthetisierten DNA-Stranges Phosphorothioate statt der natürlich auftretenden Phosphodiester sind.

30. Verfahren wie beansprucht in Anspruch 29, bei dem das Hybridisierungs- und Spaltungsverfahren auf der Spaltung einer markierten Sonde durch ein Enzym mit RNaseH-Aktivität beruht.

31. Verfahren wie beansprucht in Anspruch 29, bei dem das Hybridisierungs- und Spaltungsverfahren auf der Spaltung einer markierten DNA-Sonde durch ein Restriktionsenzym beruht, wobei besagtes Verfahren unfaßt:
Hybridisierung des synthetisierten Phosphorothioat-Derivats an eine komplementäre markierte DNA-Sonde;
Spaltung der markierten Sonde innerhalb der Sonden/Phosphorothioat-DNA-Duplex durch ein Restriktionsenzym an seiner Erkennungsstelle;
Freisetzung des intakten Phosphorothioat-Stranges;
wiederholtes Führen des Phosphorothioat-Stranges im Kreislauf durch den Reaktionsweg; und
Nachweis des Umfangs der Spaltung der Sonde im Anschluß an die Reaktion.

32. Verfahren wie beansprucht in Anspruch 1, bei dem zwei oder mehr Hybridisierungs- oder Spaltungsreaktionen gekoppelt werden, um eine exponentielle Signalverstärkung zu erreichen.

IDT Exhibit 1020.9767

33. Verfahren zum Nachweis einer RNA-Sequenz unter Verwendung eines Kompetitions-Tests, wobei besagtes Verfahren umfaßt:
Hybridisierung der Proben-RNA an ein zur Zielsequenz komplementäres DNA-Molekül;
Zugabe einer zu besagtem DNA-Molekül komplementären, markierten RNA-Sonde plus eines Enzyms mit RNaseH-Aktivität zur Testprobe;
Inkubation der Testprobe, um das Fortschreiten der Hybridisierungs- und Spaltungsreaktion zu ermöglichen; und
Messen des Umfangs der Spaltung der RNA-Sonde.

34. Verfahren wie beansprucht in Anspruch 33, wobei besagtes DNA-Molekül ein zu sowohl der Ziel-RNA-Sequenz als auch der markierten Sonde komplementäres Segment und ein zweites, nur zur Zielsequenz komplementäres Segment enthält.

35. Verfahren wie beansprucht in Anspruch 34, wobei das DNA-Molekül eine wenigstens 4 Nukleotide lange Sequenz enthält, in der die Phosphatgruppen nicht-modifiziert sind, und die restlichen Phosphatgruppen modifiziert sind, um Spaltung der komplementären Ziel- und Sonden-RNA-Sequenzen durch die RNaseH zu verhindern.

36. Verfahren wie beansprucht in Anspruch 33, bei dem besagtes DNA-Molekül zu der markierte Sonde komplementär ist, aber mit der Ziel-RNA-Sequenz eine Duplex mit einem oder mehreren Mismatches bildet, so daß Spaltung der Ziel-RNA-Sequenz durch RNaseH gehemmt ist.

37. Zur Durchführung des Verfahrens, wie beansprucht in irgendeinem vorangehenden Anspruch, brauchbare Sonde, hergestellt aus einem 15 bis 50 Nukleotide langen Oligonukleotid, das eine RNA-Sequenz und eine oder mehrere zusätzliche Sequenzen enthält, die nicht von RNaseH spaltbar sind.

38. Sonde wie beansprucht in Anspruch 37, in der besagte DNA-Sequenz von 3 bis 15 Nukleotide lang ist und in der das Nukleotid an oder nahe dem 5'- oder 3'-Ende der Kette mit einer radioaktiven, fluoreszierenden, elektrochemischen, lumineszierenden oder Enzym-Markierung oder einer anderen Reporter-Gruppe markiert ist.

39. Sonde wie beansprucht in Anspruch 38, in der das Oligonukleotid außerdem mit einer Gruppe mit hoher Affinität an oder nahe dem entgegengesetzten Ende der Kette zu demjenigen an das besagte Reporter-Gruppe gebunden ist, modifiziert ist.

40. Zur Durchführung des Verfahrens, wie in einem der Ansprüche 1 bis 35 beansprucht, brauchbare Sonde, hergestellt aus einem von 50 bis mehrere Tausend Nukleotide langen Polynukleotid, das aus RNA plus einer oder mehreren eingestreuten Sequenzen besteht, die nicht von RNaseH spaltbar sind.

41. Sonde wie beansprucht in Anspruch 40, in der die Strecken der RNA-Sequenz 3 Nukleotide lang oder länger sind und in der eine oder mehrere der eingestreuten Sequenzen, die nicht durch RNaseH spaltbar sind, mit einer radioaktiven, fluoreszierenden, elektrochemischen, lumineszierenden oder Enzym-Markierung oder einer anderen Reporter-Gruppe markiert sind.

Revendications

1. Méthode de détection de séquences d'acides nucléiques, dans laquelle lesdites séquences servent de cofacteur pour une réaction catalytique dans laquelle une sonde d'acide nucléique marquée complémentaire est clivée et détectée, ladite méthode comprenant :
l'hybridation d'une séquence-cible à une sonde d'acide nucléique marquée pour produire un duplex sonde:séquence-cible ;
le clivage de la sonde marquée à l'intérieur du duplex sonde:séquence-cible pour libérer la séquence-cible intacte ;
le recyclage de la séquence-cible de manière répétée dans le processus réactionnel ; et
la détection du degré de clivage de la sonde marquée.

2. Méthode suivant la revendication 1, ladite méthode de détection étant utilisée pour le diagnostic d'une maladie ou d'un caractère génétique chez l'homme ou d'autres espèces.

IDT Exhibit 1020.9768

3. Procédé suivant la revendication 1, ladite méthode de détection étant utilisée pour l'identification d'une séquence d'acide nucléique dans un processus industriel ou agricole.

4. Procédé suivant la revendication 1, 2 ou 3, dans lequel la sonde d'acide nucléique est marquée avec un marqueur radio-actif, fluorescent, électrochimique, luminescent ou enzymatique, ou bien avec un autre groupe reporter pour permettre sa détection.

5. Méthode suivant l'une quelconque des revendications précédentes, dans laquelle le degré de clivage de la sonde est déterminé par séparation électrophorétique des fragments de clivage plus petits portant le groupe de marquage des molécules restantes de la sonde intacte.

6. Méthode suivant l'une quelconque des revendications 1 à 4, dans laquelle le degré de clivage de la sonde est déterminé en utilisant des acides forts tels que l'acide trichloracétique pour précipiter sélectivement les molécules non clivées plus volumineuses de la sonde, en laissant les fragments clivés plus petits portant le marqueur en solution.

7. Méthode suivant l'une quelconque des revendications précédentes, dans laquelle la sonde est modifiée à la fois avec un marqueur à forte affinité et un groupe reporter, et dans laquelle un support solide auquel est fixé un groupe doué d'affinité correspondant est utilisé pour adsorber des molécules de la sonde intacte, en laissant les fragments clivés portant le marqueur en solution, ladite méthode comprenant les étapes supplémentaires :

d'addition d'une matière solide de support à l'échantillon après la réaction d'hybridation et de clivage ;

d'association du marqueur à forte affinité fixé à la sonde avec le groupe doué d'affinité correspondant sur le support solide ;

de séparation du support solide de la solution contenant les fragments clivés de la sonde portant le groupe reporter ; et

de détection des groupes reporter en solution.

8. Méthode suivant la revendication 7, dans laquelle le support solide est constitué d'un latex, d'un polystyrène, d'un dextrane réticulé ou de billes de verre, de cellulose, d'un homopolymère de tétrafluoréthylène ou d'une membrane de Nylon.

9. Méthode suivant la revendication 7 ou 8, dans laquelle le support solide est aimanté et ledit support solide est séparé de la masse de solution par application d'un champ magnétique.

10. Méthode suivant la revendication 7, 8 ou 9, dans lequel le marqueur doué d'affinité est une séquence d'acide nucléique non clivé dans la réaction d'hybridation et de clivage et le groupe doué d'affinité correspondant fixé au support solide est une séquence d'acide nucléique complémentaire.

11. Méthode suivant l'une quelconque des revendications précédentes, dans laquelle la sonde marquée est fixée à un support solide au cours de la réaction d'hybridation et de clivage et les fragments clivés portant le groupe reporter sont libérés dans la solution.

12. Méthode suivant l'une quelconque des revendications précédentes pour la détection d'une séquence d'ADN, dans laquelle le clivage de la sonde marquée s'effectue par médiation avec un enzyme à activité de RNaseH.

13. Méthode suivant la revendication 12, dans laquelle le clivage non spécifique de la sonde marquée est supprimé par l'utilisation d'inhibiteurs de ribonucléase à action sur les brins simples.

14. Méthode suivant la revendication 12 ou 13, dans laquelle la sonde marquée est clivée par une forme thermostable de RNaseH.

15. Méthode suivant la revendication 12, dans laquelle la sonde marquée est un oligonucléotide ayant une longueur de 15 à 50 nucléotides.

16. Méthode suivant la revendication 15, dans laquelle l'oligonucléotide est constitué totalement d'ARN.

IDT Exhibit 1020.9769

17. Méthode suivant la revendication 15, dans laquelle l'oligonucléotide contient une ou plusieurs séquences intercalées qui ne sont pas clivables par la RNaseH.

18. Méthode suivant la revendication 15, dans laquelle un oligonucléotide constitué d'une séquence ADN-ARN-ADN est utilisé comme sonde.

19. Méthode suivant la revendication 18, dans laquelle la sonde ADN-ARN-ADN est utilisée pour détecter une mutation ponctuelle telle qu'une mutation se produisant dans la maladie à hématies falciformes.

20. Méthode suivant la revendication 12, dans laquelle la sonde marquée est un polynucléotide ayant une longueur de 50 à plusieurs milliers de nucléotides.

21. Méthode suivant la revendication 20, dans laquelle le polynucléotide est constitué totalement d'ARN.

22. Méthode suivant la revendication 20, dans laquelle le polynucléotide contient une ou plusieurs séquences intercalées qui ne sont pas clivables par la RNaseH.

23. Méthode suivant la revendication 1 pour la détection d'une séquence d'ARN, dans laquelle l'ARN de l'échantillon est utilisé comme matrice pour la synthèse d'une copie d'ADN complémentaire de la séquence-cible avec la transcriptase inverse, ladite méthode comprenant :

l'hybridation d'une amorce d'ADN à l'ARN de l'échantillon en 5' de la séquence-cible ;

l'allongement de l'amorce avec la transcriptase inverse à travers la séquence-cible pour produire une copie d'ADN complémentaire de la séquence-cible ; et

la détection de la copie d'ADN complémentaire.

24. Méthode suivant la revendication 23, dans laquelle le procédé d'hybridation et de clivage est basé sur le clivage d'une sonde marquée avec un enzyme à activité de RNAseH.

25. Méthode suivant la revendication 23 ou 24, dans laquelle un dérivé de phosphorothioate de la copie d'ADN complémentaire de la séquence-cible est synthétisé avec la transcriptase inverse en utilisant des désoxynucléotide-α-thiotriphosphates comme substrats pour la réaction, de telle sorte que les groupes phosphate du brin d'ADN synthétisé soient des phosphorothioates au lieu des phosphodiesters naturels.

26. Méthode suivant la revendication 25, dans laquelle le procédé d'hybridation et de clivage est basé sur le clivage d'une sonde marquée avec un enzyme à activité de RNaseH.

27. Méthode suivant la revendication 25, dans laquelle le procédé d'hybridation et de clivage est basé sur le clivage d'une sonde d'ADN marquée avec un enzyme de restriction, ladite méthode comprenant :

l'hybridation du dérivé de phosphorothioate synthétisé à une sonde d'ADN marquée complémentaire :

le clivage de la sonde marquée à l'intérieur du duplex d'ADN sonde:phosphorothioate avec un enzyme de restriction au niveau de site de reconnaissance pour libérer le brin phosphorothioate intact :

le recyclage du brin phosphorothioate de manière répétée dans le processus réactionnel : et

la détection du degré de clivage de la sonde après réaction.

