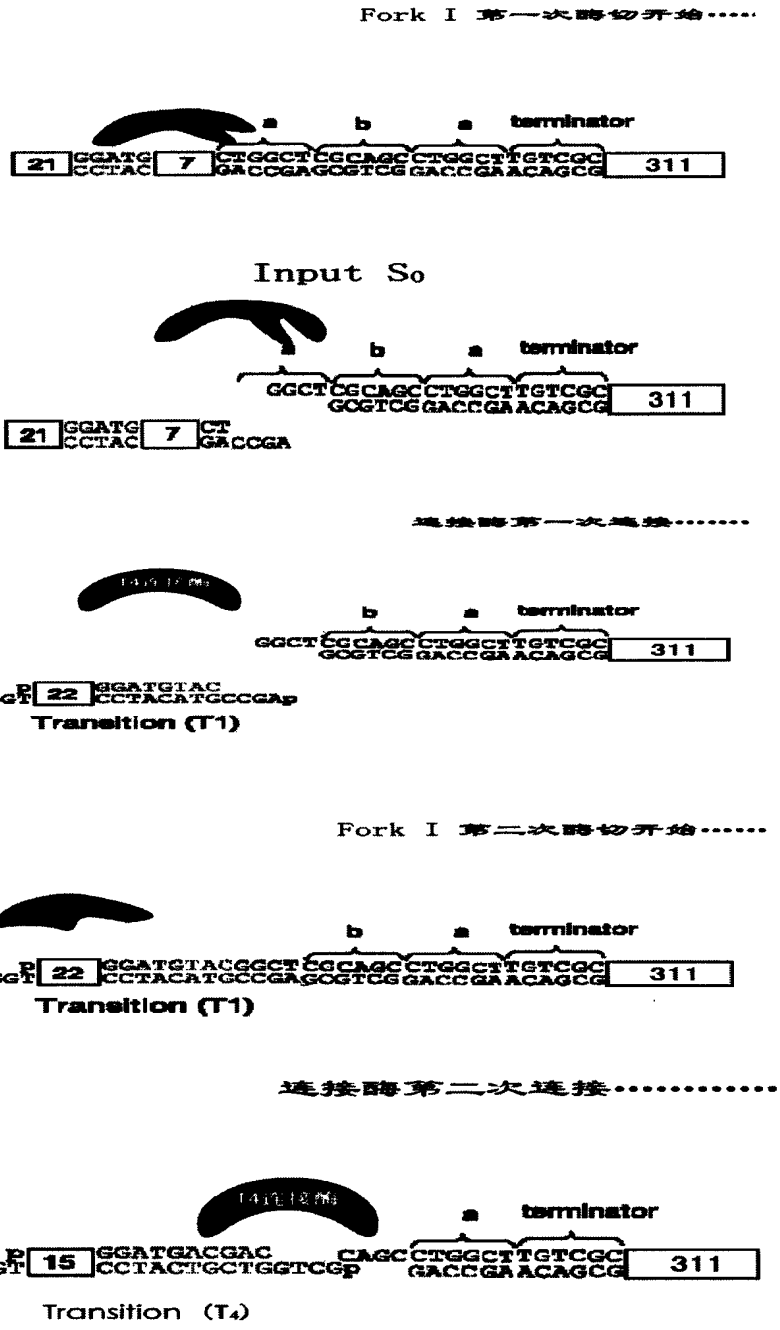


IDT Exhibit 1020.6787

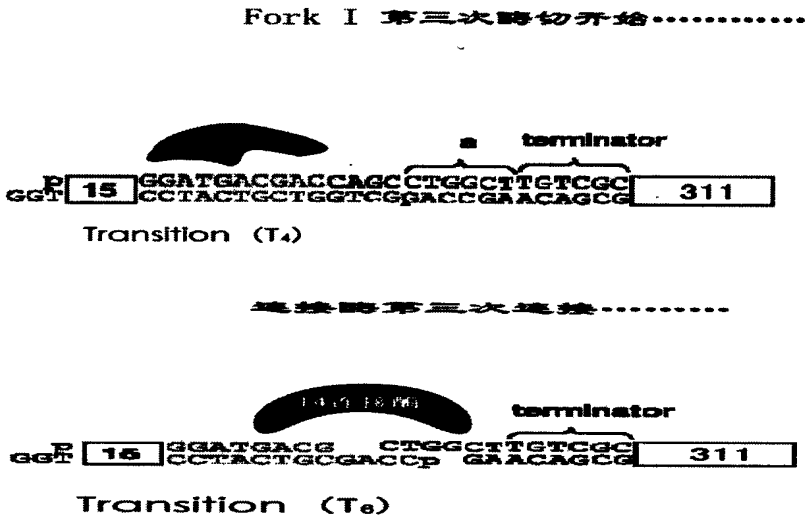


图 1

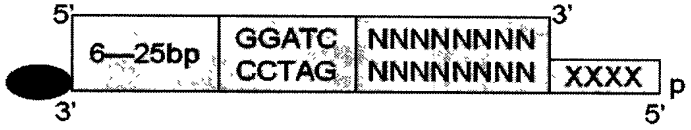


图 2

IDT Exhibit 1020.6788

Espacenet

Bibliographic data: WO0228876 (A2) — 2002-04-11

OLIGONUCLEOTIDE LINKERS COMPRISING A VARIABLE COHESIVE PORTION AND METHOD FOR THE PREPARATION OF POLYNUCLEOTIDE LIBRARIES BY USING SAID LINKERS.

Inventor(s): HAYASHIZAKI YOSHIHIDE [JP] ± (HAYASHIZAKI, YOSHIHIDE)

Applicant(s): RIKEN [JP]; HAYASHIZAKI YOSHIHIDE [JP] ± (RIKEN, ; HAYASHIZAKI, YOSHIHIDE)

Classification: - international: ***C07H21/00; C07H21/02; C07H21/04; C12N15/09; C12N15/10; C12N15/66; C12P19/34;*** *C07B61/00;* (IPC1-7): C07H21/00
- cooperative: **C12N15/1096; C12N15/66; C12P19/34;** C40B40/00

Application number: WO2001JP08805 20011005 Global Dossier

Priority number (s): JP20000306749 20001005

Also published as: WO0228876 (A3) EP1325118 (A2) EP1325118 (B1) JP2004510780 (A) JP4340779 (B2) more

Abstract of WO0228876 (A2)

A linker or population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion represented by formula (N)n, wherein N is A, C, G, T or U, or their derivatives, and n is an integer equal to or higher than 1. A linker-polynucleotide or a population of linker-polynucleotides comprising said linker or population of linkers and a target first strand polynucleotide bound to said linker. A method of preparing said linker or population of linkers and a method of preparing a linker-polynucleotide using said linker or population of linkers. Provided is a linker instead of G tailing, which can be utilized in a method of preparing a cDNA library, and a method of preparing a cDNA library using said linker.

IDT Exhibit 1020.6789

(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 02/28876 A2

(51) International Patent Classification[7]: C07H 21/00

(21) International Application Number: PCT/JP01/08805

(22) International Filing Date: 5 October 2001 (05.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2000-306749 5 October 2000 (05.10.2000) JP

(71) Applicant *(for all designated States except US)*: **RIKEN** [JP/JP]; 2-1, Hirosawa, Wako-shi, Saitama 351-0198 (JP).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **HAYASHIZAKI, Yoshihide** [JP/JP]; 22-1-201, Inarimae, Tsukuba-shi, Ibaraki 305-0061 (JP).

(74) Agents: **SHIOZAWA, Hisao** et al.; 8th Floor, Kyobashi Nisshoku Bldg., 8-7, Kyobashi 1-chome, Chuo-ku, Tokyo 104-0031 (JP).

(81) Designated States *(national)*: CA, JP, US.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

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 02/28876 A2

(54) Title: OLIGONUCLEOTIDE LINKERS COMPRISING A VARIABLE COHESIVE PORTION AND METHOD FOR THE PREPARATION OF POLYNUCLEOTIDE LIBRARIES BY USING SAID LINKERS.

(57) Abstract: A linker or population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion represented by formula (N)n, wherein N is A, C, G, T or U, or their derivatives, and n is an integer equal to or higher than 1. A linker-polynucleotide or a population of linker-polynucleotides comprising said linker or population of linkers and a target first strand polynucleotide bound to said linker. A method of preparing said linker or population of linkers and a method of preparing a linker-polynucleotide using said linker or population of linkers. Provided is a linker instead of G tailing, which can be utilized in a method of preparing a cDNA library, and a method of preparing a cDNA library using said linker.

IDT Exhibit 1020.6790

DESCRIPTION

Oligonucleotide linkers comprising a variable cohesive portion and method for the preparation of polynucleotide libraries by using said linkers.

TECHNICAL FIELD

The present invention relates to a population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion and to a method for the preparation of polynucleotide libraries comprising the use of said population of linkers. Further, the invention relates to an improved linker as a marker for specific libraries.

BACKGROUND ART

Oligonucleotide linkers and primers have been used in the prior art for priming, binding or annealing single strand polynucleotide and allowing the synthesis of the second polynucleotide complementary strand.

Carninci et al., 1996, Genomics, 37, 327-336; Carninci et al., 1997, DNA Research 4:61-66; Carninci et al., 1998, Proc.Natl.Acad.Sci USA, 95:520-4; Carninci and Hayashizaki, 1999, Methods Enzymol. 303:19-44, disclose method for the preparation of cDNA libraries. According to these protocols, a mRNA/cDNA hybrid is prepared and full-coding/full-length cDNAs are selected by mean of the Cap trapper technology, then

each single strand cDNA is ligated with G-tail and the cDNA second strand is synthesized.

However, the G-tailing methodology shows several drawbacks, for example, in the sequencing efficiency and translation efficiency when cDNAs clones are used for protein expression.

G-tailing is performed by terminal addition of dGTP using terminal deoxynucleotidyl transferase. However, the number of G added is difficult to control and it is variable, generally between 10 and 30. A long G-tail has the drawback of impairing a long read sequencing and lowering the sequencing efficiency, whilst a short G-tail has the drawback of providing a low efficient priming, with the consequence of lose of sample, and necessity of repreparing it.

During sequencing reaction, long G-stretches (long G-tail) interact with surrounding sequences and form very strong secondary structures. This may be problematic in case of interactions with 5' UTRs that are typically GC rich. In fact, a typical cDNA has 60% GC content in the 5'-UTR that is considered to act as a regulatory region. Similar problems were also observed in cloning vectors having GC rich region containing a *Sfi* I or *Not* I restriction sites next to cloning site.

Further, terminal deoxynucleotidyl transferase used for tailing reaction requires heavy metals, for example, $MnCl_2$ or $CoCl_2$. However, these heavy metals have sometimes caused degradation of cDNAs and decreased long strand,

IDT Exhibit 1020.6792

full-coding/full-length cDNA production rate.

The purpose of the present invention is to solve the several problems in the prior art and provide a novel and efficient method for the preparation of cDNA libraries.

More specifically, the purpose of the present invention is to provide a novel linker instead of G-tailing, which can be utilized in a method for the preparation of cDNA libraries and to provide a method for the preparation of cDNA libraries using said linker.

DESCRIPTION OF INVENTION

The present invention solves the above-mentioned problems by providing a linker comprising an oligonucleotide fixed portion and an oligonucleotide variable portion, the variable portion being represented as Formula (N)n wherein N is A, C, G, T or U, or their derivatives, and n is an integer equal to or higher than 1. When n is an integer equal to or higher than 2, the nucleotides (N) of the variable portion may be the same or different.

The variable portion is preferably prepared at random.

The linker according to the invention may be a single or a double strand linker.

The present invention further relates to a population of linkers comprising the two or more of the linkers of the present invention.

According to an embodiment of the present invention, it is provided a linker or

IDT Exhibit 1020.6793

population of linkers according to the present invention, which is prepared by;

(a) synthesizing a first oligonucleotide single strand comprising an oligonucleotide single strand fixed portion and an oligonucleotide single strand variable portion,

(b) synthesizing a second oligonucleotide single strand comprising an oligonucleotide single strand fixed portion complementary to the first oligonucleotide single strand fixed portion (a), and

(c) annealing the first oligonucleotide strand (a) to the second oligonucleotide strand (b) so that the variable portion protrudes outside the double strand fixed linker portion.

The present invention further relates to linker-polynucleotide product or population of linker-polynucleotide products comprising the linker or population of linkers according to the present invention and the target first strand polynucleotide bound to the linker.

The present invention still further relates to a vector comprising the linker-polynucleotide according to the present invention.

In addition, the present invention relates to a method for preparing the linker or population of linkers according to the present invention, which comprises the steps of:

(a) synthesizing a first oligonucleotide single strand comprising an oligonucleotide single strand fixed portion and an oligonucleotide single strand variable

IDT Exhibit 1020.6794

portion,

(b) synthesizing a second oligonucleotide single strand comprising an oligonucleotide single strand fixed portion complementary to the first oligonucleotide single strand fixed portion (a), and,

(c) annealing the first oligonucleotide strand (a) to the second oligonucleotide strand (b), so that the variable portion protrudes outside the double strand fixed portion.