28. Méthode suivant la revendication 1, pour la détection d'une séquence d'ADN, dans laquelle une copie de la séquence-cible est tout d'abord synthétisée avec l'ADN-polymérase au moyen de l'ADN d'échantillon comme matrice, ladite méthode comprenant :

la dénaturation de l'ADN de l'échantillon ;

l'hybridation d'une amorce d'ADN à l'ADN d'échantillon en 5' de la séquence-cible ;

l'allongement de l'amorce avec l'ADN-polymérase à travers la séquence-cible pour produire une copie du brin d'ADN complémentaire ;

la détection de la copie d'ADN.

29. Méthode suivant la revendication 28, dans laquelle un dérivé de phosphorothioate de la copie de la séquence-cible est synthétisé avec l'ADN-polymérase au moyen de désoxynucléotide-α-thiotriphosphates comme substrats pour la réaction de telle sorte que les groupes phosphate du brin d'ADN

IDT Exhibit 1020.9770

synthétisé soient des phosphorothioates au lieu des phosphodiesters naturels.

30. Méthode suivant la revendication 29, dans laquelle le procédé d'hybridation et de clivage est basé sur le clivage d'une sonde marquée avec un enzyme a activité de RNaseH.

31. Méthode suivant la revendication 29, dans laquelle le procédé d'hybridation et de clivage est basé sur le clivage d'une sonde d'ADN marquée avec un enzyme de restriction, ladite méthode comprenant :

l'hybridation du dérivé de phosphorothioate synthétisé à une sonde d'ADN marquée complémentaire ;

le clivage de la sonde marquée à l'intérieur du duplex d'ADN sonde:phosphorothioate avec un enzyme de restriction au niveau de son site de reconnaissance ;

la libération du brin de phosphorothioate intact ;

le recyclage du brin de phosphorothioate de manière répétée dans le processus réactionnel ; et

la détection du degré de clivage de la sonde après réaction.

32. Méthode suivant la revendication 1, dans laquelle deux ou plus de deux réactions d'hybridation et de clivage sont couplées pour parvenir à une amplification exponentielle du signal.

33. Méthode de détection d'une séquence d'ARN au moyen d'une analyse compétitive, ladite méthode comprenant :

l'hybridation de l'ARN de l'échantillon à une molécule d'ADN complémentaire de la séquence-cible ;

l'addition d'une sonde d'ARN marquée complémentaire de ladite molécule d'ADN plus un enzyme à activité de RNaseH à l'échantillon d'essai ;

la mise en incubation de l'échantillon d'essai pour permettre à la réaction d'hybridation et de clivage de s'effectuer ; et

la mesure du degré de clivage de la sonde d'ARN.

34. Méthode suivant la revendication 33, dans laquelle la molécule d'ADN contient un segment complémentaire à la fois de la séquence d'ARN cible et de la sonde marquée et un second segment complémentaire seulement de la séquence-cible.

35. Méthode suivant la revendication 34, dans laquelle la molécule d'ADN comprend une séquence ayant une longueur d'au moins 4 nucléotides dans laquelle les groupes phosphate sont non modifiés, et les groupes phosphate restants sont modifiés de manière à empêcher le clivage des séquences d'ARN de cible et de sonde complémentaires par la RNaseH.

36. Méthode suivant la revendication 33, dans laquelle la molécule d'ADN est complémentaire de la sonde marquée mais forme un duplex avec la séquence d'ARN cible présentant un ou plusieurs mésappariements, de telle sorte que le clivage de la séquence d'ARN cible par la RNaseH soit inhibé.

37. Sonde utile pour la mise en oeuvre de la méthode suivant l'une quelconque des revendications précédentes, constituée d'un oligonucléotide ayant une longueur de 15 à 50 nucléotides, contenant une séquence d'ARN et une ou plusieurs séquences supplémentaires non clivables par la RNaseH.

38. Sonde suivant la revendication 37, dans laquelle la séquence d'ARN possède une longueur de 3 à 15 nucléotides, et dans laquelle l'oligonucléotide est marqué au niveau ou à proximité de l'extrémité 5' ou 3' de la chaîne avec un marqueur radio-actif, fluorescent, électrochimique, luminescent ou enzymatique, ou un autre groupe reporter.

39. Sonde suivant la revendication 38, dans laquelle l'oligonucléotide est en outre modifié avec un groupe à forte affinité au niveau ou à proximité de l'extrémité de la chaîne opposée à celle auquel est fixé le groupe reporter.

40. Sonde utile pour la mise en oeuvre de la méthode suivant l'une quelconque des revendications 1 à 35, constituée d'un polynucléotide de 50 à plusieurs milliers de nucléotides de longueur, constitué d'ARN plus une ou plusieurs séquences intercalées qui ne sont pas clivables par la RNaseH.

IDT Exhibit 1020.9771

41. Sonde suivant la revendication 40, dans laquelle les brins de séquence d'ARN possèdent une longueur égale ou supérieure à 3 nucléotides, et dans laquelle une ou plusieurs des séquences intercalées non clivables par la RNaseH sont marquées avec un marqueur radio-actif, fluorescent, électrochimique, luminescent ou enzymatique, ou un autre groupe reporter.

IDT Exhibit 1020.9772

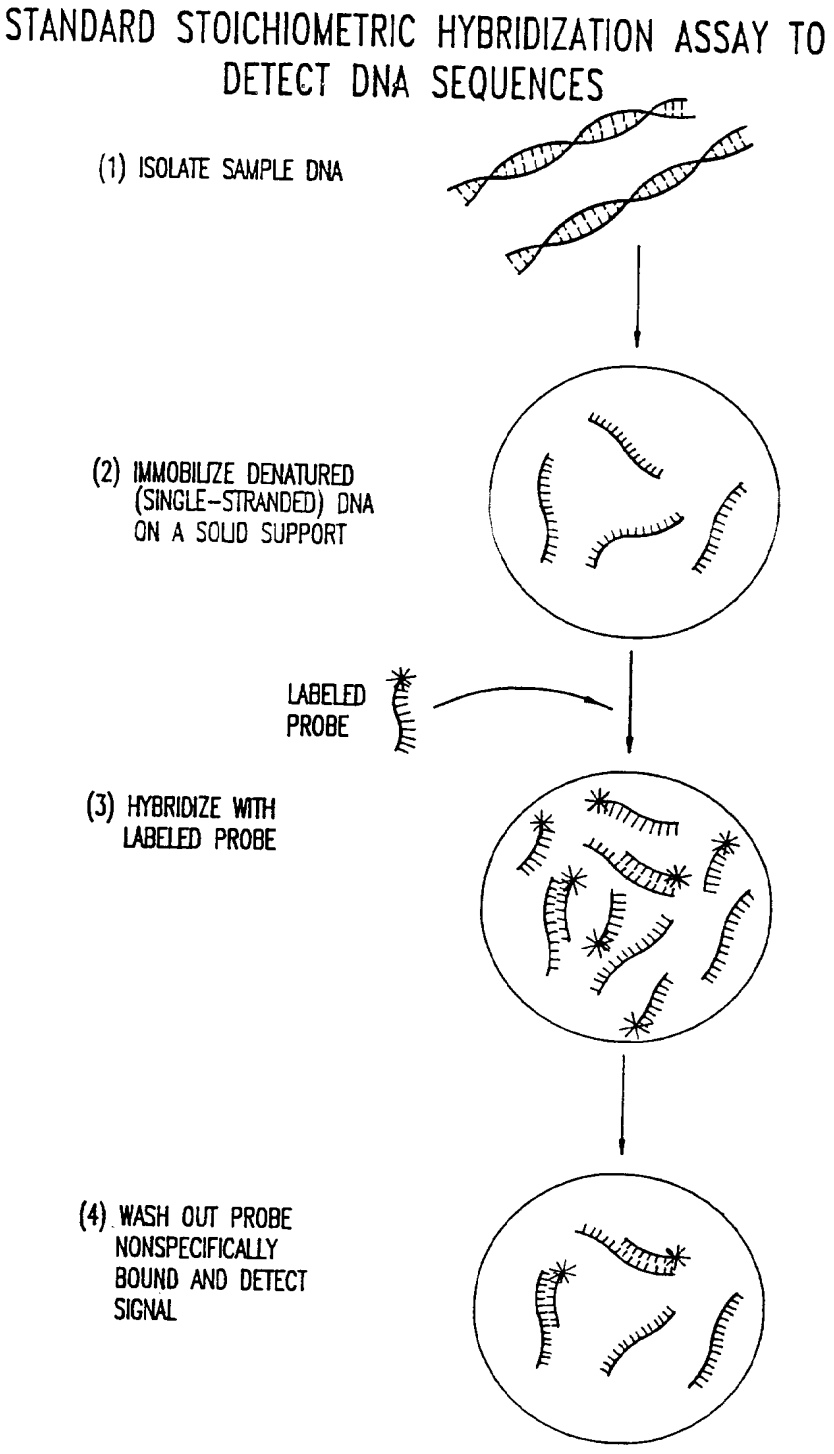


FIG.1

IDT Exhibit 1020.9773

DETECTION OF A SPECIFIC DNA SEQUENCE USING THE CHA METHOD

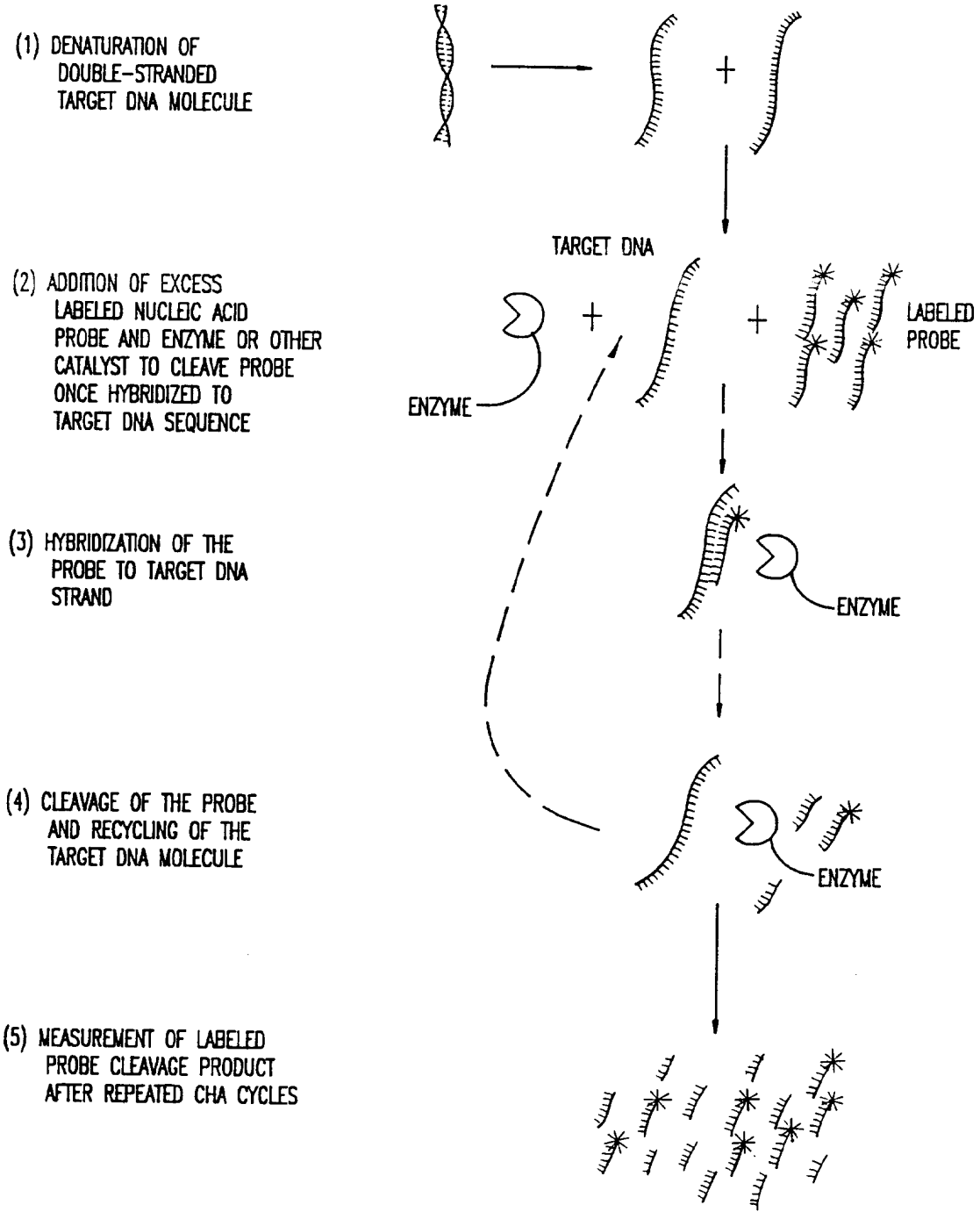


FIG.2

IDT Exhibit 1020.9774

CATALYTIC HYBRIDIZATION AMPLIFICATION IN SOLID PHASE FORMAT

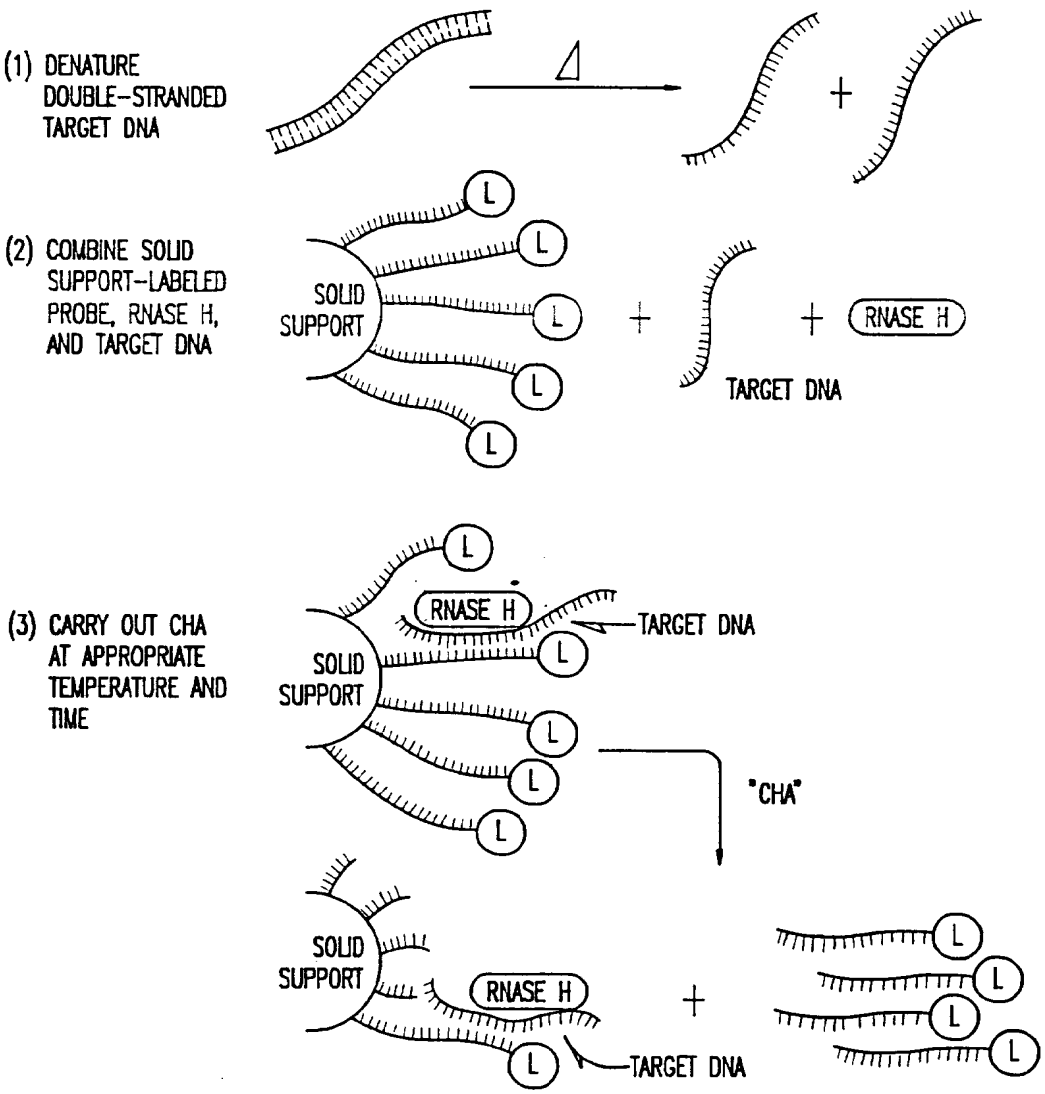
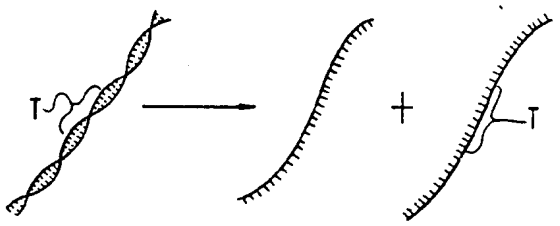


FIG.3

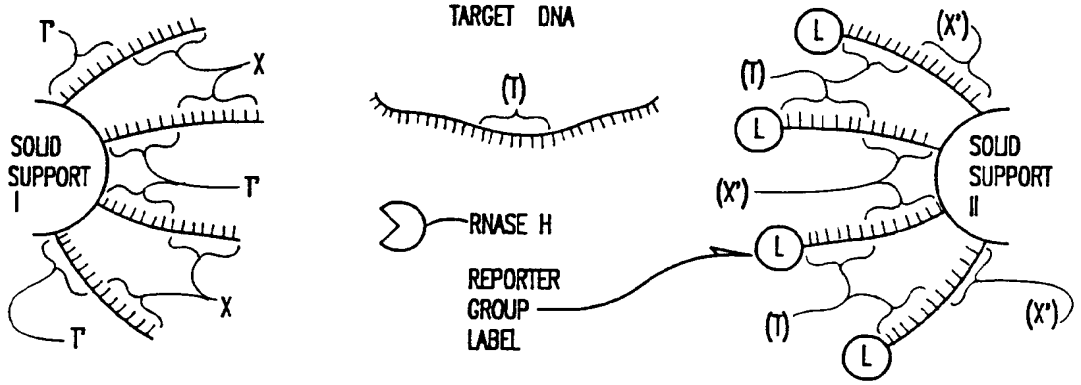
IDT Exhibit 1020.9775

EXPONENTIAL CATALYTIC HYBRIDIZATION AMPLIFICATION SYSTEM

(1) DENATURATION OF SAMPLE DNA TO FORM SINGLE-STRANDED SPECIES.
UNIQUE TARGET SEQUENCE IS REPRESENTED BY (T).



(2) ADDITION OF TARGET DNA AND RNASE H TO SOLID SUPPORT SYSTEM I AND II.
SOLID SUPPORT I CONTAINS PROBE WITH COMPLEMENTARY CLEAVABLE (T') AND (X) SEQUENCE.
SOLID SUPPORT II CONTAINS COMPLEMENTARY CLEAVABLE (X') AND LABELED (T) SEQUENCE.



(3) TARGET DNA SEQUENCE (T) HYBRIDIZES TO COMPLEMENTARY CLEAVABLE (T') ON SOLID SUPPORT I.

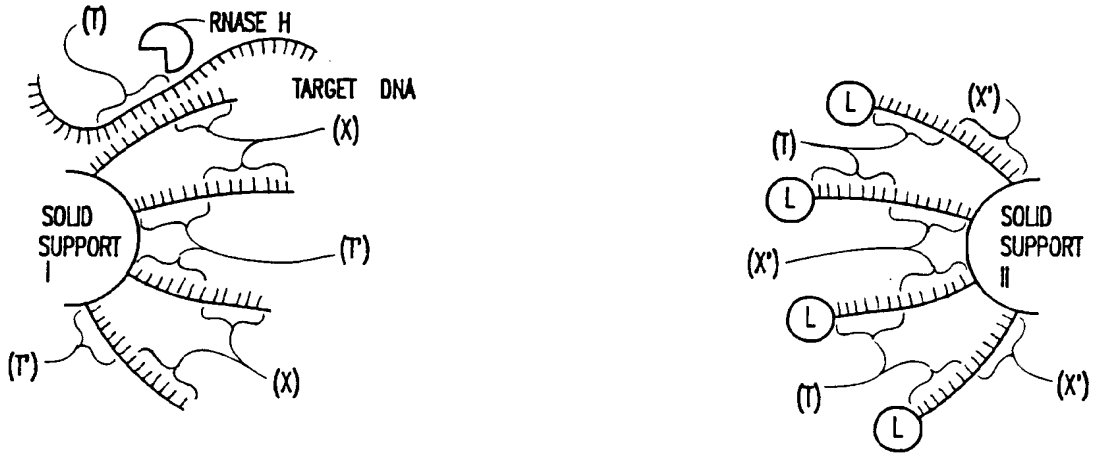
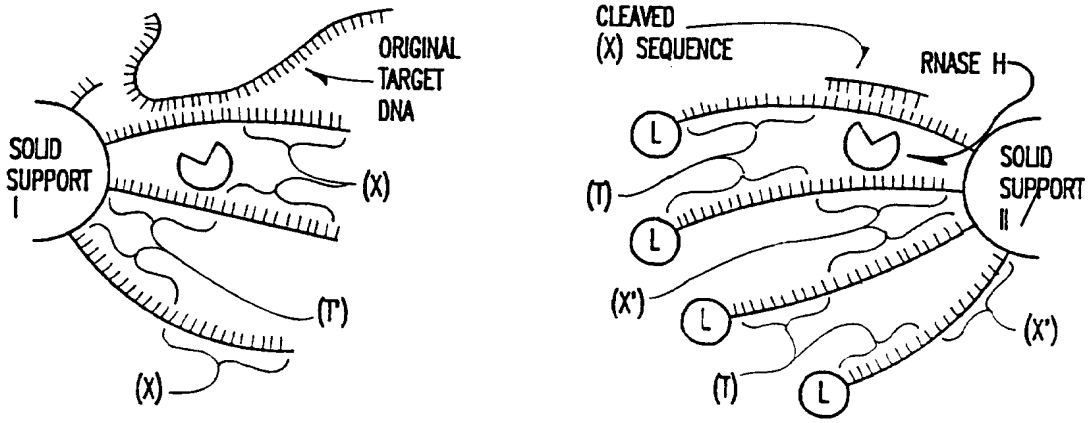


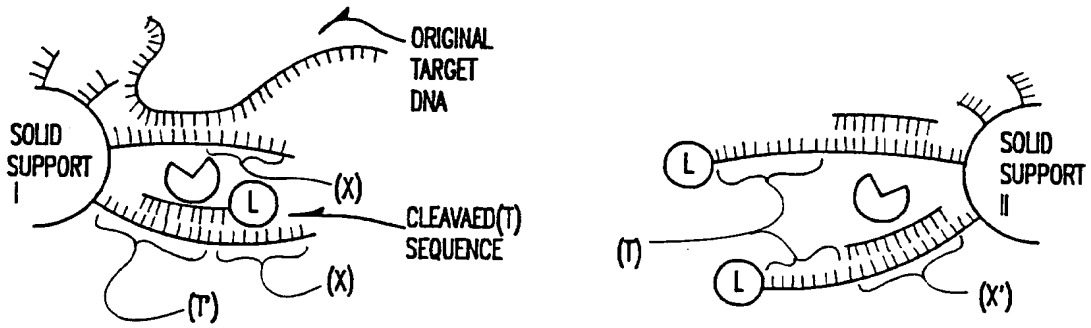
FIG.4A

IDT Exhibit 1020.9776

(4) CLEAVAGE OF (T') RELEASES SEQUENCE (X) WHICH DIFFUSES TO SOLID SUPPORT II, HYBRIDIZES TO COMPLEMENTARY CLEAVABLE SEQUENCE (X') AND RELEASES A LABELED (T) SEQUENCE. ORIGINAL TARGET SEQUENCE (T) CONTINUES TO RELEASE MORE (X) SEQUENCES.