The present invention is further directed to the following methods:

(1) a method of binding a target single strand polynucleotide to a linker comprising:

i) the preparation of the linker according to the present invention; and

ii) the step of annealing the variable portion of said linker to the target single strand polynucleotide;

(2) a method of binding a target single strand polynucleotide to a linker comprising:

i) the preparation of the linker according to the present invention; and

ii) the step of annealing the variable portion of one (first) strand of said linker to the target single strand polynucleotide and ligating the fixed portion of the other (second) strand of said linker to the target single strand polynucleotide;

(3) a method of binding a target single strand polynucleotide or a population of the polynucleotides to a population of linkers comprising:

i) the preparation of the population of the linker according to the present invention; and

IDT Exhibit 1020.6795

ii) the step of annealing the variable portion of said population of linkers to a population of the target single strand polynucleotides;

(4) a method of binding a target single strand polynucleotide or a population of the target single strand polynucleotides to a population of linkers comprising:

i) the preparation of the population of linkers according to the present invention; and

ii) the step of annealing the variable portion of the first strand of said population of the linkers to the target single strand polynucleotide or the population of the polynucleotides and ligating the fixed portion of the second strand of the population of the linker to the target single strand polynucleotide or the population of the polynucleotides;

(5) a method of preparing a linker-polynucleotide product comprising a linker and a double strand polynucleotide, comprising the steps of

i) annealing the variable portion of the linker according to the present invention to the target first strand polynucleotide, and

ii) synthesizing the second strand polynucleotide complementary to the target single strand polynucleotide;

(6) a method of preparing a linker-polynucleotide product comprising a linker and a double strand polynucleotide, comprising the steps of:

i) annealing the variable portion of the first strand of the linker according to the present invention to a target single strand polynucleotide and ligating the target single strand polynucleotide to the fixed portion of the second strand of the linker, and

ii) synthesizing the second single strand polynucleotide complementary to said

IDT Exhibit 1020.6796

target single strand polynucleotide;

(7) a method of preparing a linker-polynucleotide product comprising a linker or a population of linkers and a population of double strand polynucleotides, comprising the steps of:

i) annealing the variable portion of the linker or a population of the linkers according to the present invention to a target single strand polynucleotide or a population of the target single strand polynucleotides, and

ii) synthesizing the second strand polynucleotide complementary to said target single strand polynucleotide or a population thereof;

(8) a method of preparing a linker-polynucleotide product comprising a linker or a population of linkers and a population of double strand polynucleotides, comprising the steps of:

i) annealing the variable portion of the first strand of the linker or a population of the linkers according to the present invention to a target single strand polynucleotide or a population of the target single strand polynucleotides,

ii) ligating the target single strand polynucleotide or the population of the target single strand polynucleotides to the fixed portion of the second strand of the linker or the population of the linkers, and

iii) synthesizing the second single strand polynucleotide(s) complementary to said target single strand polynucleotide(s);

(9) a method of marking a polynucleotide library and distinguishing said library, which comprises the step of providing a population of linkers comprising a fixed portion and a variable portion (wherein the fixed portion comprises at least one marker indicating

IDT Exhibit 1020.6797

the defined tissue or species), and selecting and separating said library by said defined marker;

(10) a method of binding a linker or population of linkers to mRNA, which comprises the steps of:

(a) treating mRNA with phosphatase and removing phosphate groups from uncapped mRNA,

(b) treating a product of step (a) pyrophosphatase, which removes the CAP structure from capped mRNA, and

(c) adding an RNA ligase in the presence of the linker according to the present invention;

(11) a method of preparing a linker-polynucleotide product , which comprises the steps of:

(a) treating mRNA with phosphatase and removing phosphate groups from uncapped mRNA,

(b) treating a product of step (a) pyrophosphatase, which removes the CAP structure from capped mRNA,

(c) adding an RNA ligase in the presence of the linker according to the present invention, and

(d) adding an oligo dT and synthesizing a polynucleotide complementary to the complete sequence of said mRNA;

(12) a method of binding a linker and population of linkers according to the present invention to a target single strand polynucleotide or population of polynucleotides comprising the addition of RNA ligase;

IDT Exhibit 1020.6798

(13) a method of preparing DNA/RNA hybrids, which comprises the steps of:

i) providing a full-length/coding or long poly-A mRNAs,

ii) ligating and annealing said mRNAs to the linker according to any one of claims 1 to 41, the linker comprising a first restriction enzyme site,

iii) annealing oligo dT-primers comprising the second restriction enzyme site, to the mRNA,

iv) synthesizing cDNA strands,

v) isolating the hybrids by using restriction enzymes which recognize the two specific restriction enzyme sites introduced, and

vi) cloning;

(14) a method of preparing a linker-polynucleotide product comprising a linker and a single strand polynucleotide, comprising annealing the variable portion of the linker according to the present invention to the target first strand polynucleotide.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows an example of the procedure for preparation of full-length cDNA library using, as an example, a population of linkers comprising the variable portion GNNNNN.

PolyA+RNA is transcribed (A) and subsequently oxidized and attached to biotin. After RNase I treatment (B), only full-length cDNA has biotin and trapped with avidin coated magnetic beads (C).

IDT Exhibit 1020.6799

Figure 2 shows an example of the procedure for preparation of full-length cDNA library using, as an example, a population of linkers comprising the variable portion GNNNNN ,as a continuation of Fiure 1.

The cDNA is released from beads by alkaline treatment and recovered (D), and the linker is ligated (E). Shown is GN_5 linker. In the case of N6 linker, the variable portion is NNNNNN instead of GNNNNN. The second strand cDNA is synthesized (F) and digested with a restriction enzyme (G) and ligated into lambda phage vector (H) as well as packaged (I).

Figure 3 shows the result of the ligation between the tested cDNA and the linker.

The tested first strand cDNA prepared from 5 μg of 7.5kb poly(A)-tailed RNA (LifeTechnologies) was used as a starting material. The linker (GN_5-, Lanes 1 to 3; N_6, Lanes 4 to 6) was annealed to ligate with 50ng of 7.5kb tested cDNA. Subsequently, 10ng of linker-binding material sample was used for synthesizing the second strand cDNA, and then subjected to 0.8% alkali gel electrophoresis.

Linkers were used in different amounts: 200ng for Lanes 1 and 4; 500ng for Lanes 2 and 5, and 2 μg for Lanes 3 and 6). As a control, cDNA without linkers was used as a template of the second strand synthesis (Lane 7). Lane 8 comprises the first strand cDNA without linkers. Lane 9 comprises λ/HindDIII size marker.

Figure 4 shows the result of examining the ratio between a linker ligation cDNA of intercellular cDNA and a linker, using linkers with various molar ratios.

A linker combined with 2 μg of N_6/GN_5 (N_6:GN_5=1:4) was ligated with various quantities of cDNAs (1000ng for Lane 1, 500ng for Lane 2 and 200ng for Lane 3).

IDT Exhibit 1020.6800

They were supplied for the second strand cDNA synthesis and analyzed by 0.8% gel electrophoresis. Lane 4 comprises markers.

The indivisual lane on the left side of the Figure refers to the first strand cDNA.

Figure 5 shows a sequencing chart of a cDNA sequence having G-tail described in the prior art.

In the presence of a repetition of C in the second strand cDNA (introduced with the G-tail in the first strand), the efficiency of sequencing dropped down as shown in the chart.

Figure 6 shows a sequencing chart of a cDNA sequence ligated with N_6/GN_6 linker mixture (proportion 1:4).The sequenced clone D05_042_2-5F-ab2 in Figure 6 corresponds to the sample 2.05 in Table 1.

Figure 7 shows a sequencing chart of a cDNA sequence ligated with GN_5 linker. The sequenced clone G07_052_3-7F.ab1 in Figure 7 corresponds to the sample 3.07 in Table 1.

Figure 8 schematially shows a loop bias and the possible solutions. In Fig.8(A), the bias are explained. One of the end of the fixed or constant portion of the single strand linker, after removal of the other single strand from the double strand linker, may interact with the single strand cDNA to form a loop and block the following synthesis of the second strand cDNA.

As shown in Fig.8(B), NH_2 as a protecting group is bound to 3' end in the case of one end of the constant or fixed lower strand. There is no possibility of forming a loop, as well as the second strand cDNA synthesis is not inhibited.

IDT Exhibit 1020.6801

In Fig.8(C), 3' end of the fixed or nonvariable second strand (lower strand in the figure) and 5' end of the fixed or nonvariable first strand (upper strand in the figure) of a linker are bound together to form a loop. This prevents the possiblity of forming a loop with a single strand cDNA, as well as the second strand cDNA synthesis is not inhibited

DETAILED DESCRIPTION OF THE INVENTION

The present invention solves the above-mentioned problems in the prior art by providing a linker and a population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion. Such a linker and a population of linkers can bind to an end of a target single strand polynucleotide or a population of target polynucleotides, as well as allow the second polynucleotide strand synthesis.

According to the embodiment of the present invention, provided are a linker and a population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion. The fixed portion is preferably an oligonucleotide portion which is nonvariable in any linker of a population of linkers. The linker fixed portion can be a single strand or double strand oligonucleotide, preferably, a double strand oligonucleotide. The fixed portion preferably comprises at least one of the followings: a restriction site, a recombinational site, a polymerase promoter site, a marker or a tag.

IDT Exhibit 1020.6802

The variable portion is preferably synthesized at random. The linker of the resulting population of linkers has a nonvariable portion, preferably the common portion among the population and a variable portion which is different for each linker among the population.

The population of linkers according to the present invention may comprise, as a variable portion, one or more linkers having an oligo sequence specific for the 3' or 5' end of a target single strand polynucleotide. Said oligo sequence is specifically selected in order to bind and isolate one or more specific target polynucleotides among a population of target single strand polynucleotides.

The linker variable portion according to the present invention may also work as a primer in the synthesis of the second strand polynucleotide, as for example a long strand full-cording or full-length cDNA.

Said randomly variable single strand oligonucleotide portion can comprise any kind of nucleotide. Preferably, the variable portion has the formula (N)n, wherein N is A, C, G, T or U, or their derivatives, and n is equal to or higher than 1. When the integer n is equal or higher than 2, the nucleotides of the variable portion may be the same or different from each other.

The integer n is advantageously is from 1 to 10, preferably from 4 to 8, more preferably n is 5 or 6.

IDT Exhibit 1020.6803

As one method, in the variable portion the first to third nucleotides (beginning counting from the side of the fixed portion to the free end as shown in Fig.2) of (N)n can be G. A mixture having different strand length (that is, the length of n), and presence or absence of G is also within the object of the present invention. Preferably, it is a mixture of N_6/GN_5 with different proportion and the proportion is preferably 1:4.

The present invention further relates to a linker-polynucleotide product comprising the linker according to the present invention and a single strand or double strand polynucleotide annealed and/or ligated to the variable portion of said linker, and to a vector comprising said linker-polynucleotide product .

Preferably, said single or double strand polynucleotide is a long strand, full-coding/full-length cDNA.

Accordingly, the present invention discloses a method for the preparation of linker-polynucleotide products or polynucleotide libraries comprising the linkers according to the present invention annealed and/or ligated to target single strand polynucleotides.

The present invention further relates to a method for the preparation of polynucleotide products or libraries comprising the linkers according to the invention annealed/ligated to a double strand polynucleotides, preferably, to a method for the preparation of long strand, full-coding/length cDNA libraries.

IDT Exhibit 1020.6804

The present invention further relates to a method for marking polynucleotide libraries by providing a linker population according to the invention comprising a marker in the fixed portion. This marking system allows to distinguish and recognize libraries of different species (for instance, human, mouse, Drosophila, rice, and the like) and it can be used for distinguishing libraries of different tissues (for instance, liver, brain, lungs, and the like) for each species.

According to the present invention, provided is a method for binding a linker and a linker population described below and said linker or linker population with a target single strand polynucleotide or a population of target single strand polynucleotides, comprising the steps of:

i) preparing a linker or a population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable single strand portion;

ii) annealing a target single strand polynucleotide(s) to the variable portion(s) of said linker.

Hereinafter, in some cases, the present invention, for simplicity, will be described for a population of linkers (also indicated as a population of linkers or simply linkers) as well as to a method for binding said population to target polynucleotides and to a method for synthesizing double strand polynucleotides. However, it is clear that the present invention also includes the indivisual linkers comprised in said population of linkers and a method comprising the use of such indivisual linker.

IDT Exhibit 1020.6805

The fixed portion of each linker comprised in the population of linkers according to the present invention can be a single or double strand oligonucleotide.

Preferably, the fixed portion, as well as the linker, is a double oligonucleotide, accordingly, the method comprises the steps of:

i) preparing a linker or a population of linkers comprising an oligonucleotide double strand fixed portion and oligonucleotide variable single strand portion, wherein said variable portion is protruding outside said double strand fixed portion (therefore, the variable portion forms a cohesive protruding end);

ii) annealing a population of target single strand polynucleotides to the variable portions of the population of linkers, and preferably ligating the annealed end of said target single strand polynucleotides to the adjacent fixed portions of linkers (see steps (F) to (G) in Figure 2).