(5) LABELED (T) SEQUENCE DIFFUSES BACK TO SOLID SUPPORT I, HYBRIDIZES WITH CLEAVABLE (T'), AND RELEASES MORE (X) SEQUENCES.



(6) CHA PROCESS CONTINUES BETWEEN THE TWO SOLID SUPPORT SYSTEMS WITH THE EXPONENTIAL RELEASE FROM SOLID SUPPORT II OF LABELED PROBES FOR SUBSEQUENT DETECTION.

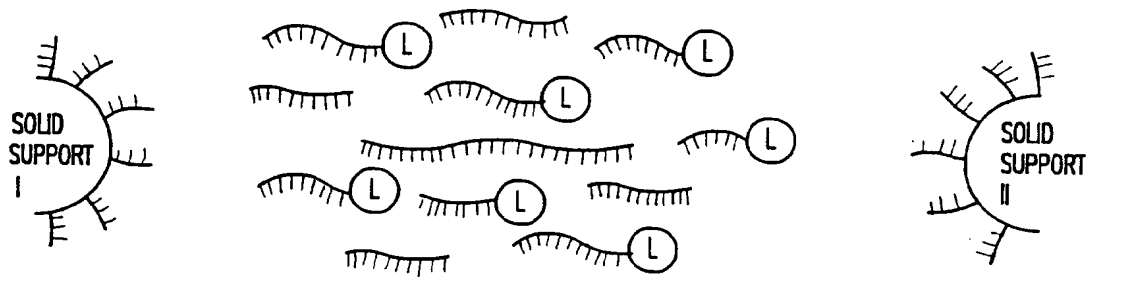


FIG.4B

IDT Exhibit 1020.9777

DETECTION OF SPECIFIC RNA SEQUENCES USING A COMPETITIVE CHA ASSAY

ADD COMPLEMENTARY DNA MOLECULE TO TEST SAMPLE OF RNA, HYBRIDIZE THEN ADD LABELED RNA PROBE AND RNASE H:

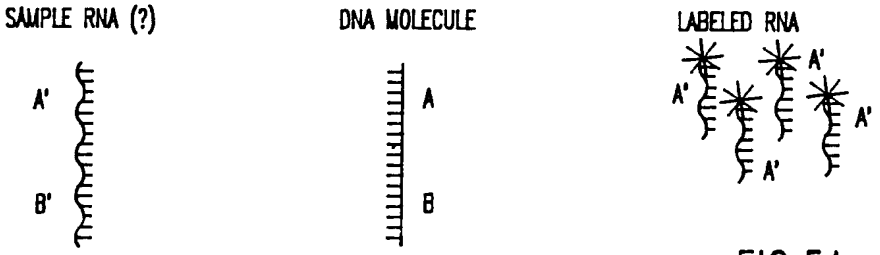
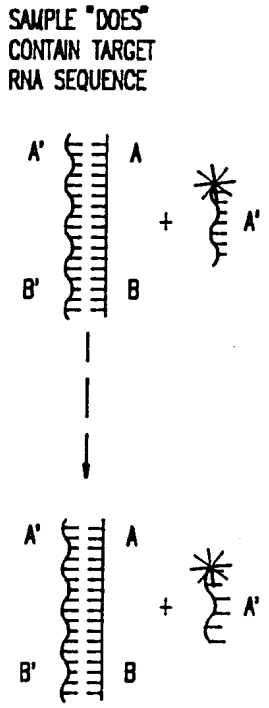
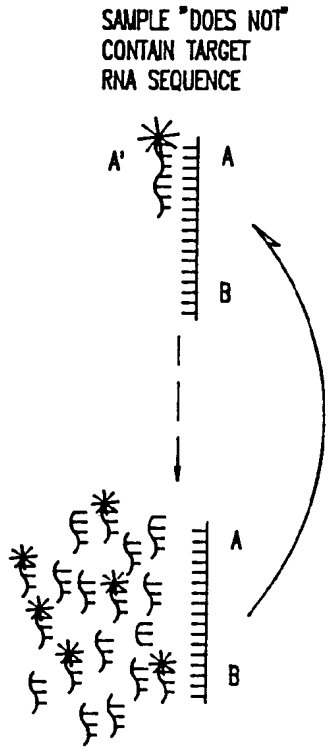


FIG.5A



NO LABELED FRAGMENTS RELEASED

FIG.5B



LABELED FRAGMENTS RELEASED

FIG.5C

IDT Exhibit 1020.9778

Espacenet

Bibliographic data: WO03106642 (A2) — 2003-12-24

COMPLEXITY MANAGEMENT OF GENOMIC DNA BY LOCUS SPECIFIC AMPLICATION

Inventor(s): JONES KEITH W [US]; SHAPERO MICHAEL [US]; LIU WEIWEI [US] ± (JONES, KEITH, W, ; SHAPERO, MICHAEL, ; LIU, WEIWEI)

Applicant(s): AFFYMETRIX INC [US]; JONES KEITH W [US]; SHAPERO MICHAEL [US]; LIU WEIWEI [US] ± (AFFYMETRIX, INC, ; JONES, KEITH, W, ; SHAPERO, MICHAEL, ; LIU, WEIWEI)

Classification: - **international:** ***C12Q1/68;*** (IPC1-7): C12N
- **cooperative:** **C12Q1/6809**; **C12Q1/6837**; **C12Q1/6858**; **C12Q1/686**; **C12Q1/6876**; C12Q2600/156 more

Application number: WO2003US18853 20030616 Global Dossier

Priority number (s): US20020272155 20021014 ; US20020389747P 20020617

Also published as: WO03106642 (A3) AU2003276392 (A1) AU2003276392 (A8) CA2489733 (A1) EP1551995 (A2) more

Abstract of WO03106642 (A2)

The present invention provides for novel methods and kits for reducing the complexity of a nucleic acid sample to interrogate a collection of target sequences. In one embodiment complexity reduction can be accomplished by extension of a locus specific capture probe followed by amplification of the extended capture probe using common primers. The locus specific capture probes may be attached to a solid support. Multiple DNA sequences may be amplified simultaneously to produce a reduced complexity sample. The invention further provides for analysis of the above sample to interrogate sequences of interest such as polymorphisms. The amplified sample may be hybridized to an array, which may be specifically designed to interrogate the desired fragments for the presence or absence of a polymorphism.

IDT Exhibit 1020.9779

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 December 2003 (24.12.2003)

PCT

(10) International Publication Number
WO 03/106642 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US03/18853

(22) International Filing Date: 16 June 2003 (16.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/389,747	17 June 2002 (17.06.2002)	US
10/272,155	14 October 2002 (14.10.2002)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	10/272,155 (CON)
Filed on	14 October 2002 (14.10.2002)
US	60/389,747 (CON)
Filed on	17 June 2002 (17.06.2002)

(71) Applicant *(for all designated States except US)*: **AFFYMETRIX, INC.** [US/US]; 3380 Central Expressway, Santa Clara, CA 95051 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JONES, Keith, W.** [US/US]; 980 Helena Drive, Sunnyvale, CA 95087 (US). **SHAPERO, Michael** [US/US]; 2449 Carson Street, Redwood City, CA 94061 (US). **LIU, Weiwei** [CN/US]; 765-86 San Antonio Road, Palo Alto, CA 94303 (US).

(74) Agent: **TREANNIE, Lisa, M.**; Hamilton, Brook, Smith & Reynolds, P.C., 530 Virginia Road, P.O. Box 9133, Concord, MA 01742-9133 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

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.

WO 03/106642 A2

(54) Title: COMPLEXITY MANAGEMENT OF GENOMIC DNA BY LOCUS SPECIFIC AMPLICATION

(57) Abstract: The present invention provides for novel methods and kits for reducing the complexity of a nucleic acid sample to interrogate a collection of target sequences. In one embodiment complexity reduction can be accomplished by extension of a locus specific capture probe followed by amplification of the extended capture probe using common primers. The locus specific capture probes may be attached to a solid support. Multiple DNA sequences may be amplified simultaneously to produce a reduced complexity sample. The invention further provides for analysis of the above sample to interrogate sequences of interest such as polymorphisms. The amplified sample may be hybridized to an array, which may be specifically designed to interrogate the desired fragments for the presence or absence of a polymorphism.

IDT Exhibit 1020.9780

COMPLEXITY MANAGEMENT OF GENOMIC DNA BY LOCUS SPECIFIC AMPLIFICATION

RELATED APPLICATIONS

This application is a continuation of U.S. Application No. 10/272,155, filed October 14, 2002, and claims the benefit of U.S. Application No. 60/389,747, filed June 17, 2002. The entire teachings of the above applications are incorporated herein by reference.

BACKGROUND OF THE INVENTION

The invention relates to enrichment and amplification of a collection of target sequences from a nucleic acid sample and methods of analyzing amplified product. In some embodiments target sequences are amplified by extension of a locus-specific primer followed by amplification of the extended locus-specific primer with a generic pair of primers. In some embodiments the locus-specific primers are attached to a solid support and extension takes place on the solid support. In some embodiments the invention relates to the preparation of target for array based analysis of genotype. The present invention relates to the fields of molecular biology and genetics.

The past years have seen a dynamic change in the ability of science to comprehend vast amounts of data. Pioneering technologies such as nucleic acid arrays

allow scientists to delve into the world of genetics in far greater detail than ever before. Exploration of genomic DNA has long been a dream of the scientific community. Held within the complex structures of genomic DNA lies the potential to identify, diagnose, or treat diseases like cancer, Alzheimer disease or alcoholism. Exploitation of genomic information from plants and animals may also provide answers to the world's food distribution problems.

Recent efforts in the scientific community, such as the publication of the draft sequence of the human genome in February 2001, have changed the dream of genome exploration into a reality. Genome-wide assays, however, must contend with the complexity of genomes; the human genome for example is estimated to have a complexity of $3x10^9$ base pairs. Novel methods of sample preparation and sample analysis that reduce complexity may provide for the fast and cost effective exploration of complex samples of nucleic acids, particularly genomic DNA.

Single nucleotide polymorphisms (SNPs) have emerged as the marker of choice for genome wide association studies and genetic linkage studies. Building SNP maps of the genome will provide the framework for new studies to identify the underlying genetic basis of complex diseases such as cancer, mental illness and diabetes. Due to the wide ranging applications of SNPs there is still a need for the development of robust, flexible, cost-effective technology platforms that allow for scoring genotypes in large numbers of samples.

SUMMARY OF THE INVENTION

The present invention provides for novel methods of sample preparation and analysis comprising managing or reducing the complexity of a nucleic acid sample by amplification of a collection of target sequences using target specific capture probes. In some embodiments the extended capture probes are attached to a solid support; in some embodiments the extended capture probes are in solution. In some embodiments the amplified collection of target sequences is analyzed by hybridization to an array that is designed to interrogate sequence variation in the target sequences. In some

embodiments the amplified collection of target sequences is analyzed by hybridization to an array of tag probes.

In one embodiment a method of amplifying a collection of target sequences from a nucleic acid sample is disclosed. A collection of capture probes is generated. The collection comprised a plurality of different species of primers wherein each species comprises a first common sequence and a 3' variable region that is specific for a target sequence in the collection of target sequences. Each target sequence is represented by at least one species of primer which hybridizes to the target sequence and the collection of capture probes is attached to a solid support so that the 3' end of the capture probes is available for extension. The nucleic acid sample is fragmented and an adapter that has a second common sequence is ligated to the fragments. Fragmentation in some embodiments is by one or more restriction enzymes. The adapter-ligated fragments are hybridized to the collection of capture probes and the capture probes are extended using the hybridized adapter-ligated fragments as template for extension and thereby incorporating the target sequence and the second common sequence into the 3' end of the extended capture probe. The extended capture probes are then amplified using first and second common sequence primers.

In some embodiments the capture probes are attached to the solid support through a covalent interaction. In another embodiment there is a tag sequence in the capture probes that is unique for each species of capture probe and the capture probes are attached to the solid support by hybridization to a collection of tag probes that are covalently attached to the solid support. In some embodiments each species of capture probe is attached to the solid support in a discrete location.

In another embodiment the extended capture probes are released from the solid support prior to amplification. Prior to releasing the extended capture probes from the solid support nucleic acids that are not covalently attached to the solid support may be removed.