Said linker or linkers can be prepared with any methodology known in the prior art, for example, with one oligonucleotide strand having the direction 5'-3', comprising a fixed nucleic acid sequence at the 5' portion and a variable nucleic acid sequence end at the 3' portion. Then, the other oligonucleotide strand comprising a fixed nucleic acid sequence is prepared. Finally, the fixed portion of one strand and the fixed portion of the other strand are annealed so that the variable portion of one strand protrudes outside the double strand fixed portion. As a matter of course, the linker can also be prepared with inverted order of steps and with other methodologies.

IDT Exhibit 1020.6806

When the linker according to the present invention is a double strand linker, for the purpose of the present application, the strand comprising the fixed portion and the variable portion is also referred to as "the first strand", while the other strand complementary to the fixed portion of the first strand is referred to as "second strand".

The present invention further relates to a linker comprising only one oligonucleotide strand comprising a fixed portion and a variable portion (that is only comprising the "first strand" and not comprising the "second strand") . If the linker has the direction 5'-3', the variable portion of the end of this single strand oligonucleotide anneales to the 3' end of a target single strand polynucleotide. Then, the second strand polynucleotide, complementary to this target single strand polynucleotide, is synthesized.

One or more linkers can also be prepared in such a way that the strand comprising a variable portion has a direction 3'-5'. As a result, the variable portion is positioned at the 5' end. The 5' end variable portion of the linker thus prepared anneals to the 5' end of the target polynucleotide. If the other strand is present, its lower strand can be ligated to the 5' end of the target polynucleotide.

The present invention also provides a method for preparing linker-polynucleotide products comprising the linker or the population of linkers according to the present invention and a double strand polynucleotide, comprising the steps of:

IDT Exhibit 1020.6807

i) preparing a population of linkers comprising an oligonucleotide double strand fixed portion and an oligonucleotide variable single strand portion, wherein said variable portion is protruding outside the double strand fixed portion (therefore, the variable portions form cohesive protruding ends);

ii) annealing a population of the target single strand polynucleotides to the variable portions of said population of linkers, and ligating said population of target single strand polynucleotides to the adjacent (second strand) fixed portions of said linkers; and

iii) synthesizing second single strand polynucleotides, complementary to the target single strand, by using the variable portions as primers.

The polynucleotide sequence can also be prepared using only one strand of the linker according to the present invention. In this case, comprised are the steps of:

i) preparing a population of single strand linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable single strand portion;

ii) annealing a population of the target first strand polynucleotides to a variable portion of the population of linkers;

iii) synthesizing second single strand polynucleotides, complementary to the target first strand, by using the variable portions of linkers as primers.

The fixed oligonucleotide portion of the linker or population of linkers can be any oligonucleotide sequence. This fixed sequence is preferably a nonvariable portion and it is therefore the common for all the linkers of the same population. This fixed

IDT Exhibit 1020.6808

portions can also comprise oligo sequences consisting of one or more groups and therefore the fixed portions, in this case, can show some differences among the same population. However, since these oligo sequences consisting of one or more optional groups will not change the general structure of the fixed portions, for the purpose of the present invention, the fixed portion, comprising or not comprising the variable oligo sequences consisting of one or more groups will be, for simplcity, indicated as fixed portion.

The fixed portion can be a nonvariable portion even in the population of linkers. That is, it can be the same one for any linkers comprised in the population.

The fixed portion can be any oligonucleotide sequence (DNA or RNA), and it is preferably the same or almost the same for the linker or the population of linkers used in the specific experiments or for the specific library.

The linker fixed portion can therefore be intended both as a single or double strand oligonucleotide, preferably, it is a double strand oligonucleotide. In this case, the single strand variable portion constitutes a protruding end. The variable portion can act as a primer in the second strand polynucleotide synthesizing process.

The fixed (or nonvariable) portion preferably comprises one or more restriction site, homologous recombinational site, polymerase promoter site, a marker and/or a tag. Preferable restriction sites are, for example, BamHI, XhoI, SstI, SaII or NotI and others,

IDT Exhibit 1020.6809

for example those disclosed in Hyone-myong Eun, Chapter "Restriction endonucleases and modification methylases" .

Examples of homologous recombinational sites are attB, Gateway™ (Life Technologies), Cre-lox (Qinghua Liu, et al., 1998, Current Biology, 8:1300-1309) Flp/FRT (J. Wild, et al, 1996, Gene, 179:181-188).

Further, as for the polymerase promoter site, it can be a RNA polymerase promoter site, for example one of those cited in Hyone-Myong Eun, page 521. Preferably, it can be T3, T7,SP6, K11 and/or BA14 RNA polymerase promoter site.

A marker can be any sequence or sequences of nucleotides, for example a sequence specific for a particular tissue or species.

As a tag, any group or molecule able to be bound to an end of the fixed portion of one strand or the other strand can be used. In fact, when a single strand of the linker is removed, for example by increasing temperature, the end of the other single strand could form a loop with the target single strand polynucleotide (Fig. 8). Preferably, a protecting group is bound to the 3' end of the strand consisting of only this fixed portion in order to avoid that the end of the strand consisting of only the fixed portion forms a loop with the target single strand polynucleotide mainly ligated to this strand, as well as inhibits the synthesis of the second strand polynucleotide (Figure. 8A). Therefore, any

IDT Exhibit 1020.6810

group that does not have a 3'-OH and cannot be ligated nor extended by DNA polymerases, can be used for the purpose of the present invention.

As a protecting group, for example ddNTPs can be used. Preferably, a NH_2 group is also used as a protecting group (Figure 8, B).

As further particular solution, in order to avoid the problem of the loop bias, both ends can be bounded together so that the ends of both strands of the fixed portion of the linker positioned opposite to a variable portion form a loop (Figure 8, C). With this solution, the ends of the fixed portion cannot form a loop with the target single strand polynucleotide, and the synthesis of the second strand polynucleotide is not inhibited.

The oligonucleotide variable portion of the linker or the population of linkers is, preferably, randomly synthesized. Accordingly, in a population of linkers, the variable portion of any linker is preferably synthesized at random and the sequence of the variable portion and/or the number of base comprised in each linker differs each other. A population of linkers, therefore, comprises protruding ends having a high number of different sequences. Such a population of linkers comprises a high variation of random protruding ends. These recognize, anneal and/or ligate to the complementary ends of a population of target single strand polynucleotides. That is, this is a population of full-length cDNAs forming polynucleotide sequences comprising the linker and a target single strand polynucleotide (see Figures 1 and 2).

IDT Exhibit 1020.6811

The present invention therefore also relates to a population of linkers comprising at least two linkers prepared according to the present invention. Preferably, the invention relates to a population of linkers comprising at least two subpopulations of linkers.

The population of linker can be that one in which the fixed portions comprised in all the linkers are an oligonucleotide portion having the same sequence.The population of linkers may also comprises two or more subpopulations of linkers, wherein one subpopulation of linkers comprises linkers in which the fixed portion is an oligonucleotide portion having the same sequence, and other subpopulations of which differ each other in the fixed portion sequence.

Preferably, in the population or subpopulation of linkers the variable portions of the linkers are synthesized at random. Preferably, in the population or subpopulation, the sequence of the variable portions of the linkers are different from each other.

The variable portion can also be a specific oligonucleotide sequence complementary for an end (preferably 3' end) of a target single strand polynucleotide.

Preferably, the linker population of the present invention comprises among the variable portions, one or more specifically determined portions able to recognize and anneal end(s) of specific target polynucleotides which are intended to select from the population of target polynucleotides.

IDT Exhibit 1020.6812

The end of a target single strand polynucleotide anneals to the protruding variable end of the linker. When a population of linkers is added to a population of target single strand poynucleotides, the variable portions (protruding ends), preferably randomly synthesized, in the linkers recognize and anneal to the ends of the population of target single strand polynucleotides.

Preferably, the linkers according to the invention are double strand oligonucleotides comprising a fixed portion (preferably a nonvariable portion for all the linkers of a population) and a variable portion, which is different for any linker of the population. According to the first embodiment, the 3' end of the target single strand polynucleotide anneals to the 3' end of protruding end of the variable portion and ligates to the 5' end off the fixed portion of the other strand adjacent to the 3' end of said target single strand polynucleotide.

The linker can also be constituted, according to the second embodiment, by the fixed portions of one strand having the direction 3'-5' and the other strand. In this case, the 5' end of the target single strand polynucleotide anneals and ligates to this variable portion of the linker.

The variable single strand oligonucleotide portion of the linker can comprises any kind on nucleic acid. Preferably, said variable portion has the formula (N)n, wherein N is A, C, G, T or U or their derivatives and n is equal to or higher than 1, and if n is an integer equal to or higher than 2, the nucleotides (N) of the variable portion may

IDT Exhibit 1020.6813

be the same or differ from each other. Preferably, $1 \leqq n \leqq 10$ and more preferably $4 \leqq n \leqq 8$. As a particular preferred linker, n is 5 or 6, that is, N_6 or N_5.

The first, second and/or third N, closest to the fixed portion (that is, the nucleotides of the variable portion coming from the 5' end of the linker in case of the first embodiment) can also be a G, according to the formula $(G)_m(N)_{n-m}$, wherein m = 1 to 3. Preferably, the linker variable portion can be GN_4, GN_5, G_2N_3, G_2N_4, G_3N_2, G_3N_3, N_5, N_6 or a mixer thereof.

More specifically, the linker population according to the present invention is a mixture of $(N)_n$ linker and $(G)_m(N)_{n-m}$, preferably N_6/GN_5, N_6/G_2N_4 or N_6/G_3N_3 having different proportion. The proportion of the N_6/GN_5 mixture linker can be 0:1- 1:0, preferably 1:3-1:5, more preferably 1:4. The ligation can be realized with any ligation method known in the prior art, preferably by using a DNA ligase, more preferably a T4 DNA ligase or E. coli DNA ligase (see for example, Hyone-Mong Eun, under the chapter "Ligases") or using a RNA ligase (Maruyama et al. 1995).

Preferably, the ligation reaction according to the present invention comprises the addition of ligase stimulating agents. Preferably, as a ligase stimulating agent, PEG (Polyethylene glycol), preferably at 6000-8000 molecular weight, is used.

After the annealing and/or ligation step, the variable linker portion (that is, the protruding or free 3' end of the linker of the first embodiment) can act as a primer for the

IDT Exhibit 1020.6814

synthesis of the second strand polynucleotide, forming a polynucleotide sequence comprising the linker according to the invention and a double strand polynucleotide.

The ligation of the linker according to the invention to a target single strand polynucleotide can also be realized by using the oligo-capping technique (K.Maruyama et al., 1995, Gene, 138:171-174; and S.Kato et al., 1995, Gene, 150:243-250). The oligo-capping essentially comprises the following steps: i) mRNAs extracted from cells are treated with a phosphatase enzyme, preferably bacterial alkaline phosphatase for removing phosphates from non full-length mRNAs (that is, forming 5' ends of uncapped RNA having an hydroxyl at the 5' end, but not removing the CAP structure from the capped full-length RNA); ii) the mixture obtained in i) is treated with a pyrophosphatase, preferably tobacco acid pyrophosphatase (TAP), which removes the CAP structure from full-length mRNAs and leaves the full-length 5' ends with a phosphate group; iii) the full-length mRNA having a phosphate group at the 5' end is ligated to a specific RNA or a DNA adapter with a RNA ligase; and iv) an oligo dT is added and the complementary strand is synthesized.

The method for binding the linker according to the present invention to a target polynucleotide and/or the method of preparing a polynucleotide sequence according to the present invention can also be performed using, as ligation step, a modified oligo-capping method as follow.

IDT Exhibit 1020.6815

Accordingly, a target single strand polynucleotide (which can be RNA, mRNA prepared as described as the oligo-capping method, or cDNA) can be ligated to the linker according to the present invention in presence of a ligase.

In particular, when the linker of the present invention is a double strand linker, an end of the target single strand polynucleotide ligates to the second strand (which consists of the only fixed portion) and anneals to the variable portion of the first portion of the linker.

As an another possibility, the target polynucleotide ligates to the variable portion of the linker (which can be both a single or a double strand). In both cases, an oligo dT is added and a complementary polynucleotide, preferably cDNA, is synthesized.

The use of RNA ligase is not limited to RNA or mRNA as above described, but can also be used to ligate DNA.

The ligation method using RNA ligase in order to bind the linker according to the present invention and a target single strand polynucleotide, can therefore bind:
i) a single strand DNA to a single strand DNA; ii) a single strand RNA to a single strand RNA; and iii) a single strand DNA to a single strand RNA or a single strand RNA to a single strand DNA.

IDT Exhibit 1020.6816

According to an embodiment, the polynucleotide is a long strand, full-coding/full-length mRNA and the linker is DNA (but can also be RNA) and comprises a first restriction enzyme site.