In another embodiment the extended capture probes are enriched prior to amplification. In some embodiments capture probes are enriched by incorporation of labeled nucleotides into the extended capture probes followed by isolation of labeled

capture probes by affinity chromatography. In some embodiments capture probes are labeled with biotin and avidin, streptavidin or an anti-biotin antibody, which may be monoclonal, may be used to isolate extended capture probes. In another embodiment extended capture probes are made double stranded and single stranded nucleic acid in the sample is digested by, for example a nuclease, such as, for example Exonuclease I. In another embodiment the extended capture probes are circularized prior to amplification and uncircularized nucleic acid in the sample is digested by, for example, a nuclease, such as, for example, Exonuclease III. In some embodiments the extended capture probes are circularized by hybridizing an oligonucleotide splint to the extended capture probes so that the 5' and 3' ends of extended capture probes are juxtaposed and then ligating the ends of the extended capture probes.

In one embodiment a method of genotyping one or more polymorphic locations in a sample is disclosed. An amplified collection of target sequences from the sample is prepared and hybridized to an array designed to interrogate at least one polymorphic location in the collection of target sequences. The hybridization pattern is analyzed to determine the identity of the allele or alleles present at one or more polymorphic location in the collection of target sequences.

In another embodiment a method for analyzing sequence variations in a population of individuals is disclosed. A nucleic acid sample is obtained from each individual and a collection of target sequences from each nucleic acid sample is amplified. Each amplified collection of target sequences is hybridized to an array designed to interrogate sequence variation in the collection of target sequences to generate a hybridization pattern for each sample and the hybridization patterns are analyzed or compared to determine the presence or absence of sequence variation in the population of individuals.

In another embodiment a method of amplifying a collection of target sequences from a nucleic acid sample in solution is disclosed. A collection of capture probes is generated. The collection comprised a plurality of different species of primers wherein each species comprises a first common sequence and a 3' variable region that is specific for a target sequence wherein each target sequence in a collection of target sequences is

represented by at least one species of primer which hybridizes to the target sequence. The nucleic acid sample is fragmented and an adapter is ligated to the fragments so that the strand that is ligated to the 5' end of the fragment strands comprises a second common sequence and the strand that is ligated to the 3' end of the fragments lacks the second common sequence and is blocked from extension at the 3' end. The adapter-ligated fragments are hybridized to the collection of capture probes and the capture probes are extended using the hybridized adapter-ligated fragments as template for extension and thereby incorporating the target sequence and the complement of the second common sequence into the extended capture probes. The extended capture probes are then amplified with first and second common sequence primers.

In one embodiment an amino group is used to block extension at the 3' end of the adapter strand.

In some embodiments fragmentation of the nucleic acid sample is by digestion with one or more restriction enzymes.

In another embodiment a method for genotyping one or more polymorphisms in a nucleic acid sample is disclosed. The nucleic acid sample is fragmented and an adaptor comprising a first common priming sequence is ligated to the fragments. A collection of capture probes is ligated to the fragments. The capture probes have a second common priming sequence, a tag sequence unique for each species of capture probe, a first locus specific sequence, a Type IIs restriction enzyme recognition sequence, and a second locus specific sequence. The Type IIs restriction enzyme recognition sequence is positioned so that the enzyme will cut immediately 5' of the polymorphic base in a target sequence. The capture probes are extended to generate single-stranded extension products and then amplified using the first and second common sequence primers. The the amplified product is digested with a Type IIs restriction enzyme and the fragments are extended in the presence of one or more type of labeled ddNTP. In one embodiment the extension is done is four separate reactions, one for each ddNTP and the ddNTPs may be labeled with the same label. The extended fragments are then hybridized to four separate arrays. In another embodiment the ddNTPs are differentially labeled with at least two different labels and the extension

reactions may be done in less than four reactions and each reaction may be hybridized to a separate array. The arrays are arrays of tag probes that hybridize to the tag sequences in the capture probes. The hybridization pattern on each of the arrays is analyzed to determine at least one genotype.

In some embodiments the ddNTPs are labeled with biotin.

In another embodiment one of the common sequence primers is resistant to nuclease digestion and the sample is treated with a nuclease that cleaves 5' to 3'after the fragments are extended in the presence of labeled ddNTP. In one embodiment the primer is resistant to nuclease digestion because it contains phosphorothioate linkages. In some embodiments the nuclease is T7 Gene 6 Exonuclease.

In another embodiment a method for screening for sequence variations in a population of individuals is disclosed. A nucleic acid sample from each individual is provided and the sample is amplified and genotyped by one of the method of the invention and the genotypes from the samples are compared to determine the presence or absence of sequence variation in the population of individuals.

In another embodiment a kit for amplifying a collection of target sequences is disclosed. The kit has a collection of capture probes that is specific for a collection of target sequences and has a first common sequence that is common to all of the capture probes, an adapter that has a second common sequence; and a pair of first and second common sequence primers. In another embodiment the collection of capture probes in the kit is covalently attached to a solid support so that the 3' end of the capture probes is available for extension. In another embodiment the kit also provides a restriction enzyme, buffer, DNA polymerase and dNTPs. In some embodiments the restriction enzyme is a Type IIs restriction enzyme. In another embodiment the kit also contains a ligase, dNTPs, ddNTPs, buffer and DNA polymerase. In some embodiments one of the common sequence primers is resistant to nuclease digestion.

In another embodiment the capture probes also have a tag sequence unique for each species of capture probe and a Type IIs restriction enzyme recognition sequence. In another embodiment the adapter has a first strand comprising a common sequence and a second strand that does not contain the complement of that common sequence and

the second strand is blocked from extension at the 3' end by, for example, an amino group.

In another embodiment a collection of capture probes attached to a solid support is disclosed. The solid support may be arrays, beads, microparticles, microtitre dishes or gels.

In another embodiment a plurality of oligonucleotides attached to a solid support is disclosed. The solid support may be arrays, beads, microparticles, microtitre dishes or gels. The oligonucleotides may be released and used for a variety of analysis. The plurality of oligonucleotides may comprise a collection of capture probes.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention.

Figure 1 shows a method of amplifying specific target sequences using a capture probe that is locus specific and genomic DNA that has been ligated to an adapter. The capture probes are attached to a solid support and extended to incorporate the sequence of interest and the adapter sequence. The extended capture probes are released from the solid support and amplified with a single primer pair.

Figure 2 shows a method where the capture probes are attached to a solid support by hybridization to a probe that is covalently attached to the solid support. The probes on the array are complementary to a tag sequence in the 5' region of the capture probe. The capture probe hybridizes so that the 3' end is available for extension.

Figure 3 shows a schematic of solution-based multiplexed SNP genotyping. A sample is fragmented and ligated to an adaptor so that the adaptor sequence that hybridizes to the 3' end of the strands of the fragments is blocked from extension.

Locus specific capture probes are hybridized to the fragments and extended in solution then amplified by PCR using primers to A1 and A2. Prior to amplification the extended capture probes may be enriched by, for example, removal of non-extended products or by positive selection of extended products.

Figures 4A-4B show a method of multiplexed anchored runoff amplification wherein the alleles present at different polymorphic positions are analyzed by hybridization to an array of tag probes. The capture probe includes a recognition site for a Type IIs restriction enzyme so that the enzyme cuts immediately upstream of the polymorphic locus. The capture probe is extended by one labeled nucleotide and the identity of the nucleotide is determined by hybridization to an array or probes that are complementary to the tag sequences in the capture probes.

Figure 5 shows an enrichment scheme. Biotin is incorporated into the extended capture probes and biotin labeled extended capture probes are selected by affinity chromatography.

Figure 6 shows another enrichment scheme using nuclease that is specific for single stranded nucleic acid. Capture probes that are fully extended through the adapter site on the genomic DNA fragment are converted to double stranded DNA by annealing and extension of a primer that hybridizes to the adapter sequence.

Figure 7 shows another enrichment scheme. The ends of the extended capture probes are ligated together to form a circle using a splint oligonucleotide that is complementary to the primer sites at the ends of the extended capture probes. The sample is digested with an exonuclease so circularized sequences are protected from digestion.

DETAILED DESCRIPTION OF THE INVENTION

A description of preferred embodiments of the invention follows.

(A) General

The present invention has many preferred embodiments and relies on many patents, applications and other references for details known to those of the art.

Therefore, when a patent, application, or other reference is cited or repeated below, it should be understood that it is incorporated by reference in its entirety for all purposes as well as for the proposition that is recited.

As used in this application, the singular form "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an agent" includes a plurality of agents, including mixtures thereof.

An individual is not limited to a human being but may also be other organisms including but not limited to mammals, plants, bacteria, or cells derived from any of the above.

Throughout this disclosure, various aspects of this invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible sub-ranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed sub-ranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. The same holds true for ranges in increments of 10^5, 10^4, 10^3, 10^2, 10, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, or 10^{-5}, for example. This applies regardless of the breadth of the range.

The practice of the present invention may employ, unless otherwise indicated, conventional techniques and descriptions of organic chemistry, polymer technology, molecular biology (including recombinant techniques), cell biology, biochemistry, and immunology, which are within the skill of the art. Such conventional techniques include polymer array synthesis, hybridization, ligation, and detection of hybridization using a label. Specific illustrations of suitable techniques can be had by reference to the example herein below. However, other equivalent conventional procedures can, of course, also be used. Such conventional techniques and descriptions can be found in standard laboratory manuals such as *Genome Analysis: A Laboratory Manual Series (Vols. I-IV)*, *Using Antibodies: A Laboratory Manual*, *Cells: A Laboratory Manual*,

PCR Primer: A Laboratory Manual, and *Molecular Cloning: A Laboratory Manual* (all from Cold Spring Harbor Laboratory Press), *Stryer* (anyone have the cite), *Gait, "Oligonucleotide Synthesis: A Practical Approach" 1984, IRL Press, London,* Nelson and Cox (2000), *Lehninger, Principles of Biochemistry* 3rd Ed., W.H. Freeman Pub., New York, NY and Berg et al. (2002) *Biochemistry*, 5th Ed., W.H. Freeman Pub., New York, NYall of which are herein incorporated in their entirety by reference for all purposes.

The present invention can employ solid substrates, including arrays in some preferred embodiments. Methods and techniques applicable to polymer (including protein) array synthesis have been described in U.S.S.N 09/536,841, WO 00/58516, U.S. Patents Nos. 5,143,854, 5,242,974, 5,252,743, 5,324,633, 5,384,261, 5,424,186, 5,451,683, 5,482,867, 5,491,074, 5,527,681, 5,550,215, 5,571,639, 5,578,832, 5,593,839, 5,599,695, 5,624,711, 5,631,734, 5,795,716, 5,831,070, 5,837,832, 5,856,101, 5,858,659, 5,936,324, 5,968,740, 5,974,164, 5,981,185, 5,981,956, 6,025,601, 6,033,860, 6,040,193, 6,090,555, and 6,136,269, in PCT Applications Nos. PCT/US99/00730 (International Publication Number WO 99/36760) and PCT/US 01/04285, and in U.S. Patent Applications Serial Nos. 09/501,099 and 09/122,216 which are all incorporated herein by reference in their entirety for all purposes.

Patents that describe synthesis techniques in specific embodiments include U.S. Patents Nos. 5,412,087, 6,147,205, 6,262,216, 6,310,189, 5,889,165 and 5,959,098 which are each incorporated herein by reference in their entirety for all purposes. Nucleic acid arrays are described in many of the above patents, but the same techniques are applied to polypeptide arrays.

The present invention also contemplates many uses for polymers attached to solid substrates. These uses include gene expression monitoring, profiling, library screening, genotyping, and diagnostics. Gene expression monitoring and profiling methods can be shown in U.S. Patents Nos. 5,800,992, 6,013,449, 6,020,135, 6,033,860, 6,040,138, 6,177,248 and 6,309,822. Genotyping and uses therefore are shown in USSN 10/013,598, and U.S. Patents Nos. 5,856,092, 6,300,063, 5,858,659, 6,284,460, 6,361,947, 6,368,799 and 6,333,179 which are each incorporated herein by reference.

Other uses are embodied in U.S. Patents Nos. 5,871,928, 5,902,723, 6,045,996, 5,541,061, and 6,197,506 which are incorporated herein by reference.