Accordingly, it is provided a method for the preparation of a single or double strand cDNA comprising the steps of:

(I) providing a long strand, full-coding or full-length mRNA comprising a poly-A;

(II) providing a double strand linker comprising a first restriction enzyme site;

(III) ligating the 5' end of the mRNA (by using a ligase, for instance RNA ligase) to the fixed portion of the second strand of the linker and annealing the 5' end to variable portion of the first strand of the linker;

(IV) providing an oligo dT-primer comprising a second restriction enzyme site, and annealing the oligo dT-primer to the poly-A of the mRNA;

(V) synthesizing the cDNA by addition of a reverse transcriptase and NTPs; during this step, the new synthesized cDNA displaces the linker first strand (that one comprising the fixed and the variable portion);

(VI) removing the mRNA and obtaining a single strand cDNA.

Further, a primer can be added to the 3' end of the cDNA, and in presence of a polymerase a complementary DNA is synthesized forming a double strand cDNA. The double strand formed therefore comprises a first restriction enzyme site at one end and a second restriction enzyme site at the other end.

IDT Exhibit 1020.6817

The removal of mRNA at step VI) may be performed by addition of a RNase H or other enzyme which cuts RNA in fragments and remove them, or by addition of alkali (for instance NaOH) according to the methodology known in the art (Sambrook et al, 1989).

The double strand polynucleotide sequence is then cleaved at the first and second restriction enzyme sites, specifically introduced with the linkers, by using the specific restriction enzymes, resulting in forming protruding ends. The double strand polynucleotide with protruding ends is then inserted in a plasmid or phage expression vector or in a sequencing vector (for example, as described in, for example, Sambrook et al, 1989, Molecular Cloning, Cold Spring Harbor Laboratory; Invitrogen Catalog 1999; Stragene Catalog 1999, etc.). The double strand polynucleotide can also be cloned by site-specific recombination (for example attB-attP) or by blunt-ends methodology (Sambrook et al., 1989).

Examples of phage vectors are lambda-ZAP, lambda-Dash (Stratagene).

The invention, therefore, is also related, but not limited, to a phage or plasmid expression or sequencing vector comprising the polynucleotide sequence according to the present invention.

The single or double strand polynucleotide according to the present invention is RNA or DNA, or also a DNA/RNA hybrid. Preferably, a long strand full-coding and/or

IDT Exhibit 1020.6818

full-length cDNA. The 3' end of said long strand full-coding/full-length cDNA corresponds to the 5' Cap end of mRNA.

For the purposes of the present invention, with the wording full-length cDNA is intended a cDNA comprising the 5' and 3' UTR sequences and the oligo dT-primer (that is, complementary to a mRNA comprising the poly-A). It may also comprises additional sequences for cloning, such as restriction enzyme sites. With a full-coding cDNA, a cDNA sequence comprises at least the start and stop codon. And with long strand cDNA, it is understood a cDNA sequence which is almost full-coding/full-length, lacking of one or few nucleotides at the 3' end (corresponding to the 5' end of mRNA) or at the 5' end if considering a cDNA strand complementary to the cDNA complementary to the mRNA (that is, having the same direction of the gene). Such a stop of synthesis reaction during cDNA synthesis may be caused by the formation of secondary structure of the mRNA, for example, the level of the Cap structure. However, also fragments of genes, nucleotides, cDNAs, RNA or mRNA are not excluded from the purpose of the application of the present invention.

A DNA/RNA hybrid can be prepared by:

(I) providing a long strand, full-coding or full-length mRNA comprising a poly-A;

(II) providing a double strand linker comprising a first restriction enzyme site;

(III) ligating the 5' end of the mRNA (by using a ligase, for instance RNA ligase) to the fixed portion of the second strand of the linker and annealing the 5' end to variable portion of the first strand of the linker;

IDT Exhibit 1020.6819

(IV) providing an oligo dT-primer comprising a second restriction enzyme site, and annealing the oligo dT-primer to the poly-A of the mRNA;

(V) synthesizing the cDNA by addition of a reverse transcriptase and NTPs; during this step, the new synthesized cDNA displaces the linker first strand (that one comprising the fixed and the variable portion);

(VI) adding an oligonucleotide complementary to the second restriction enzyme site of the oligo dT-primer and ligated this oligonucleotide to the poly-A; an hybrid double strand polynucleotide is then formed.

The hydrid double strand polynucleotide can be cleaved by specific restriction enzymes as above said and inserted into a vector as above.

The target single strand polynucleotide, which anneals the variable linker portion and/or ligates to the adjacent fixed linker portion of the linker, can be prepared with any technique known in the prior art.

Preferably, the long strand full-coding/full-length single strand cDNAs are prepared according the technique of 5' mRNA Cap trapping, disclosed in Carninci et al., 1996, Genomics, 37, 327-336; Carninci et al., 1997, DNA Research 4:61-66; Carninci et al., 1998, Proc.Natl.Acad.Sci USA, 95:520-4; Carninci and Hayashizaki, 1999, Methods Enzymol. 303:19-44.

IDT Exhibit 1020.6820

Preferably, all the steps described in the above prior art documents are followed, with the exception that instead of the G-tailing step, the population of linkers according to the invention is provided.

Preferably, the Cap-trapping method described in Figures 1 and 2 is used, however, the target single strand polynucleotide is not limited to that prepared with this technology. For example, other methods of isolation of first strand cDNA such as that described in Edery et al., 1995, Mol Cell Biol, 15:3363-71 or the oligo-capping method (K.Maruyama et al., 1995, Gene, 138:171-174; and S.Kato et al., 1995, Gene, 150:243-250) can also be used.

The target single strand polynucleotides according to the present invention can also be normalized and/or subtracted (for example, Soares et al., 1994, Proc. Natl. Acad. Sci. 91:9228-9232 and Bonaldo et al., 1996, 6:791-806). The recovered normalized and/or subtracted polynucleotides, preferably cDNAs, more preferably long strand full-coding/full-length cDNAs, are preferably prepared according to the Cap-trapping technology, and then ligated to the population of linkers according to the present invention, or said isolated cDNAs are first annealed and/or ligated to the population of linkers of the present invention, and then normalized and/or subtracted.

The target single strand polynucleotide may show a bias due to the formation of a loop or hairpin-loop. For example, the 3' end of a synthesized intercellular cDNA

IDT Exhibit 1020.6821

may form a loop with an internal portion of itself, preventing the following annealing and ligation with the linker according to the present invention.

In order to solve this problem, the target single strand polynucleotide is optionally subjected to high temperature, from 25°C to up the boiling point of the solution (about 100°C), preferably at 65°C, and then cooled down, preferably in ice, before annealing and/or ligation with the linker according to the present invention.

As a modified method, the secondary structure can be deleted with chemical agents, such as solutions consisting of NaOH (for example 0.1 N), formamide 50-99% and Urea 6-8M or similar agents known to delete/reduce the secondary structure of nucleic acids or denature the double strand nucleic acids. In this case, such agents must be removed, usually by ethanol precipitation, prior to subsequent enzymatic reactions.

As a further modified method, the target polynucleotide annealed and/or ligated to the linker can be subjected to high temperature (hot start) in order to remove possibility of hairpin-loop formation. The temperature range is from 25°C up to the boiling point of the solution (about 100°C), preferably 65°C.

However, the increase of temperature may remove one strand of the linker (that is, the strand comprising the fixed portion and the variable portion), therefore later the same strand linker or any primer can be added to the other strand of the linker and the polynucleotide sequence comprising the target single strand polynucleotide. There is a

IDT Exhibit 1020.6822

possibility of the formation of hairpin-loop, but it can be avoided using the solutions described in Figure 8, B and C.

Using the method according to the invention, the annealing and ligation step result to be very efficient, so that the following cloning step allows the preparation of high-titer libraries without PCR amplification.

The method according to the invention allows the preparation of libraries more advantageously compared to the method in the prior art and in particular to the method of G-tailing.

The G-tailing method in fact has a serious drawback during the sequencing process, as shown in Fig. 5. The second strand cDNA comprises a repetition of C, complementary to the G-tail sequence (the length of which cannot be easily controlled and therefore may reach the length of 20-30 G). This excessive C strength repetition makes the sequencing process to stop, preventing the DNA sequencing.

The method using the linker according the invention does not have this drawback (even if the random variable portion comprises G, they are statistically within a small number) and can allow an efficient sequencing as described in Figures 6 and 7.

The clone of Figure 6 comprises a portion of a linker N6 (marked in the box of Fig.6) corresponding to nucleotide 12 to 49 of SEQ ID NO:3. The nucleotides 1 to 11

IDT Exhibit 1020.6823

(included) were cleaved as shown at step G of Figure 2. The variable portion of linker of Fig. 6 is GGCGAA (as shown in the marked box).

The clone of Figure 7 comprises a portion of a linker GN5 (marked in the box of Fig. 7) corresponding to nucleotide 12 to nucleotide 49 of SEQ ID NO:1. The nucleotides 1 to 11 (included) of SEQ ID NO:1 were cleaved as shown at step G) of Figure 2. The variable portion of linker of Fig.7 is GGCGAA (as shown in the marked box).

Then, the long G-stretches of the G-tailing methodology may interact with surrounding sequences and form very strong secondary structures, and this phenomenon affects the efficiency of sequencing, transcription and translation. On the contrary, the linker according to the present invention does not have these drawbacks.

Further, terminal deoxynucleotidyl transferase used for G-tailing reaction requires the presence of heavy metals, like $MnCl_2$ or $CoCl_2$. These heavy metals cause degradation of cDNAs and decreased long strand full-coding/full-length cDNA content rate. Also this problem is solved using the linker according to the present invention, which does not require heavy metals and can be performed at low temperature, for instance, 4-37°C, preferably 12-20°C, or preferably 16°C.

According to another embodiment of the present invention, the constant portion of the linker of the present invention can comprises a marker. For example, a specific

IDT Exhibit 1020.6824

oligonucleotide sequence, a specific sequence or combination of sequences, easily recognizable.

The presence of this marker is very useful in order to distinguish and not to confuse libraries of different tissues (for instance, liver, brain, lungs, and the like) for the same or for different species, or libraries of different species (for instance, human, mouse, Drosophila melanogaster, rice, and the like).

In fact, when many kinds of libraries obtained from different tissues and/or species are constructed in the same laboratory and used for large-scale sequencing, there is the risk of confusing or contaminating the libraries or clones at any stage of colony picking, DNA preparation, sequencing determination, clones banking, re-arraying, etc.

Individual marking of cDNAs allows the preparation of different marked cDNAs from several tissues, allowing tissue expression profiling by sequencing 3' ends (complementary to the 5' mRNA end) of mixed cDNA libraries.

[Example]

The method and embodiments according to the present invention will now be illustrated with reference to the following examples.

IDT Exhibit 1020.6825

Example 1

Linker evaluation using a Test cDNA

Linkers preparation

The population of linker oligonucleotides were purchased by Gibco-BRL Life technologies. The oligonucleotides were distinguished in one single strands (single upper strands) (indicated as A and C, comprising the variable portion) and the other single strands (single lower strands) (indicated as B). Then, one of A and C and the B were bound together in order to form two different population of double strands. The population of linkers A comprises linkers having a fixed portion oligonucleotide (in this case, bases 1-43 of SEQ ID NO:1) and a variable portion (GN_5), wherein the first base is G (that is, the base number 44) and the following bases NNNNN (bases from 45 to 49) different for each linker of the population and prepared at random.

The population of linkers C, comprises a constant portion oligonucleotide (bases 1-43 of SEQ ID NO:3) and the variable portion NNNNNN (bases 44 to 49) different for each linker of the population and prepared at random.

A) GN_5 A strand,

5'-AGAGAGAGAGCTCGAGCTCTATTTAGGTGACACTATAGAACCAGNNNNN-3' (SEQ ID NO:1);

B) B strand,

5'-TGGTTCTATAGTGTCACCTAAATAGAGCTCGAGCTCTCTCTCT-3' (SEQ ID NO:2);

IDT Exhibit 1020.6826

The B strand was also phosphorylated at the 5' end when it was synthesized.

C) N_6 C strand,

5'-AGAGAGAGAGCTCGAGCTCTATTTAGGTGACACTATAGAACCANNNNNN-3'

(SEQ ID NO:3).

For degenerate nucleotides, V stands for A, G or C and N stands for any nucleotide, according to the international convention and to the Patentin Standard 2.1 Manual.

These oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis (Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.) in order to remove contaminants that may have non-specific or lack annealing sites. Two populations of linkers named GN_5 and N_6 were prepared. The linker GN_5 was made by oligonucleotides A/B (SEQ ID NO:1/SEQ ID NO:2) and the linker N_6 was made by oligonucleotides C/B (SEQ ID NO:3/SEQ ID NO:2). They were prepared by mixing the oligonucleotides with NaCl (final concentration, 100 mM) and incubating at 65℃ for 5 min, 45℃ for 5 min, 37℃ for 10 min and 25℃ for 10 min.

The linkers prepared were then used for annealing to and ligating with single strand DNA(s).

IDT Exhibit 1020.6827

Test cDNA

To establish the appropriate linker when preparing cDNA libraries, a test first strand cDNA has been generated from 5 μg 7.5-kb poly(A)-tailed RNA (Life Technologies) according to the method described in Carninci and Hayashizaki, 1999, except that CAP-Trapping was omitted. [α-^{32}P]dGTP was incorporated at the reverse-transcription step. The amount of the produced first strand cDNA was estimated according to the incorporation ratio of radioactivity. Then, 50 ng of the 7.5-kb cDNA and various amounts (200 ng to 2 μg) (see also description of Figure 3) of linker (N_6 or GN_5), prepared in the above step of Example 1, were combined together and ligated in 30 μL reaction volume. The reactions were performed by the incubation overnight at 10° C.

After ligation, to remove excess linkers, linker-bound single strand cDNA samples were incubated with 0.2 mg/mL proteinase K in 10 mM EDTA/0.2% SDS (reaction volume, 40 μL) at 45° C for 15 min. The reaction products were extracted by using phenol/chloroform 40 μL. Then, the phenol/chloroform mixture was treated with 60 μL column buffer (10 mM Tris-HCl, 1 mM EDTA, 0.1 M NaCl, 0.1% SDS; pH 7.5) in order to extract the reaction product which was still remaining in the interface of the phenol/chloroform mixture. The reaction products extracted were loaded on gel-filtration columns Sephacryl™-300 (Amersham Pharmacia Biotech) and purified by centrifugation at 400 x g for 2 min. The eluted fractions (comprising the purified linker-first strand cDNA samples) were precipitated by using isopropanol.

IDT Exhibit 1020.6828

The control sample (comprising the single strand Test cDNA but no linker) was used for the synthesis of the second strand cDNA. To assess the ability of the present method to support the synthesis of second-strand cDNA, 10 ng of the purified linker-ligated samples (Lanes 1-6 of Figure 3) and unligated 7.5-kb first-strand cDNA as a control (Lane 7 of Figure 3) were independently combined in a 10-μL reaction solution containing 1 μL 10X ExTaq™ buffer, 1 μL 2.5 mM dNTPs, 0.5 μL [α-^{32}P]dGTP, and 0.5 μL Ex-Taq™ (Takara). The obtained samples were incubated at 65° C for 5 min, 68° C for 30 min, and 72° C for 10 min and then analyzed by using alkaline gel electrophoresis.

The alkaline gel electrophoresis was performed by adding 5 μ1 of the samples to 1 μ1 of 6 x Alkaline dye (Sambrook, Molecular Cloning, 6.7, 6.12). Used were the electrophoresis gel contained 0.8 % of agarose, 50 mM NaOH and 5 mM EDTA and the buffer contained 50 mM NaOH and 5 mM EDTA (Sambrook, Molecular Cloning).

Results

Lanes 1-3 indicates the ligation of 50 ng between single strand Test cDNA and respectively 200 ng, 500 ng and 2 μg of linker GN_5.

Lanes 4-6 indicates the ligation between 50 ng of single strand Test cDNA and respectively 200 ng, 500 ng and 2 μg of linker N_6.

Lane 7 is the control. 10 ng of single first strand cDNA were added to 10-μL reaction solution comprising 1 μL 10X ExTaq™ buffer, 1 μL 2.5 mM dNTPs, 0.5 μL [α-^{32}P]dGTP, and 0.5 μL Ex-Taq™ (Takara). The sample was incubated at 65° C for

IDT Exhibit 1020.6829

5 min, 68° C for 30 min, and 72° C for 10 min. The single strand cDNA was extended by forming a hairpin structure to form a second strand cDNA. This was detected at 15 kb in the alkaline gel electrophoresis.

Lane 8 comprises the first strand cDNA (23 ng) (without linker). This is detected at 7.5 kb.

Lane 9 represents the markers.

The electrophoresis of Figure 3 shows that in Lanes 1-6, the ligation was particularly efficient (spots at level of 7.5 kb) and the amount of no ligation was negligible (spots at level of 15 kb).

Example 2

Full-length cDNA library preparation and cDNA analysis

Linker preparation

Linkers were prepared as above described in Example 1.

Preparation of RNA

Slices of mouse liver tissue (0.5-1g) were homogenized in 10 ml of a suspension and extracted with 1 ml of 2M sodium acetate (pH 4.0) and the same amount of a mixture of phenol/chloroform (volume ratio 5:1). After the extraction, the same volume of isopropanol was added to the aqueous layer to precipitate RNA. This sample was incubated on ice for an hour and centrifuged at 4000 rpm for 15 minutes with cooling to collect the precipitates. The resulting precipitates were washed with 70% ethanol and dissolved in 8 ml of water. By adding 2 ml of 5M NaCl and 16 ml of an aqueous solution (pH 7.0) containing 1% CTAB (cetyltrimethylammonium bromide), 4M urea

IDT Exhibit 1020.6830

and 50 mM Tris, RNA was precipitated and polysaccharides were removed (CTAB precipitate). After centrifugation at 4000 rpm for 15 minutes at room temperature, the RNA was dissolved in 4 ml of 7M guanidine-Cl. Then, two-fold amount of ethanol was added to the solution, incubated for an hour on ice and centrifuged at 4000 rpm for 15 minutes. The resulting precipitates were washed with 70% ethanol and collected. The precipitates were again dissolved in water and purity of RNA was determined by measuring OD ratio 260/280 (>1.8) and 230/260 (<0.45). The total RNA thus obtained was then purified by using the mRNA isolation kit for total RNA MACS™ (Miltenyi Biotech, Germany) and those containing poly A+ were concentrated.

cDNA synthesis

5 to 10 μg of this polyA+ rich RNA, 5 μg of the first-strand primer containing an BamHI site 5'-$(GA)_5$AGGATCCAAGAGCTC$(T)_{16}$VN-3') (SEQ ID NO:4) and 11.2 μl 80% glycerol have been combined in a total volume of 24 μl. RNA/primer mixture was denatured at 65°C for 10 min. In parallel, we combined, in a final volume of 76 μl, 18.2 μl 5X first-strand synthesis buffer, 9.1 μl 0.1 M DTT, 6.0 μl 10 mM (each) dTTP, dGTP, dATP, and 5-methyl-dCTP (instead of dCTP), 29.6 μl saturated trehalose (approximately 80%, low metal content; Fluka Biochemika), and 10.0 μl Superscript II reverse transcriptase (200 U/μl). We placed 1.0 μl [α-^{32}P]dGTP in a third tube. The mRNA, glycerol, and primers were mixed on ice with the solution containing the Superscript, and an aliquot (20%) was quickly added to the tube containing the [α-^{32}P]dGTP. First-strand cDNA syntheses were performed in a thermocycler with a heated lid (e.g., MJ Research) according to the following program: step 1, 45°C for 2

IDT Exhibit 1020.6831

min; step 2, gradient annealing: cool to 35°C over 1 min; step 3, complete annealing: 35°C for 2 min; step 4, 50°C for 5 min; step 5, increase to 60°C at 0.1°C per second; step 6, 55°C for 2 min; step 7, 60°C for 2 min; step 8, return to step 6 for 10 additional cycles. Incorporation of radioactivity allowed the estimation of the yield of cDNA (Carninci and Hayashizaki, 1999). The cDNA was treated with proteinase K, phenol/chloroform and chloroform-extracted, and ethanol-precipitated by using ammonium acetate as the salt (Carninci and Hayashizaki, 1999).

mRNA biotinylation

Before biotinylation, the diol group of the cap and 3' end of mRNA was oxidized in a reaction solution in a final volume of 50 μ l, containing the resuspended mRNA/cDNA comprising first-strand cDNA, 66 mM sodium acetate (pH 4.5), and 5 mM $NaIO_4$. Samples were incubated on ice in the dark for 45 min. mRNA/cDNA hybrids were then precipitated by adding 0.5 μl of 10% SDS, 11 μl NaCl, and 61 μl of isopropanol. After incubation in the dark on ice for 45 min, the sample was centrifuged for 10 min at 15,000 rpm. Finally the mRNA/cDNA hybrids were rinsed twice with 70% ethanol and resuspended in 50 μl of water. Subsequently, the cap was biotinylated in a final volume reaction solution of 210 μl by adding 5 μl 1 M sodium acetate (pH 6.1), 5 μl 10% SDS, and 150 μl of 10 mM biotin hydrazide long-arm (Vector Biosystem). After overnight (13 hours) incubation at room temperature, the mRNA/cDNA hybrids were precipitated by adding 75 μl 1 M sodium acetate (pH 6.1), 5 μl 5 M NaCl, and 750 μl absolute ethanol and incubated on ice for 1 hour. The mRNA/cDNA hybrids were pelleted by centrifugation at 15,000 rpm for 10 min; then the pellet was washed once with

IDT Exhibit 1020.6832

70% ethanol and once with 80% ethanol. The mRNA/cDNA hybrids were then resuspended in 70 μl 0.1X TE (1 mM Tris [pH 7.5], 0.1 mM EDTA).

Adsorption and release of full-length cDNA

500 μl of MPG-streptavidin beads and 100 μg DNA-free tRNA were combined and the obtained mixture incubated on ice for 30 min with occasional mixing. The beads were separated by using a magnetic stand for 3 minutes, and the supernatant was removed. The beads were then washed three times with 500 μL washing/binding solution (2 M NaCl, 50 mM EDTA [pH 8.0]).

At the same time, 1 unit of RNase I (Promega) per 1 μg of starting material mRNA was added to the mRNA/cDNA hybrid sample in the buffer attached to the product (final volume, 200 μl); the sample was incubated at 37°C for 15 min. To stop the reaction, the sample was put on ice and 100 μg tRNA and 100 μl of 5 M NaCl were added. To adsorb the full-coding/full-length mRNA/cDNA hybrids, the biotinylated, RNase I-treated mRNA/cDNA and the washed beads, which were resuspended in 40 μl of the washing/binding solution were combined. After mixing, the tube was gently rotated for 30 min at room temperature. Full-coding/full-length cDNA was adsorbed on the beads, and the shortened cDNAs did not. The beads were separated from the supernatant with a magnetic stand. The beads were gently washed to remove the nonspecifically adsorbed cDNAs. Two washes with washing/binding solution were performed: one with 0.4% SDS, 50 μg/ml tRNA; one with 10 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, 40 μg/ml tRNA, 10 mM NaCl, and 20% glycerol; as well as with 50 μg/ml

IDT Exhibit 1020.6833

tRNA in water.

The cDNA was released from the beads by adding 50 μl 50 mM NaOH, 5 mM EDTA and incubating for 10 min at room temperature with occasional mixing. The beads then were removed magnetically, and the eluted cDNA was transferred on ice to a tube containing 50 μl 1 M Tris-HCl, pH 7.0. The elution cycle was repeated once or twice with 50 μl-aliquots of 50 mM NaOH, 5 mM EDTA until most of the cDNA (80 to 90%, as measured by monitoring the radioactivity with a hand-held monitor) were recovered from the beads.

To remove traces of RNA, 1 μl RNase 1 (10U/μl) to the recovered cDNA on ice was quickly added; the sample was then incubated at 37°C for 10 min. The cDNA was treated with proteinase K, and then phenol/chloroform-extracted, and back-extracted. Then, the samples were concentrated by using one round of ultrafiltration with a Microcon 100 (Millipore) for 40-60 min at 2000 rpm.

CL-4B spin-column fractionation of cDNA

The cDNA samples were then treated with CL-4B chromatography (Carninci and Hayashizaki, 1999) according the manual (S-400 spin column, for example of Amersham-Pharmacia, can also be used).

cDNAs-linker ligation

Cap-Trapper full-length single strand cDNAs, prepared as above, were divided

IDT Exhibit 1020.6834

in three different tubes. One for G-tailing, the second one for GN_5 linker and the last one for N_6/GN_5 mixed linker. An aliquot of 200 ng of cDNA were tailed with dG homopolymer as described in the prior art and used for the control cDNA library preparation (Carninci et al., Genomics, 1996).

300 ng of Cap-Trapper full-length first strand cDNAs were used as substrate for the linker-ligation using the linkers prepared as above by Gibco-BRL/Life Technologies, and cDNA libraries were constructed (shown in Figures 1 and 2).

The 300 ng of the single strand cDNA were added to 800 ng of a mixture of N_6/GN_5 linkers at proportion 1:4, and to 800 ng of GN_5 linker.

Ligation substrates (the cDNA/linker prepared as above), Solution I and Solution II (Ligation Kit, Takara) were mixed in ratio 1:2:1 and all the process were performed as described in the manual attached to the product. The reaction volume therefore was of 30 μl and contained 7.5 μl of sample, 15 μl of Solution I and 7.5 μl of Solution II. The reaction run overnight at 10℃ (Fig. 2E).

Isolation from excess linkers.

After annealing and ligation between cDNA and linker, gel filtration was carried out. 30 μl of linker-ligation samples, as above, were treated with 0.2 mg/ml proteinase K in the presence of 10 mM EDTA and 0.2% SDS. They were incubated at 45℃ for 15 min, followed by the phenol/chloroform extraction. The samples were back extracted

IDT Exhibit 1020.6835

with 60 μ1 of column buffer (10 mM Tris-HCl, 1 mM EDTA, 0.1 M NaCl, 0.1% SDS, pH 7.5). Subsequently, the samples were subjected to the spun column for gel filtration Sephacryl S 300 (Amersham Pharmacia Biotech). In the step for the spun column, the centrifugation was carried out at 400x g for 2 min. The eluted fraction was recovered and precipitated with isopropanol.