The present invention also contemplates sample preparation methods in certain preferred embodiments. For example, see the patents in the gene expression, profiling, genotyping and other use patents above, as well as USSN 09/854,317, U.S. Patent Nos. 5,437,990, 5,215,899, 5,466,586, 4,357,421, and Gubler et al., 1985, Biochemica et Biophysica Acta, Displacement Synthesis of Globin Complementary DNA: Evidence for Sequence Amplification.

Prior to or concurrent with analysis, the nucleic acid sample may be amplified by a variety of mechanisms, some of which may employ PCR. *See,* e.g., *PCR Technology: Principles and Applications for DNA Amplification* (Ed. H.A. Erlich, Freeman Press, NY, NY, 1992); *PCR Protocols: A Guide to Methods and Applications* (Eds. Innis, et al., Academic Press, San Diego, CA, 1990); Mattila et al., *Nucleic Acids Res.* 19, 4967 (1991); Eckert et al., *PCR Methods and Applications* 1, 17 (1991); *PCR* (Eds. McPherson et al., IRL Press, Oxford); and U.S. Patent Nos. 4,683,202, 4,683,195, 4,800,159 4,965,188,and 5,333,675, each of which is incorporated herein by reference in their entireties for all purposes. The sample may be amplified on the array. See, for example, U.S Patent No 6,300,070 and U.S. patent application 09/513,300, which are incorporated herein by reference.

Other suitable amplification methods include the ligase chain reaction (LCR) *(e.g.,* Wu and Wallace, *Genomics* 4, 560 (1989), Landegren et al., *Science* 241, 1077 (1988) and Barringer et al. *Gene* 89:117 (1990)), transcription amplification (Kwoh et al., *Proc. Natl. Acad. Sci. USA* 86, 1173 (1989) and WO88/10315), self-sustained sequence replication (Guatelli et al., *Proc. Nat. Acad. Sci. USA,* 87, 1874 (1990), WO/88/10315 and WO90/06995), selective amplification of target polynucleotide sequences (U.S. Patent No 6,410,276), consensus sequence primed polymerase chain reaction (CP-PCR) (U.S. Patent No 4,437,975), arbitrarily primed polymerase chain reaction (AP-PCR) (U.S. Patent No 5, 413,909, 5,861,245) and nucleic acid based sequence amplification (NABSA). (*See,* US patents nos. 5,409,818, 5,554,517, and 6,063,603, each of which is incorporated herein by reference). Other amplification

methods that may be used are described in, U.S. Patent Nos. 5,242,794, 5,494,810, 4,988,617 and in USSN 09/854,317, each of which is incorporated herein by reference.

Additional methods of sample preparation and techniques for reducing the complexity of a nucleic sample are described in Dong et al., *Genome Research* 11, 1418 (2001), in U.S. Patent No 6,361,947, 6,391,592 and U.S. Patent application Nos. 09/512,300, 09/916,135, 09/920,491, 09/910,292, and 10/013,598, which are incorporated herein by reference in their entireties.

The present invention also contemplates detection of hybridization between ligands in certain preferred embodiments. See U.S. Pat. Nos. 5,143,854, 5,578,832; 5,631,734; 5,834,758; 5,936,324; 5,981,956; 6,025,601; 6,141,096; 6,185,030; 6,201,639; 6,218,803; and 6,225,625 and in PCT Application PCT/US99/ 06097 (published as WO99/47964), each of which also is hereby incorporated by reference in its entirety for all purposes.

The practice of the present invention may also employ conventional biology methods, software and systems. Computer software products of the invention typically include computer readable medium having computer-executable instructions for performing the logic steps of the method of the invention. Suitable computer readable medium include floppy disk, CD-ROM/DVD/DVD-ROM, hard-disk drive, flash memory, ROM/RAM, magnetic tapes and etc. The computer executable instructions may be written in a suitable computer language or combination of several languages. Basic computational biology methods are described in, e.g. Setubal and Meidanis et al., *Introduction to Computational Biology Methods* (PWS Publishing Company, Boston, 1997); Salzberg, Searles, Kasif, (Ed.), *Computational Methods in Molecular Biology*, (Elsevier, Amsterdam, 1998); Rashidi and Buehler, *Bioinformatics Basics: Application in Biological Science and Medicine* (CRC Press, London, 2000) and Ouelette and Bzevanis *Bioinformatics: A Practical Guide for Analysis of Gene and Proteins* (Wiley & Sons, Inc., 2nd ed., 2001).

The present invention may also make use of various computer program products and software for a variety of purposes, such as probe design, management of data, analysis, and instrument operation. See, U.S. Pat. Nos. 5,593,839, 5,795,716,

5,733,729, 5,974,164, 6,066,454, 6,090,555, 6,185,561, 6,188,783, 6,223,127, 6,229,911 and 6,308,170.

Additionally, the present invention may have preferred embodiments that include methods for providing genetic information over the internet. See U.S. patent applications and provisional applications 10/063,559, 60/349,546, 60/376,003, 60/394,574, and 60/403,381

The present invention provides a flexible and scalable method for analyzing complex samples of nucleic acids, such as genomic DNA. These methods are not limited to any particular type of nucleic acid sample: plant, bacterial, animal (including human) total genome DNA, RNA, cDNA and the like may be analyzed using some or all of the methods disclosed in this invention. The word "DNA" may be used below as an example of a nucleic acid. It is understood that this term includes all nucleic acids, such as DNA and RNA, unless a use below requires a specific type of nucleic acid. This invention provides a powerful tool for analysis of complex nucleic acid samples. From experimental design to isolation of desired fragments and hybridization to an appropriate array, the invention provides for fast, efficient and inexpensive methods of complex nucleic acid analysis.

(B) Definitions

Nucleic acids according to the present invention may include any polymer or oligomer of pyrimidine and purine bases, preferably cytosine, thymine, and uracil, and adenine and guanine, respectively. (*See* Albert L. Lehninger, *Principles of Biochemistry*, at 793-800 (Worth Pub. 1982) which is herein incorporated in its entirety for all purposes). Indeed, the present invention contemplates any deoxyribonucleotide, ribonucleotide or peptide nucleic acid component, and any chemical variants thereof, such as methylated, hydroxymethylated or glucosylated forms of these bases, and the like. The polymers or oligomers may be heterogeneous or homogeneous in composition, and may be isolated from naturally occurring sources or may be artificially or synthetically produced. In addition, the nucleic acids may be DNA or

RNA, or a mixture thereof, and may exist permanently or transitionally in single-stranded or double-stranded form, including homoduplex, heteroduplex, and hybrid states.

An "oligonucleotide" or "polynucleotide" is a nucleic acid ranging from at least 2, preferably at least 8, 15 or 20 nucleotides in length, but may be up to 50, 100, 1000, or 5000 nucleotides long or a compound that specifically hybridizes to a polynucleotide. Polynucleotides of the present invention include sequences of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) or mimetics thereof which may be isolated from natural sources, recombinantly produced or artificially synthesized. A further example of a polynucleotide of the present invention may be a peptide nucleic acid (PNA). (*See* U.S. Patent No. 6,156,501 which is hereby incorporated by reference in its entirety.) The invention also encompasses situations in which there is a nontraditional base pairing such as Hoogsteen base pairing which has been identified in certain tRNA molecules and postulated to exist in a triple helix. "Polynucleotide" and "oligonucleotide" are used interchangeably in this application.

The term "fragment," "segment," or "DNA segment" refers to a portion of a larger DNA polynucleotide or DNA. A polynucleotide, for example, can be broken up, or fragmented into, a plurality of segments. Various methods of fragmenting nucleic acid are well known in the art. These methods may be, for example, either chemical or physical in nature. Chemical fragmentation may include partial degradation with a DNase; partial depurination with acid; the use of restriction enzymes; intron-encoded endonucleases; DNA-based cleavage methods, such as triplex and hybrid formation methods, that rely on the specific hybridization of a nucleic acid segment to localize a cleavage agent to a specific location in the nucleic acid molecule; or other enzymes or compounds which cleave DNA at known or unknown locations (see, for example, USSN 09/358,664). Physical fragmentation methods may involve subjecting the DNA to a high shear rate. High shear rates may be produced, for example, by moving DNA through a chamber or channel with pits or spikes, or forcing the DNA sample through a restricted size flow passage, e.g., an aperture having a cross sectional dimension in the micron or submicron scale. Other physical methods include sonication and

nebulization. Combinations of physical and chemical fragmentation methods may likewise be employed such as fragmentation by heat and ion-mediated hydrolysis. *See* for example, Sambrook et al., "Molecular Cloning: A Laboratory Manual," 3rd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (2001) ("Sambrook et al.) which is incorporated herein by reference for all purposes. These methods can be optimized to digest a nucleic acid into fragments of a selected size range. Useful size ranges may be from 100, 200, 400, 700 or 1000 to 500, 800, 1500, 2000, 4000 or 10,000 base pairs. However, larger size ranges such as 4000, 10,000 or 20,000 to 10,000, 20,000 or 500,000 base pairs may also be useful.

A number of methods disclosed herein require the use of restriction enzymes to fragment the nucleic acid sample. In general, a restriction enzyme recognizes a specific nucleotide sequence of four to eight nucleotides and cuts the DNA at a site within or a specific distance from the recognition sequence. For example, the restriction enzyme *Eco*RI recognizes the sequence GAATTC and will cut a DNA molecule between the G and the first A. The length of the recognition sequence is roughly proportional to the frequency of occurrence of the site in the genome. A simplistic theoretical estimate is that a six base pair recognition sequence will occur once in every 4096 (4^6) base pairs while a four base pair recognition sequence will occur once every 256 (4^4) base pairs. *In silico* digestions of sequences from the Human Genome Project show that the actual occurrences may be more or less frequent, depending on the sequence of the restriction site. Because the restriction sites are rare, the appearance of shorter restriction fragments, for example those less than 1000 base pairs, is much less frequent than the appearance of longer fragments. Many different restriction enzymes are known and appropriate restriction enzymes can be selected for a desired result. (For a description of many restriction enzymes *see*, New England BioLabs Catalog which is herein incorporated by reference in its entirety for all purposes).

Type-IIs endonucleases are a class of endonuclease that, like other endonucleases, recognize specific sequences of nucleotide base pairs within a double stranded polynucleotide sequence. Upon recognizing that sequence, the endonuclease will cleave the polynucleotide sequence, generally leaving an overhang of one strand of

the sequence, or "sticky end." The Type-IIs endonucleases are unique because they generally do not require palindromic recognition sequences and they generally cleave outside of their recognition sites. For example, the Type-IIs endonuclease *Ear*I recognizes and cleaves in the following manner:

```
                    ↓
5'-C-T-C-T-T-C-N-N-N-N-N-3' (SEQ ID NO:1)
3'-G-A-G-A-A-G-n-n-n-n-n-5' (SEQ ID NO:2)
                          ↑
```

where the recognition sequence is -C-T-C-T-T-C-, N and n represent complementary, ambiguous base pairs and the arrows indicate the cleavage sites in each strand. As the example illustrates, the recognition sequence is non-palindromic, and the cleavage occurs outside of that recognition site.

Type-IIs endonucleases are generally commercially available and are well known in the art. Specific Type-IIs endonucleases which are useful in the present invention include, e.g., *Bbv*I, *Bce*AI, *Bfu*AI, *Ear*I, *Alw*I, *Bbs*I, *Bsa*I, *Bsm*AI, *Bsm*BI, *Bsp*MI, , *Hga*I, *Sap*I, *Sfa*NI, *Bsm*FI, *Fok*I, and *Ple*I. Other Type-IIs endonucleases that may be useful in the present invention may be found, for example, in the New England Biolabs catalogue. In some embodiments Type-IIs enzymes that generate a recessed 3' end are particularly useful.