After purification step, the second strand cDNA synthesis was carried out (Fig. 2F).

To synthesize the second-strand cDNA, all purified linker-ligated samples were used. 6 μ1 of 10 x ExTaq buffer Takara, and 6 μ1 of 10 mM dNTPs and 0.5 μ1 [α-^{32}P]dGTP were added to the tubes in 60 μ1. The samples were pre-incubated at 72℃ for 15 sec and then 0.5 μ1 of ExTaq were added. Then they were incubated at 72℃ for 30 min.

The samples were analyzed by alkali gel electrophoresis, that is, 0,5 μ1 of the samples comprising the synthesized second-strand were added to 1 μ1 of 6 x Alkaline dye (Sambrook, Molecular Cloning, 6.7, 6.12) in final volume of 6 μ1 and the electrophoresis was performed.

The electrophoresis is performed using an agarose gel containing 0.8 % of agarose, 50 mM NaOH, 5 mM EDTA and electrophoresis buffer containing 50 mM NAOH and 5 mM EDTA (Sambrook, Molecular Cloning).

IDT Exhibit 1020.6836

The samples were purified with phenol/chloroform followed by ethanol precipitation under standard condition (Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Mannual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.).

Subsequently, cDNA was cleaved with *Bam* HI (25 U/μg of cDNAs) and *Xho* I (25 U/μg of cDNAs) at 37℃ for 1h and extracted with phenol/chloroform. The aqueous phase was purified with CL4B gel filtration spin column (Amersham Pharmacia Biotech) followed by ethanol precipitation as described in presence of 2 μg of glycogen (Carninci and Hayashizaki, 1999).

Construction of pBS IV vector

10 ng pBS SK+ (stratagene), 20μl 10x NEB buffer 2 (New England Biolabs, Inc), 20μl 10 mg/ml bovine serum albumin (NEB), 30 units Not I (NEB), 30 units Kpn I(NEB) and 10 units Xho I(NEB) were mixed in the volume of 200μl and incubated at 37 °C for 2 hours. Then, this mixture was subjected to electrophoresis at 50V for 1 hour on 0.8% SeaPlaque. agarose gel (FMC Bioproducts)/1x TAE buffer/0.5μg/ml ethidium bromide (8cm x 8cm) in 1x TAE/0.5μg/ml ethidium bromide buffer to separate a long plasmid part from a short DNA section (Molecular Cloning). The long plasmid part was cut out from the gel and the gel was transferred to a tube. Cleaved plasmid was extracted and purified by the GENECLEAN II™ Kit (Bio 101 Inc.). Concentration and purity of the plasmid were checked by agarose gel electrophoresis, comparing with standard plasmid, which concentration was already known.

IDT Exhibit 1020.6837

Double strand oligonucleotide preparation

The oligonucleotides used were custom-synthesized (Life Technologies-Life Tech Oriental, Tokyo, Japan) and then purified by using denaturing polyacrylamide gel electrophoresis in order to remove contaminants (Maniatis etc.). The Not/Kpn double strand oligonucleotide was prepared by mixing the following two single strand oligonucleotides:

one strand (Upper-strand)

(5'GGCCGCATAACTTCGTATAGCATACATTATACGAAGTTATGGATCAGGCCAAA TCGGCCGAGCTCGAATTCGTCGACGAGAGACTGCAGGAGAGAGGATCCGGTA C-3')(SEQ ID NO:6); and

the other strand (Lower-strand)

(5'CGGATCCTCTCTCCTGCAGTCTCTCGTCGACGAATTCGAGCTCGGCCGATTT GGCCTGATCCATAACTTCGTATAATGTATGCTATACGAAGTTATGC -3')(SEQ ID NO:7) in NaCl (final concentration, 100 mM).

This mixture was then incubated at 65 ° C for 5 min, 45 ° C for 5 min, 37 ° C for 10 min, and 25° C for 10 min.

Vector-Oligonucleotide ligation

100 ng of this plasmid, which has Kpn I and Not I sites at the end were mixed with 3 ng of a Not/ Kpn double strand oligonucleotide, 1 μ1 10 x ligation buffer (NEB) and T4 DNA ligase (NEB) in 10 μ1.

IDT Exhibit 1020.6838

Cell transformation

A tube comprising the ligation sample was then incubated overnight at 16 °C. The ligation sample was mixed with 250 mM NaCl, 1 μg glycogen and isopropanol, then precipitated to remove buffer. Then, it was dissolved with 10 μl sterile water. 1 μl of the obtained sample was supplied to transform to a suspension of *E.coli* cells DH10B (Life tech oriental) by electroporation (following the Protocol of the Manufacturer). The transformed cell was selected on LB plate containing 100 μg/ml ampicillin. The ampicillin-resistant clone was cultured in the LB liquid medium containing 100 μg/ml ampicillin at 37 °C for 16 h with shaking. The recombinant plasmid used for the insertion of the synthesized double strand oligonucleotide was purified by alkali-SDS method (Molecular Cloning). The inserted sequence was confirmed with the M13 forward primer (SEQ ID NO:5) and the Big dye kit by ABI377 DNA sequencer (PE-Applied BioSystems).

Vector preparation

10 μg of the modified pBS SK (+) plasmid obtained as above (named pBS IV) were mixed with 20 μl 10x Bam HI buffer, 20 μl 10 mg/ml bovine serum albumin (New England Biolabs, Inc), 30 units BamH I (New England Biolabs, Inc), 30 units Sal I(New England Biolabs, Inc), adjusted to a volume of 200 μl and incubated at 37 °C for 1.5 hours. Then, 10 units Pst I (New England Biolabs, Inc) were added to the mixture in a tube and incubated for 30 min. Furthermore, dephosphorylation of the plasmid end was carried out by 0.5 units thermo-sensitive alkaline phosphatase TsAP (Life Technologies).

IDT Exhibit 1020.6839

TsAP makes background depending on partial cut plasmid low. Dephosphorylated ends cannot ligate each other. The tube was incubated at 37 °C for 30 min. To inactivate TsAP, EDTA (final concentration of 20 mM) was added and incubated at 65 °C for 30 min. The restrict enzyme/TsAP treated plasmid was separated as described above in the vector construction step. The band corresponding to linear plasmid, which has Bam HI and Sal I site at the end, was cut out from the gel and sliced to small pieces. It was put in a tube containing 500 μl 1x β agarase buffer (NEB) and left on ice for 30 min. The buffer was changed once and left on ice more 30 min.

This tube was incubated at 65 °C for 10 min to melt the gel. β-agarase buffer was added to bring the solution to 100 μl. Then, it was cooled at 45 °C for 3 min and β-agarase (NEB) was added to the concentration of 3U/100 μl reaction. This reaction solution was incubated at 45 °C for 6 h. 10 μl of 5M NaCl and 100 μl of phenol/chloroform were added to the tube. The tube was inverted gently for 5 min and centrifuged at 15 krpm for 3 min at room temperature. The aqueous phase was recovered and followed by chloroform extraction and isopropanol precipitation. The tube was centrifuged at 15 krpm for 10 min at 4 °C and the obtained pellet was washed with 80% ethanol twice. Finally, the pellet was dissolved with sterile water to a final concentration of 100 ng/μl. The vector concentration and purity were checked by agarose gel electrophoresis, comparing with standard plasmid, which concentration was already known.

IDT Exhibit 1020.6840

Cloning

10 ng of cDNA, obtained in the previous step, were ligated overnight to 190 ng of the aforementioned modified vector pBluescript KS (+)(Stratagene).

The cDNA-vector ligations were precipitated with EtOH 2.5 times volume. The samples were introduced into *E.coli* DH10B (Gibco BRL) by electroporation. The transformed cells were applied on LB plate containing 100 μg/ml ampicillin and cultured overnight at 37℃. 36 colonies were picked up randomly and cultured in LB ampicillin (100 μg/ml) liquid medium overnight at 37℃. Recombinant plasmids were extracted from 3 cultures (Sambrook et al., 1989). These 3 purified plasmids were sequenced from 5' end with the M13 forward primer TGTAAAACGACGGCCAGT (SEQ ID NO:5) with the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE – ABI) by using the ABI3700 DNA sequencer (PE-Applied BioSystems) according to the Kit Manual Instruction.

Example 3

Efficiency of ligation

In the ligation of linkers prepared as above to the mouse liver derived cDNAs, 2 μg of mixed linker (N_6: GN_5 = 1: 4) were ligated against various amounts of cDNAs (1 μg, 0.5 μg and 0.2 μg respectively in Lanes 1, 2 and 3 of Figure 4). Then, 50 ng of ligated cDNAs were used for the second strand synthesis and analyzed by alkaline gel electrophoresis. All electrophoresis patterns and incorporation-rate were the same (Lanes 1-3), suggesting that 2 μg of linker were efficiently ligated to any of the different amount of cDNAs. If the linker amount was not appropriate, excessively expressed

IDT Exhibit 1020.6841

cDNA bands would be shifted up to twice size by forming hairpin structure. Instead, the first-strand cDNA and all second-strand cDNAs showed the same pattern (the same size).

Example 4

Efficiency of linker-ligation full-length cDNAs preparation

Liver mouse cDNA libraries were prepared in the same way as above described using the CAP-trapper technology.

Whether those prepared by the linker method as described in the above example are full-length cDNA was checked by confirming the presence of ATG starting codon after sequencing step. In fact, those cDNAs containing the full-coding sequence from the starting ATG were accepted as full-length cDNAs. The 5' sequences were compared with the public nucleotide database using BLAST (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J., 1990, "Basic local alignment search tool." J. Mol. Biol. 215:403-410).

Nucleotide Sequences were determined using Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE – ABI) and the Perkin Elmer-Applied Biosystems ABI 3700 according to the Kit Manual Instructions.

Sequencing primer used is the M13 primer on the 5' side (SEQ ID NO:5).

The data are reported in Table 1. The presence of ATG (from which the transcription starts) has been marked with the corresponding position. For example, with reference to

IDT Exhibit 1020.6842

sample 2.01, the adenosine of ATG codon has the position 63; this indicates that this cDNA sequence has 62 bp 5'-UTR.

TABLE 1

Sample code	Gene	Public db ID	RIKEN db ID	Existence and position of ATG
1) N6/GN5 mixed linker (1:4)				
2.01	*M. musculus* alpha fetoprotein	gi|6680667	ZX00047P08	63
2.05	Human cDNA FLJ10604 fis	gi|7022741	ZX00036I09	172
2.22	Mouse mMCM2	gi|2381484	——	37
2.25	*M. musculus* epsilon 14-3-3 isoform	gi|57965	——	104
2.26	*M. musculus* SH2-containing inositol phosphatase SHIP	gi|1255351	——	105
2.28	*M. musculus* calmodulin 3	gi|6680833	R000011E16	176
2.30	*M. musculus* EAT/MCL-1	gi|2522268	——	61
2.31	*M. musculus* heterogenous nuclear ribonucleoprotein U	gi|3329495	——	199
2.33	*H. sapiens* hDj9	gi|6567165	ZX00052D18	166
2.36	*M. musculus* Lsc (lsc) oncogene	gi|1389755	ZX00048L03	214
2) GN5 linker				
3.07	*M. musculus* ornithine decarboxylase antizyme	gi|1279827	ZX00047L07	81
3.09	*R. norvegicus* guanosine monophosphate reductase	gi|3907578	R000001H02	131
3.10	Mouse calmodulin	gi|469421	R000011E16	192
3.13	*M. musculus* protein phosphatase 5	gi|2407636	——	89
3.16	*Rattus sp.* calcium-regulated heat stable protein CRHSP-24	gi|4583308	R000016L20	59
3.20	Murine mRNA with homology to yeast L29 ribosomal protein	gi|50320	ZX00047I16	30
3.21	*M. musculus* ribosomal protein S3	gi|439521	ZX00048P23	39
3.23	*M. musculus* melanome X-actin	gi|6671508	ZX00035N19	88
3.29	*H. sapiens* CGI-47 protein	gi|4929562	ZX00048M14	199
3.33	*M. musculus* mini chromosome maintenance deficient 6	gi|6678831	——	131
3.34	*M. musculus* membrane protein TMS-1	gi|5853318	R000009P22	60

IDT Exhibit 1020.6843

These data show that cDNAs are efficiently prepared and sequenced using the linker methodology according to the present invention.

The sequencing of clone 2.05 of Table 1 is shown in the chart of Figure 6.

The sequencing of clone 3.07 of Table 1 is shown in the chart of Figure 7.

Advantages of the CAP-Trapping-linker versus the conventional G-tailing CAP-Trapping

I) Control of G-tail length has been difficult over years. To anneal the second strand primer, cDNA clones have at least 11 of dGs with an average of 13-15 dGs. Although G-tailing reaction is self-limiting to 15-30 nt (Hyone-Myong Eun, 1996, Enzymology Primer for Recombinant DNA Technology, page 477), G-stretches longer than about 20 bases, which were often obtained, caused a dramatic decrease of sequencing yield or in the worst case failure, while shorter G stretches impaired long sequence reading (see Figure 5). During sequencing, long G-stretches may interact with surrounding sequences and form very strong secondary structures. This may be problematic in case of interactions with 5' UTRs that are typically GC rich. This is especially serious in full-length cDNAs synthesis like in case of Cap-Trapping libraries.