"Adaptor sequences" or "adaptors" are generally oligonucleotides of at least 5, 10, or 15 bases and preferably no more than 50 or 60 bases in length; however, they may be even longer, up to 100 or 200 bases. Adaptor sequences may be synthesized using any methods known to those of skill in the art. For the purposes of this invention they may, as options, comprise primer binding sites, recognition sites for endonucleases, common sequences and promoters. The adaptor may be entirely or substantially double stranded. A double stranded adaptor may comprise two oligonucleotides that are at least partially complementary. The adaptor may be phosphorylated or unphosphorylated on one or both strands. Adaptors may be more efficiently ligated to fragments if they comprise a substantially double stranded region and a short single stranded region which is complementary to the single stranded region

created by digestion with a restriction enzyme. For example, when DNA is digested with the restriction enzyme *Eco*RI the resulting double stranded fragments are flanked at either end by the single stranded overhang 5'-AATT-3', an adaptor that carries a single stranded overhang 5'-AATT-3' will hybridize to the fragment through complementarity between the overhanging regions. This "sticky end" hybridization of the adaptor to the fragment may facilitate ligation of the adaptor to the fragment but blunt ended ligation is also possible. Blunt ends can be converted to sticky ends using the exonuclease activity of the Klenow fragment. For example when DNA is digested with *Pvu*II the blunt ends can be converted to a two base pair overhang by incubating the fragments with Klenow in the presence of dTTP and dCTP. Overhangs may also be converted to blunt ends by filling in an overhang or removing an overhang.

Methods of ligation will be known to those of skill in the art and are described, for example in Sambrook et at. (2001) and the New England BioLabs catalog both of which are incorporated herein by reference for all purposes. Methods include using T4 DNA Ligase which catalyzes the formation of a phosphodiester bond between juxtaposed 5′ phosphate and 3′ hydroxyl termini in duplex DNA or RNA with blunt and sticky ends; *Taq* DNA Ligase which catalyzes the formation of a phosphodiester bond between juxtaposed 5′ phosphate and 3′ hydroxyl termini of two adjacent oligonucleotides which are hybridized to a complementary target DNA; *E.coli* DNA ligase which catalyzes the formation of a phosphodiester bond between juxtaposed 5′ -phosphate and 3′ -hydroxyl termini in duplex DNA containing cohesive ends; and T4 RNA ligase which catalyzes ligation of a 5′ phosphoryl-terminated nucleic acid donor to a 3′ hydroxyl-terminated nucleic acid acceptor through the formation of a 3′ –>5′ phosphodiester bond, substrates include single-stranded RNA and DNA as well as dinucleoside pyrophosphates; or any other methods described in the art.

When a fragment has been digested on both ends with the same enzyme or two enzymes that leave the same overhang, the same adaptor may be ligated to both ends. Digestion with two or more enzymes can be used to selectively ligate separate adaptors to either end of a restriction fragment. For example, if a fragment is the result of digestion with EcoRI at one end and BamHI at the other end, the overhangs will be 5'-

AATT-3' and 5'GATC-3', respectively. An adaptor with an overhang of AATT will be preferentially ligated to one end while an adaptor with an overhang of GATC will be preferentially ligated to the second end.

An adaptor may be ligated to one or both strands of the fragmented DNA. In some embodiments a double stranded adaptor is used but only one strand is ligated to the fragments. Ligation of one strand of an adaptor may be selectively blocked. Any known method to block ligation of one strand may be employed. For example, one strand of the adaptor can be designed to introduce a gap of one or more nucleotides between the 5' end of that strand of the adaptor and the 3' end of the target nucleic acid. Adapters can be designed specifically to be ligated to the termini produced by restriction enzymes and to introduce gaps or nicks. For example, if the target is an *Eco*RI digested fragment an adapter with a 5' overhang of TTA could be ligated to the AATT overhang left by *Eco*RI to introduce a single nucleotide gap between the adaptor and the 3' end of the fragment. Phosphorylation and kinasing can also be used to selectively block ligation of the adaptor to the 3' end of the target molecule. Absence of a phosphate from the 5' end of an adaptor will block ligation of that 5' end to an available 3'OH. For additional adaptor methods for selectively blocking ligation *see* U.S. Patent 6,197,557 and USSN 09/910,292 which are incorporated by reference herein in their entirety for all purposes.

Adaptors may also incorporate modified nucleotides that modify the properties of the adaptor sequence. For example, phosphorothioate groups may be incorporated in one of the adaptor strands. A phosphorothioate group is a modified phosphate group with one of the oxygen atoms replaced by a sulfur atom. In a phosphorothioated oligo (often called an "S-Oligo"), some or all of the internucleotide phosphate groups are replaced by phosphorothioate groups. The modified backbone of an S-Oligo is resistant to the action of most exonucleases and endonucleases. Phosphorothioates may be incorporated between all residues of an adaptor strand, or at specified locations within a sequence. A useful option is to sulfurize only the last few residues at each end of the oligo. This results in an oligo that is resistant to exonucleases, but has a natural DNA center.

A genome is all the genetic material of an organism. In some instances, the term genome may refer to the chromosomal DNA. Genome may be multichromosomal such that the DNA is cellularly distributed among a plurality of individual chromosomes. For example, in human there are 22 pairs of chromosomes plus a gender associated XX or XY pair. DNA derived from the genetic material in the chromosomes of a particular organism is genomic DNA. The term genome may also refer to genetic materials from organisms that do not have chromosomal structure. In addition, the term genome may refer to mitochondria DNA. A genomic library is a collection of DNA fragments representing the whole or a portion of a genome. Frequently, a genomic library is a collection of clones made from a set of randomly generated, sometimes overlapping DNA fragments representing the entire genome or a portion of the genome of an organism.

The term "chromosome" refers to the heredity-bearing gene carrier of a living cell which is derived from chromatin and which comprises DNA and protein components (especially histones). The conventional internationally recognized individual human genome chromosome numbering system is employed herein. The size of an individual chromosome can vary from one type to another with a given multi-chromosomal genome and from one genome to another. In the case of the human genome, the entire DNA mass of a given chromosome is usually greater than about 100,000,000 bp. For example, the size of the entire human genome is about 3×10^9 bp. The largest chromosome, chromosome no. 1, contains about 2.4×10^8 bp while the smallest chromosome, chromosome no. 22, contains about 5.3×10^7 bp.

A "chromosomal region" is a portion of a chromosome. The actual physical size or extent of any individual chromosomal region can vary greatly. The term "region" is not necessarily definitive of a particular one or more genes because a region need not take into specific account the particular coding segments (exons) of an individual gene.

An allele refers to one specific form of a genetic sequence (such as a gene) within a cell, an individual or within a population, the specific form differing from other forms of the same gene in the sequence of at least one, and frequently more than one, variant sites within the sequence of the gene. The sequences at these variant sites that

differ between different alleles are termed "variances", "polymorphisms", or "mutations". At each autosomal specific chromosomal location or "locus" an individual possesses two alleles, one inherited from one parent and one from the other parent, for example one from the mother and one from the father. An individual is "heterozygous" at a locus if it has two different alleles at that locus. An individual is "homozygous" at a locus if it has two identical alleles at that locus.

The term genotyping refers to the determination of the genetic information an individual carries at one or more positions in the genome. For example, genotyping may comprise the determination of which allele or alleles an individual carries for a single SNP or the determination of which allele or alleles an individual carries for a plurality of SNPs. For example, a particular nucleotide in a genome may be an A in some individuals and a C in other individuals. Those individuals who have an A at the position have the A allele and those who have a C have the C allele. In a diploid organism the individual will have two copies of the sequence containing the polymorphic position so the individual may have an A allele and a C allele or alternatively two copies of the A allele or two copies of the C allele. Those individuals who have two copies of the C allele are homozygous for the C allele, those individuals who have two copies of the A allele are homozygous for the C allele, and those individuals who have one copy of each allele are heterozygous. The array may be designed to distinguish between each of these three possible outcomes. A polymorphic location may have two or more possible alleles and the array may be designed to distinguish between all possible combinations.

Polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A polymorphic marker or site is the locus at which divergence occurs. Preferred markers have at least two alleles, each occurring at frequency of preferably greater than 1%, and more preferably greater than 10% or 20% of a selected population. A polymorphism may comprise one or more base changes, an insertion, a repeat, or a deletion. A polymorphic locus may be as small as one base pair. Polymorphic markers include restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions,

minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, insertion elements such as Alu or small insertions or deletions, for example, deletions or insertions of 1-10 bases. The first identified allelic form is arbitrarily designated as the reference form and other allelic forms are designated as alternative or variant alleles. The allelic form occurring most frequently in a selected population is sometimes referred to as the wild type form. Diploid organisms may be homozygous or heterozygous for allelic forms. When an organism carries two identical alleles the organism is homozygous at that position. When an organism carries two different alleles the organism is heterozygous at that position. Normal cells that are heterozygous at one or more loci may give rise to tumor cells that are homozygous at those loci. This loss of heterozygosity may result from structural deletion of normal genes or loss of the chromosome carrying the normal gene, mitotic recombination between normal and mutant genes, followed by formation of daughter cells homozygous for deleted or inactivated (mutant) genes; or loss of the chromosome with the normal gene and duplication of the chromosome with the deleted or inactivated (mutant) gene.

Single nucleotide polymorphisms (SNPs) are positions at which two alternative bases occur at appreciable frequency (>1%) in the human population, and are the most common type of human genetic variation. The site is usually preceded by and followed by highly conserved sequences of the allele (e.g., sequences that vary in less than 1/100 or 1/1000 members of the populations).

A single nucleotide polymorphism usually arises due to substitution of one nucleotide for another at the polymorphic site. A transition is the replacement of one purine by another purine or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine by a pyrimidine or vice versa. Single nucleotide polymorphisms can also arise from a deletion of a nucleotide or an insertion of a nucleotide relative to a reference allele.

A diallelic polymorphism has two forms in a population. A triallelic polymorphism has three forms. A polymorphism between two nucleic acids can occur naturally, or be caused by exposure to or contact with chemicals, enzymes, or other

agents, or exposure to agents that cause damage to nucleic acids, for example, ultraviolet radiation, mutagens or carcinogens.

Linkage disequilibrium or allelic association means the preferential association of a particular allele or genetic marker with a specific allele, or genetic marker at a nearby chromosomal location more frequently than expected by chance for any particular allele frequency in the population. For example, if locus X has alleles a and b, which occur equally frequently, and linked locus Y has alleles c and d, which occur equally frequently, one would expect the combination ac to occur with a frequency of 0.25. If ac occurs more frequently, then alleles a and c are in linkage disequilibrium. Linkage disequilibrium may result from natural selection of certain combination of alleles or because an allele has been introduced into a population too recently to have reached equilibrium with linked alleles. A marker in linkage disequilibrium can be particularly useful in detecting susceptibility to disease (or other phenotype) notwithstanding that the marker does not cause the disease. For example, a marker (X) that is not itself a causative element of a disease, but which is in linkage disequilibrium with a gene (including regulatory sequences) (Y) that is a causative element of a phenotype, can be detected to indicate susceptibility to the disease in circumstances in which the gene Y may not have been identified or may not be readily detectable.

Capture probes are oligonucleotides that have a 5' common sequence and a 3' locus or target specific region or primer. The locus or target specific region is designed to hybridize near a region of nucleic acid that includes a region of interest so that the locus or target specific region of the capture probe can be used as a primer and be extended through the region of interest to make a copy of the region of interest. The common sequence in the capture probe may be used as a priming site in subsequent rounds of amplification using a common primer or a limited number of common primers. The same common sequence may be present in many or all or the capture probes in a collection of capture probes. Capture probes may also comprise other sequences, for example, tag sequences that are unique for different species of capture probes, and endonuclease recognition sites.

A tag or tag sequence is a selected nucleic acid with a specified nucleic acid sequence. A tag probe has a region that is complementary to a selected tag. A set of tags or a collection of tags is a collection of specified nucleic acids that may be of similar length and similar hybridization properties, for example similar T_m. The tags in a collection of tags bind to tag probes with minimal cross hybridization so that a single species of tag in the tag set accounts for the majority of tags which bind to a given tag probe species under hybridization conditions. For additional description of tags and tag probes and methods of selecting tags and tag probes see USSN 08/626,285 and EP/0799897, each of which is incorporated herein by reference in their entirety.

A collection of capture probes may be designed to interrogate a collection of target sequences. The collection would comprise at least one capture probe for each target sequence to be amplified. There may be multiple different capture probes for a single target sequence in a collection of capture probes, for example, there may be a capture probe that hybridizes to one strand of the target sequence and a capture probe that hybridizes to the opposite strand of the target sequence, these may be referred to as a forward locus or target specific primer and a reverse locus or target specific primer. There also may be two or more capture probes that hybridize at different locations downstream of the target sequence.