The method using the linker according the present invention, different from conventional methods, does not have such a drawback (even if the random variable portion comprises G, there are statistically within a small number) and can allow an efficient sequencing (Figure 6 and 7).

IDT Exhibit 1020.6844

II) G-stretch is expected to affect the efficiency of translation in case of functional studies, for instance where protein expression is required, such as in expression cloning (King RW, Lustig KD, Stukenberg PT, McGarry TJ, Kirschner MW. Expression cloning in the test tube. Science. 1997;277:973-4.). On the other hand, the linker sequence of the present invention does not inhibit transcription and translation.

INDUSTRIAL APPLICABILITY

The present invention provides a novel and efficient method for the preparation of cDNA libraries. More specifically, the present invention provides a novel linker instead of G-tailing, which can be utilized in a method for the preparation of cDNA libraries and to provide a method for the preparation of cDNA libraries using said linker.

IDT Exhibit 1020.6845

CLAIMS

1. A linker or a population of linkers comprising an oligonucleotide fixed portion and an oligonucleotide variable portion, the variable portion being represented as Formula (N)n wherein N is A, C, G, T or U, or their derivatives, and n is an integer equal to or higher than 1, and if n is an integer equal to or higher than 2, the nucleotides (N) may be the same or different from each other.

2. The linker of claim 1, wherein the variable portion is prepared at random.

3. The linker according to claim 1, wherein n is an integer ranging from 1 to 10.

4. The linker according to claim 1, wherein n is an integer ranging from 4 to 8.

5. The linker according to claim 1, wherein n is 5 or 6.

6. The linker according to any one of claims 1 to 5, wherein the variable portion is $(G)_m(N)_{n-m}$, and m is an integer ranging from 1 to 3.

7. The linker according to any one of claims 1 to 6, wherein the variable portion is GN_4, GN_5, G_2N_3, G_2N_4, G_3N_2, G_3N_3, N_5 or N_6.

8. The linker according to any one of claims 1 to 7, wherein the oligonucleotide fixed portion is a single strand or double strand oligonucleotide.

9. The linker according to claim 8, wherein the fixed portion is a double strand oligonucleotide and the single strand variable portion (N)n protrudes outside the double strand fixed portion.

10. The linker according to any one of claims 1 to 9, which is a single strand oligonucleotide having the sequence of SEQ ID NO:1 or SEQ ID NO:3.

11. The linker according to any one of claims 1 to 9, which is a double strand

IDT Exhibit 1020.6846

oligonucleotide having the sequence SEQ ID NO:1/SEQ ID NO:2 or SEQ ID NO:3/SEQ ID NO:2.

12. A double strand oligonucleotide linker or a population of linkers, comprising

a) a first oligonucleotide single strand comprising an oligonucleotide single strand fixed portion and an oligonucleotide single strand variable portion, and

b) a second oligonucleotide single strand comprising an oligonucleotide single strand fixed portion annealed to the complementary first oligonucleotide single strand fixed portion (a),

so that the variable portion protrudes outside the double strand fixed portion of the linker; and wherein the variable portion is represented by the formula (N)n wherein N is A, C, G, T or U, or their derivatives and n is an integer equal to or higher than 1.

13. The linker of claim 12, wherein the variable portion is prepared at random.

14. The linker of claim 12, wherein n is an integer ranging from 1 to 10.

15. The linker of claim 12, wherein n is an integer ranging from 4 to 8.

16. The linker of claim 12, wherein n is 5 or 6.

17. The linker according to any one of claims 12 to 16, wherein the variable portion is (G)m(N)n-m, and m is from 1 to 3.

18. The linker according to any one of claims 12 to 17, wherein the variable portion is GN_4, GN_5, G_2N_3, G_2N_4, G_3N_2, G_3N_3, N_5 or N_6.

19. The linker according to any one of claims 12 to 18, which has the sequence SEQ ID NO:1/SEQ ID NO:2 or SEQ ID NO:3/SEQ ID NO:2.

20. A population of linkers comprising at least two linkers according to any one of claims 1 to 19.

IDT Exhibit 1020.6847

21. A population of linkers comprising at least two subpopulations of linkers according to any one of claims 1 to 20.

22. The population of linkers according to any one of claims 1 to 21, wherein the fixed portions comprised in all linkers are an oligonucleotide portion having the same sequence.

23. The population of linkers according to any one of claims 1 to 22, comprising two or more subpopulations of linkers, wherein one subpopulation of linkers comprise linkers in which the fixed portion is an oligonucleotide portion having the same sequence, and other subpopulations of linkers which differ each other in the fixed portion sequence.

24. The population of linkers according to any one of claims1 to 23, wherein the variable portion of the linkers is synthesized at random.

25. The population of linkers according to any one of claims 1 to 24, wherein the sequence of variable portion comprised in each linker differs from each other.

26. The population of linkers according to any one of claims 1 to 25, which is a mixture of the (N)n linker and the (G)m(N)n-m linker, represented as (N)n/(G)m(N)n-m mixture.

27. The population of linkers according to any one of claims 1 to 26, which is a mixture of two or more kinds of linkers having any variable portion selected from the group consisting of GN_4, GN_5, G_2N_3, G_2N_4, G_3N_2, G_3N_3, N_5 or N_6.

28. The population of linkers according to claim 27, which is a N_6/GN_5 mixture, N_6/G_2N_4 mixture or N_6/G_3N_3 mixture.

29. The population of linkers according to claim 28, wherein the mixing ratio ranges from 0:1 to 1:0.

IDT Exhibit 1020.6848

30. The population of linkers according to claim 29, wherein the mixing ratio is 1:4.

31. The linker or population of linkers according to any one of claims 1 to 30, which is prepared by;

(a) synthesizing a first oligonucleotide single strand comprising an oligonucleotide single strand fixed portion and an oligonucleotide single strand variable portion,

(b) synthesizing a second oligonucleotide single strand comprising an oligonucleotide single strand fixed portion complementary to the first oligonucleotide single strand fixed portion (a), and

(c) annealing the first oligonucleotide strand (a) to the second oligonucleotide strand (b) so that the variable portion protrudes outside the double strand fixed linker portion.

32. The linker or population of linkers according to any one of claims 1 to 31, wherein the fixed portion has one or more sequences of restriction enzyme sites, recombinational sites, RNA polymerase promoter sites, markers or tags.

33. The linker or population of linkers according to claim 32, wherein the restriction enzyme site is BamHI, XhoI, SstI, SalI or NotI.

34. The linker or population of linkers according to claim 32, wherein the recombinational site is attB, cre-lox or Gateway™.

35. The linker or population of linkers according to claim 32, wherein the polymerase promoter site is selected from the group consisting of T7, T3, K11, SP6 and BA14 RNA polymerase.

36. The linker or population of linkers according to claim 32, wherein the markers

IDT Exhibit 1020.6849

are specific sequences or plural oligonucleotides

37. The linker or population of linkers according to claim 32, wherein the tag binds to one or both ends of two strands of the fixed portion, opposite to the variable portion, and said end is a protecting group preventing the formation of binding loop between said linker and an annealed target single strand polynucleotide.

38. The linker or population of linkers according to claim 37, wherein the protecting group is a NH_2 group.

39. The linker or population of linkers according to claim 38, wherein the ends of two strands of the fixed portion of said linker, opposite to the variable portion, are bound together to form a loop so as to prevent said linker from binding to the annealed target single strand polynucleotide.

40. The population of linkers according to any one of claims 1 to 39, which is a population of linkers comprising at least one linker in which the variable oligonucleotide sequence is a specific oligonucleotide sequence, wherein the specific oligonucleotide sequence being able to specifically annealing to the end of at least one defined target polynucleotide to be selected from a population of target polynucleotides.

41. The linker or population of linkers according to any one of claims 1 to 40, which is DNA or RNA or a mixture thereof.

42. A linker-polynucleotide or population of linker-polynucleotides comprising the linker or population of linkers according to any one of claims 1 to 41 and the target first strand polynucleotide bound to the linker.

43. The linker-polynucleotide or population of linker-polynucleotides according to claim 42, wherein one end of the target first strand polynucleotide is annealed and/or

IDT Exhibit 1020.6850

ligated to the linker.

44. The linker-polynucleotide or population of linker-polynucleotides according to claim 42, wherein one end of the target first strand polynucleotide is annealed to the variable portion of the first strand of the linker and ligated to the second strand fixed portion of the linker.

45. The linker-polynucleotide or population of linker-polynucleotides according to any one of claims 42-44, wherein the linker is ligated to the target first strand polynucleotide with ligase.

46. The linker-polynucleotide or population of linker-polynucleotides according to claim 45, wherein the ligase is a DNA ligase or RNA ligase.

47. The linker-polynucleotide or population of linker-polynucleotides according to any one of claims 42 to 46, which further comprises a second strand polynucleotide complementary to the target first strand polynucleotide.

48. The linker-polynucleotide or population of linker-polynucleotides according to claim 47, wherein the second strand polynucleotide is synthesized from a free end of the variable portion, with the variable portion acting as a primer.

49. The single strand or double strand polynucleotide or population of polynucleotides according to any one of claims 45 to 48, which is RNA or DNA, or a mixture thereof.

50. The single strand or double strand polynucleotide or population of polynucleotides according to claim 49, wherein DNA is cDNA.

51. The single strand or double strand polynucleotide or population of polynucleotides according to claim 50, which is a long strand, full-coding and/or

IDT Exhibit 1020.6851

full-length cDNA.

52. A vector comprising the linker-polynucleotide according to any one of claims 42-51.

53. The vector according to claim 52, which is a phage, plasmid vector or vector for sequencing.

54. A method for preparing the linker or population of linkers according to any one of claims 1 to 41, which comprises the steps of:

(a) synthesizing a first oligonucleotide single strand comprising an oligonucleotide single strand fixed portion and an oligonucleotide single strand variable portion,

(b) synthesizing a second oligonucleotide single strand comprising an oligonucleotide single strand fixed portion complementary to the first oligonucleotide single strand fixed portion (a), and

(c) annealing the first oligonucleotide strand (a) to the second oligonucleotide strand (b), so that the variable portion protrudes outside the double strand fixed portion of the linker.

55. A method of binding a target single strand polynucleotide to a linker comprising the steps of:

i) preparing the linker according to any one of claims 1 to 41; and

ii) annealing the variable portion of said linker to the target single strand polynucleotide.

56. A method of binding a target single strand polynucleotide to a linker comprising the steps of:

IDT Exhibit 1020.6852

i) preparing the linker according to any one of claims 1 to 41; and

ii) annealing the variable portion of the first strand of the linker to the target single strand polynucleotide and ligating the fixed portion of the second strand of the linker to the target single strand polynucleotide.

57. A method of binding a target single strand polynucleotide or a population of the polynucleotides to a population of linkers comprising the steps of:

i) preparing the population of linkers according to any one of claims 1 to 41; and

ii) annealing the variable portion of the population of linkers to the target single strand polynucleotide or to the population of the polynucleotides.

58. A method of binding a target single strand polynucleotide or a population of the polynucleotides to a population of linkers comprising the steps of:

i) preparing the population of linkers according to any one of claims 1 to 41; and

ii) annealing the variable portion of the first strand of said population of linkers to the target single strand polynucleotide or the population of the polynucleotides and ligating the fixed portion of the second strand of the population of linkers to the target single strand polynucleotide or the population of the polynucleotides.

59. A method of preparing a linker-polynucleotide comprising a linker and a double strand polynucleotide, comprising the steps of:

i) annealing the variable portion of the linker according to any one of claims 1 to 41 to the target first strand polynucleotide, and

ii) synthesizing the second strand polynucleotide complementary to the target single strand polynucleotide.

60. A method of preparing a linker-polynucleotide comprising a linker and a double

IDT Exhibit 1020.6853

strand polynucleotide, comprising the steps of:

i) annealing the variable portion of the first strand of the linker according to any one of claims 1 to 41 to a target single strand polynucleotide and ligating the target single strand polynucleotide to the fixed portion of the second strand of the linker, and

ii) synthesizing the second single strand polynucleotide complementary to said target single strand polynucleotide.

61. A method of preparing a linker-polynucleotide comprising a linker and a double strand polynucleotide, comprising the steps of:

i) annealing the variable portion of the linker or a population of the linkers according to any one of claims 1 to 41 to a target single strand polynucleotide or a population of the target single strand polynucleotides, and

ii) synthesizing the second strand polynucleotide complementary to said target single strand polynucleotide or a population thereof.

62. A method of preparing a linker-polynucleotide comprising a linker and a double strand polynucleotide, comprising the steps of:

i) annealing the variable portion of the first strand of the linker or a population of the linkers according to any one of claims 1 to 41 to a target single strand polynucleotide or a population of the target single strand polynucleotides,

ii) ligating the target single strand polynucleotide or the population of the target single strand polynucleotides to the fixed portion of the second strand of the linker or the population of the linkers, and

iii) synthesizing the second single strand polynucleotide(s) complementary to said target single strand polynucleotide(s).