A collection of capture probes may be used to amplify a subset of a genome. The collection of capture probes may be initially used to generate a copy of the target sequences in the genomic sample and then the copies may be amplified using common primers. The amplification may be done simultaneously in the same reaction and often in the same tube.

The term "target sequence", "target nucleic acid" or "target" refers to a nucleic acid of interest. The target sequence may or may not be of biological significance. As non-limiting examples, target sequences may include regions of genomic DNA which are believed to contain one or more polymorphic sites, DNA encoding or believed to encode genes or portions of genes of known or unknown function, DNA encoding or believed to encode proteins or portions of proteins of known or unknown function, and DNA encoding or believed to encode regulatory regions such as promoter sequences,

splicing signals, polyadenylation signals, etc. The number of sequences to be interrogated can vary, but preferably are from about 1000, 2,000, 5,000, 10,000, 20,000 or 100,000 to 5000, 10,000, 100,000, 1,000,000 or 3,000,000 target sequences.

An "array" comprises a support, preferably solid, with nucleic acid probes attached to the support. Preferred arrays typically comprise a plurality of different nucleic acid probes that are coupled to a surface of a substrate in different, known locations. These arrays, also described as "microarrays" or colloquially "chips" have been generally described in the art, for example, U.S. Pat. Nos. 5,143,854, 5,445,934, 5,744,305, 5,677,195, 5,800,992, 6,040,193, 5,424,186 and Fodor et al., *Science*, 251:767-777 (1991). Each of which is incorporated by reference in its entirety for all purposes.

Arrays may generally be produced using a variety of techniques, such as mechanical synthesis methods or light directed synthesis methods that incorporate a combination of photolithographic methods and solid phase synthesis methods. Techniques for the synthesis of these arrays using mechanical synthesis methods are described in, e.g., U.S. Pat. Nos. 5,384,261, and 6,040,193, which are incorporated herein by reference in their entirety for all purposes. Although a planar array surface is preferred, the array may be fabricated on a surface of virtually any shape or even a multiplicity of surfaces. Arrays may be nucleic acids on beads, gels, polymeric surfaces, fibers such as fiber optics, glass or any other appropriate substrate. (*See* U.S. Patent Nos. 5,770,358, 5,789,162, 5,708,153, 6,040,193 and 5,800,992, which are hereby incorporated by reference in their entirety for all purposes.)

Arrays may be packaged in such a manner as to allow for diagnostic use or can be an all-inclusive device; e.g., U.S. Patent Nos. 5,856,174 and 5,922,591 incorporated in their entirety by reference for all purposes.

Preferred arrays are commercially available from Affymetrix under the brand name GeneChip® and are directed to a variety of purposes, including genotyping and gene expression monitoring for a variety of eukaryotic and prokaryotic species. (*See* Affymetrix Inc., Santa Clara and their website at affymetrix.com.)

Hybridization probes are oligonucleotides capable of binding in a base-specific manner to a complementary strand of nucleic acid. Such probes include peptide nucleic acids, as described in Nielsen et al., *Science* 254, 1497-1500 (1991), and other nucleic acid analogs and nucleic acid mimetics. See US Patent Application No. 08/630,427-filed 4/3/96.

The term "hybridization" refers to the process in which two single-stranded polynucleotides bind non-covalently to form a double-stranded polynucleotide; triple-stranded hybridization is also theoretically possible. The resulting double-stranded polynucleotide is a "hybrid." The hybrid may have double-stranded regions and single stranded regions.

Hybridizations are usually performed under stringent conditions, for example, at a salt concentration of no more than 1 M and a temperature of at least 25°C. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM NaPhosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C are suitable for allele-specific probe hybridizations. For stringent conditions, see, for example, Sambrook et al., (2001) which is hereby incorporated by reference in its entirety for all purposes above.

An individual is not limited to a human being, but may also include other organisms including but not limited to mammals, plants, bacteria or cells derived from any of the above.

(C.) Multiplexed Anchored Runoff Amplification

Generally, the invention provides methods for highly multiplexed locus specific amplification of nucleic acids and methods for analysis of the amplified products. In some embodiments the invention combines the use of capture probes that comprise a common sequence and a locus-specific region with adapter-modified sample nucleic acid; the adapter comprises a second common sequence. The capture probes are extended to produce copies of the sample DNA that contain common priming sequences flanking the target sequence. The copies are amplified with a generic set of primers that

recognize the common sequences. The amplified product may be analyzed by hybridization to an array of probes.

In one embodiment the steps of the invention comprise: generating capture probes; digesting a nucleic acid sample; ligating adaptors to the fragmented sample; mixing the fragments and the capture probes under conditions that will allow hybridization of the fragments and the capture probes; extending the capture probes in the presence of dNTPs and polymerase; amplifying the extended capture probes; and detecting the presence or absence of target sequences of interest.

One embodiment of the methods is illustrated in Figure 1. Capture probes are designed with a locus specific region ($LS1_F$ and $LS1_R$) that hybridizes near a target sequence of interest and a common sequence (A1) that is 5' of the locus specific region. The common priming site may be present in a plurality of capture probes so that a primer to A1 may be used for amplification of a plurality of different targets in subsequent steps. The capture probes are attached to a solid support so that they have a free 3' end. A plurality of a single species of capture probes may be synthesized at a discreet location on an array and may form a discrete feature of an array. Each feature of the array may contain a different species of locus specific capture probe.

Genomic DNA is fragmented and adapters comprising a second common sequence (A2) are ligated to the fragments. The adapter-ligated fragments are then mixed with the capture probes under conditions that allow hybridization of the fragments to the capture probes on the array. The capture probes are then extended using the adapter-ligated fragments as template. The extension product has a common sequence, A1, near its 5' end and a second common sequence A2 near its 3' end. These common sequences flank a region of interest. The capture probes are then released from the array and extended capture probes are amplified by PCR using primers to the common sequences A1 and A2. The amplified product may then be analyzed by, for example, hybridization to an array. Information about the region of interest can be determined by analysis of the hybridization pattern.

A second embodiment of the methods is illustrated in Figure 2. Capture probes are designed with a locus specific region (LS1 or LS2) and a common sequence (A1) as

in figure 1. In this embodiment the capture probes further comprise a tag sequence that is unique for each species of capture probe designed. (For a description of tags and tag probes, *see*, USSN 08/626,285.) The capture probes are attached to the array through hybridization of the tag sequence to a substantially complementary tag probe sequence that is attached to the array. The tag probes may be attached to the array in discrete locations. Different species of tag probes are present at different discrete, spatially addressable locations. Adapter-ligated genomic DNA is hybridized to the array so that the capture probes hybridize to target sequences in the sample. The capture probes are extended as in figure 1 to incorporate the target sequence and common sequence A2. The extended capture probes are released and amplified using primers A1 and A2. The amplified product may then be analyzed by, for example, hybridization to an array. Information about the region of interest can be determined by analysis of the hybridization pattern. The amplified sample may be analyzed by any method known in the art, for example, MALDI-TOF mass spec, capillary electrophoresis, OLA, dynamic allele specific hybridization (DASH) or TaqMan® (Applied Biosystems, Foster City, CA). For other methods of genotyping analyses *see* Syvanen, *Nature Rev. Gen.* 2:930-942 (2001) which is herein incorporated by reference in its entirety.

In some embodiments the capture probes are attached to a solid support prior to hybridization and hybridization takes place while the capture probes are attached to the solid support. In some embodiments the capture probes are synthesized on a solid support. Any suitable solid support known in the art may be used, for example, arrays, beads, microparticles, microtitre dishes and gels may be used. In some embodiments the capture probes are synthesized on an array in a 5' to 3' direction.

In some embodiments hybridization and extension of capture probes are done while the capture probes are attached to a solid support. Following extension of the capture probes nucleic acids that are not covalently attached to the solid support may be washed away. In some embodiments the extended capture probes are released from the solid support prior to amplification. In another embodiment amplification takes place while the extended capture probes are attached to the solid support. The extended capture probes may be released from the solid support by, for example, using a

reversible linker or an enzymatic release, such as an endonuclease or by a change in conditions that results in disruption of an interaction between the capture probe and the solid support, for example, when capture probes are associated with the solid support through base pairing between a tag in the capture probe and a tag probe on the solid support, disruption of the base pairing interaction releases the capture probes from the solid support. Enzymatic methods include, for example, use of uracil DNA glycosylase (UDG) or (UNG). UNG catalyzes the hydrolysis of DNA that contains deoxyuridine at the site the uridine is incorporated. Incorporation of one or more uridines in the capture probe followed by treatment with UNG will result in release of the capture probe from the solid support. A thermolabile UNG may also be used

In some embodiments a collection of target sequences is analyzed. A plurality of capture probes is designed for a plurality of target sequences. In some embodiments target sequences contain or are predicted to contain a polymorphism, for example, a SNP. The polymorphism may be, for example, near a gene that is a candidate marker for a phenotype, useful for diagnosis or a disorder or for carrier screening or the polymorphism may define a haplotype block (see, Daly et al. *Nat Genet.* 29:229-32 (2001), and Rioux et al. *Nat Genet.* 29:223-8 (2001) and U.S. Patent application 10/213,272, each of which is incorporated herein by reference in its entirety). A collection of capture probes may be designed so that capture probes hybridize near a polymorphism, for example, within 1, 5, 10, or 100 to 5, 10, 100, 1000, 10,000 or 100,000 bases from the polymorphism. The capture probes hybridize to one strand of the target sequence and can be extended through the polymorphic site or region so that the extension product comprises a copy of the polymorphic region.

Many amplification methods are most efficient at amplification of smaller fragments. For example, PCR most efficiently amplifies fragments that are smaller than 2 kb (see, Saiki et al. 1988). In one embodiment capture probes and fragmentation conditions are selected for efficient amplification of a selected collection of target sequences. The size of the amplified fragments is dependent on where the target specific region of the capture probe hybridizes to the target sequence and the 5' end of the fragment strand that the capture probe is hybridized to. In some embodiments of the

present methods capture probes and fragmentation methods are designed so that the target sequence of interest can be amplified as a fragment that is, for example, less than 20,000, 2,000, 800, 500, 400, 200 or 100 base pairs long. The capture probe can be designed so that the 3' end of the target specific region hybridizes to the base that is just 3' of a position to be interrogated in the target sequence. For example, if the sequence to be interrogated is a polymorphism and the sequence is 5'-GCTXATCGG-3', where X is the polymorphic position, the target specific region of the capture probe may have the sequence 5'---CCGAT-3'. When the sample is fragmented with site specific restriction enzymes the length of the fragments will also depend on the position of the nearest recognition site for the enzyme or enzymes used for fragmentation. A collection of target sequences may be selected based on proximity to restriction sites. In some embodiments target sequences are selected for amplification and analysis based on the presence of a sequence of interest, such as a SNP, and proximity to a cleavage site for a selected restriction enzyme. For example, SNPs that are within 200, 500, 800, 1,000, 1,500, 2,000 or 20,000 base pairs of either a restriction site, such as, for example, an *Eco*RI site, a *BglI* site, an *Xba*I site or any other restriction enzyme site may be selected to be target sequences in a collection of target sequences. In another method a fragmentation method that randomly cleaves the sample into fragments that are 30,100, 200, 500 or 1,000 to 100, 200, 500, 1,000 or 2,500 base pairs on average may be used.

In another embodiment, illustrated in Figure 3, the capture probes are in solution and hybridization and extension take place in solution. In this embodiment the nucleic acid sample is fragmented and adapter containing common sequences A2 and A3 is ligated to the fragments. In some embodiments one strand of the adapter, the strand that is ligated to the 3' end of the fragment strands lacks common sequence A2 and is blocked from extension at the 3' end. Ligation of the blocked adapter strand to the 3' end of the fragment strands prevents the fragments from being extended to incorporate A2 at both ends, thus preventing amplification of the fragments by primer A2 in the subsequent PCR amplification step. Capture probes with locus specific regions and common sequence A1 are mixed with the adapter-ligated fragments under conditions that allow hybridization of the capture probes to the adapter ligated fragments. The