IDT Exhibit 1020.6854

63. The method according to any one of claims 54-62, wherein one end of the target first strand polynucleotide is annealed to the variable portion of the linker, and ligated to the fixed portion of the second strand of said linker.

64. The method according to any one of claims 54-63, wherein the variable portion is synthesized at random.

65. The method according to any one of claims 55-64, wherein the population of linkers comprises at least one linker in which the variable oligonucleotide portion is a specific oligonucleotide portion, this specific oligonucleotide portion being complementary to the end of at least one specific polynucleotide target to be selected from the group of target polynucleotides.

66. The method according to any one of claims 55-65, wherein ligation is performed by ligase.

67. The method according to claim 66, wherein the ligase is a DNA ligase or RNA ligase.

68. The method according to claim 67, wherein the DNA ligase is T4 DNA ligase or *E. coli* DNA ligase.

69. The method according to any one of claims 66-68, wherein the ligation is performed in the presence of a ligase-stimulating agent.

70. The method according to claim 69, wherein the ligase-stimulating agent is PEG (polyethylene glycol).

71. The method according to any one of claims 55-70, wherein the linker and the target first strand polynucleotide and/or the second strand polynucleotide complementary to the target is DNA.

IDT Exhibit 1020.6855

72. The method according to claim 71, wherein the target first strand polynucleotide or the double strand polynucleotide is a long strand, full-coding and/or full-length cDNA.

73. The method according to claim 72, wherein the first strand cDNA is obtained from the Cap trapping at the 5' end of mRNA.

74. The method according to claim 73, wherein the Cap-trapped cDNA is further normalized or subtracted before or after the ligation to linker.

75. The method according to any one of claims 55-74, which comprises the step of increasing temperature before annealing the linker to the target first strand polynucleotide and/or after synthesizing polynucleotide second strand.

76. The method according to claim 75, wherein the temperature ranges from 25 to 100℃.

77. The method according to claim 76, wherein the temperature is 65℃.

78. The method according to any one of claims 55-77, wherein at least one end of the fixed portion of the linker, opposite to the variable portion, is tagged with a protective group.

79. The method of claim 78, wherein the protective group is NH_2.

80. The method of any one of claims 55-79, wherein the ends of the two strands of the linker fixed portion are bound together by making a loop.

81. The method according to any one of claims 55-80, which further comprises the step in which said linker-polynucleotide is cleaved at both ends in restriction enzyme sites and inserted into a vector.

82. The method according to any one of claims55-80, which further comprises the step in which said linker-polynucleotide is cleaved remaining blunt ends and inserted into

IDT Exhibit 1020.6856

a vector.

83. A method of marking a polynucleotide library and distinguishing said library, comprising the step of providing a population of linkers, comprising a fixed portion and a variable portion, wherein the fixed portion comprises at least one marker indicating a specific or defined tissue or species, and selecting and separating said library by said specific or defined marker.

84. The method according to claim 83, wherein said linker or population of linkers is the one according to any one of claims 1 to 41.

85. The method according to claim 83 or 84, wherein the fixed portions comprised in all linkers among the population of linkers are oligonucleotide portions having the same sequence.

86. A method of binding a linker or population of linkers to mRNA, which comprises the steps of:

(a) treating mRNA with phosphatase and removing phosphate groups from uncapped mRNA,

(b) treating a product of step (a) with pyrophosphatase, which removes the CAP structure from capped mRNA, and

(c) adding an RNA ligase in the presence of the linker according to any one of claims 1 to 41.

87. A method of preparing a linker-polynucleotide, which comprises the steps of:

(a) treating mRNA with phosphatase and removing phosphate groups from uncapped mRNA,

IDT Exhibit 1020.6857

(b) treating a product of step (a) with pyrophosphatase, which removes the CAP structure from capped mRNA,

(c) adding an RNA ligase in the presence of the linker according to any one of claims 1 to 41, and

(d) adding an oligo dT and synthesizing a polynucleotide complementary to the complete sequence of said mRNA.

88. The method according to claim 87, wherein the polynucleotide complementary to mRNA is cDNA.

89. The method according to any one of claims 86-88, wherein the phosphatase is bacterial alkaline phosphatase (BAP).

90. The method according to any one of claims 86-88, wherein the phosphatase is tobacco acid pyrophosphatase (TAP).

91. The method according to any one of claims 86-90, wherein the linker of step (c) is DNA and further comprising the steps of

(e) adding RNase H,

(f) adding DNA polymerase I, and

(g) synthesizing a cDNA strand.

92. A method of binding a linker or a population of linkers according to any one of claims 1 to 41 to a target single strand polynucleotide or population of polynucleotides comprising the step of adding RNA ligase.

93. The method according to claim 92, which further comprises the addition of an oligonucleotide primer or population of primers complementary to the variable portion of the linker or population of linkers.

IDT Exhibit 1020.6858

94. The method according to claim 93, wherein the primer comprised in the population of primers is randomly synthesized.

95. A method of preparing a second strand polynucleotide, which comprises the step according to any one of claims 92-94 and further synthesizing the second strand polynucleotide complementary to a target polynucleotide.

96. A method of preparing a DNA/RNA hybrid, which comprises the steps of:

i) providing a full-length/coding or long poly-A mRNAs,

ii) ligating and annealing said mRNA to the linker according to any one of claims 1 to 41, the linker comprising a first restriction enzyme site,

iii) annealing an oligo dT-primer comprising a second restriction enzyme site to the mRNA,

iv) synthesizing a cDNA strand,

v) isolating the hybrid by using restriction enzymes which recognize the two specific restriction enzyme sites introduced, and

vi) cloning.

97. A method of preparing a linker-polynucleotide product comprising a linker and a single strand polynucleotide, comprising annealing the variable portion of the linker according to the present invention to the target first strand polynucleotide.

IDT Exhibit 1020.6859

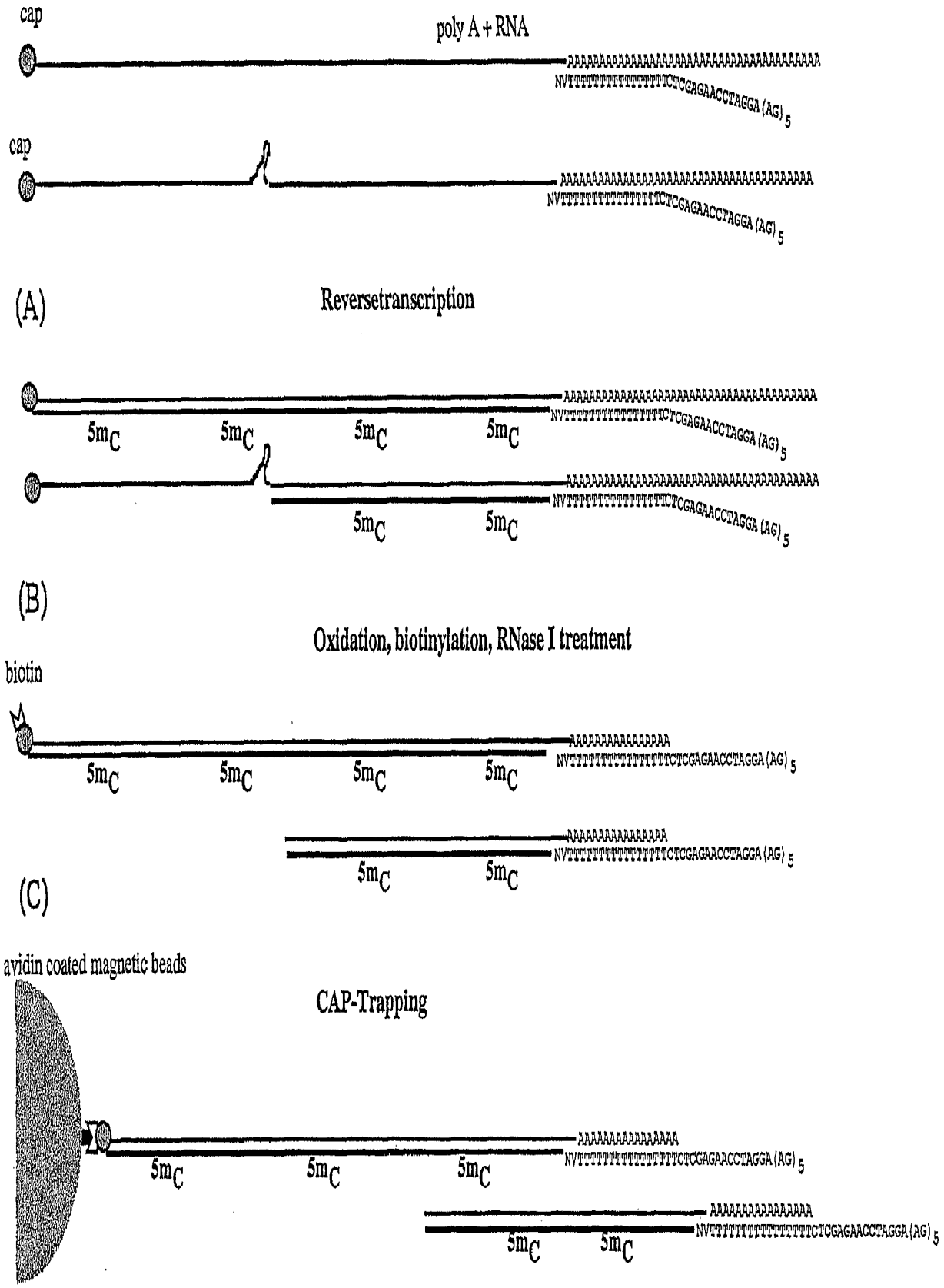
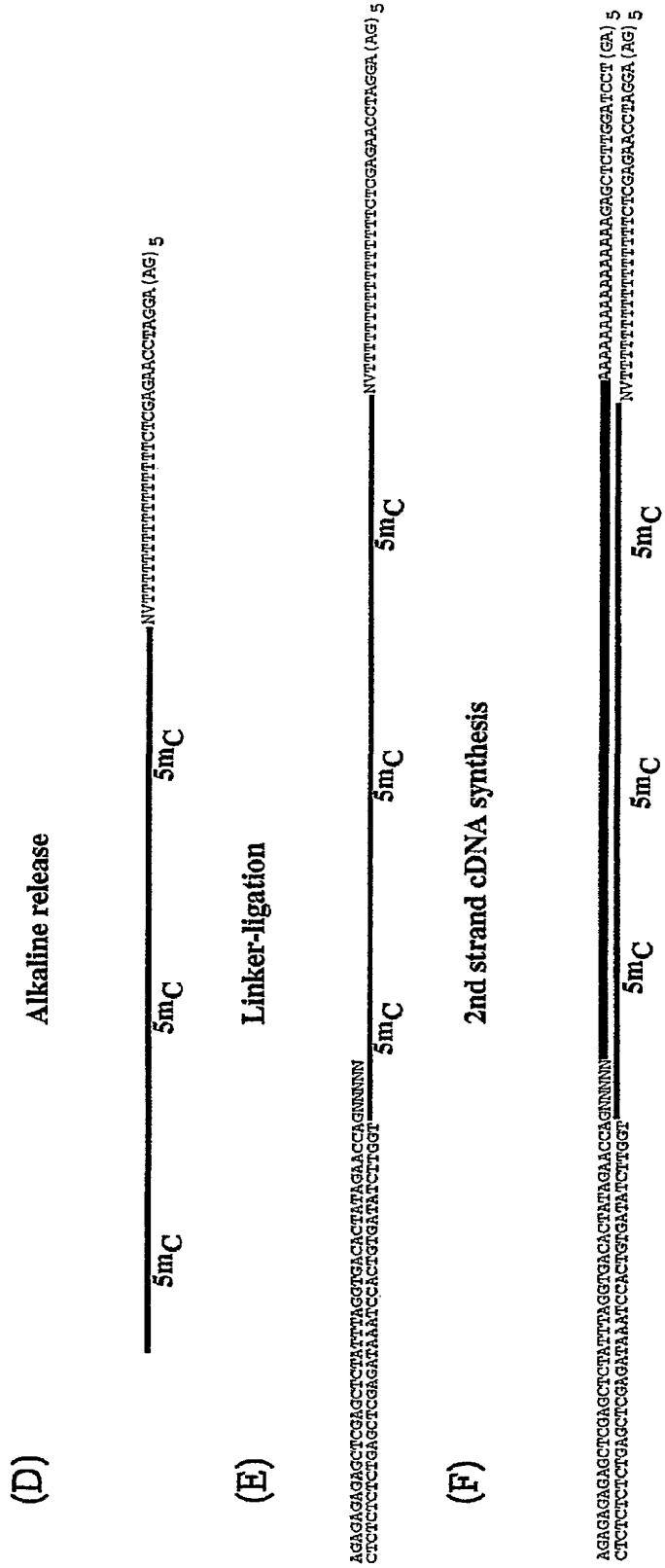


Fig.1

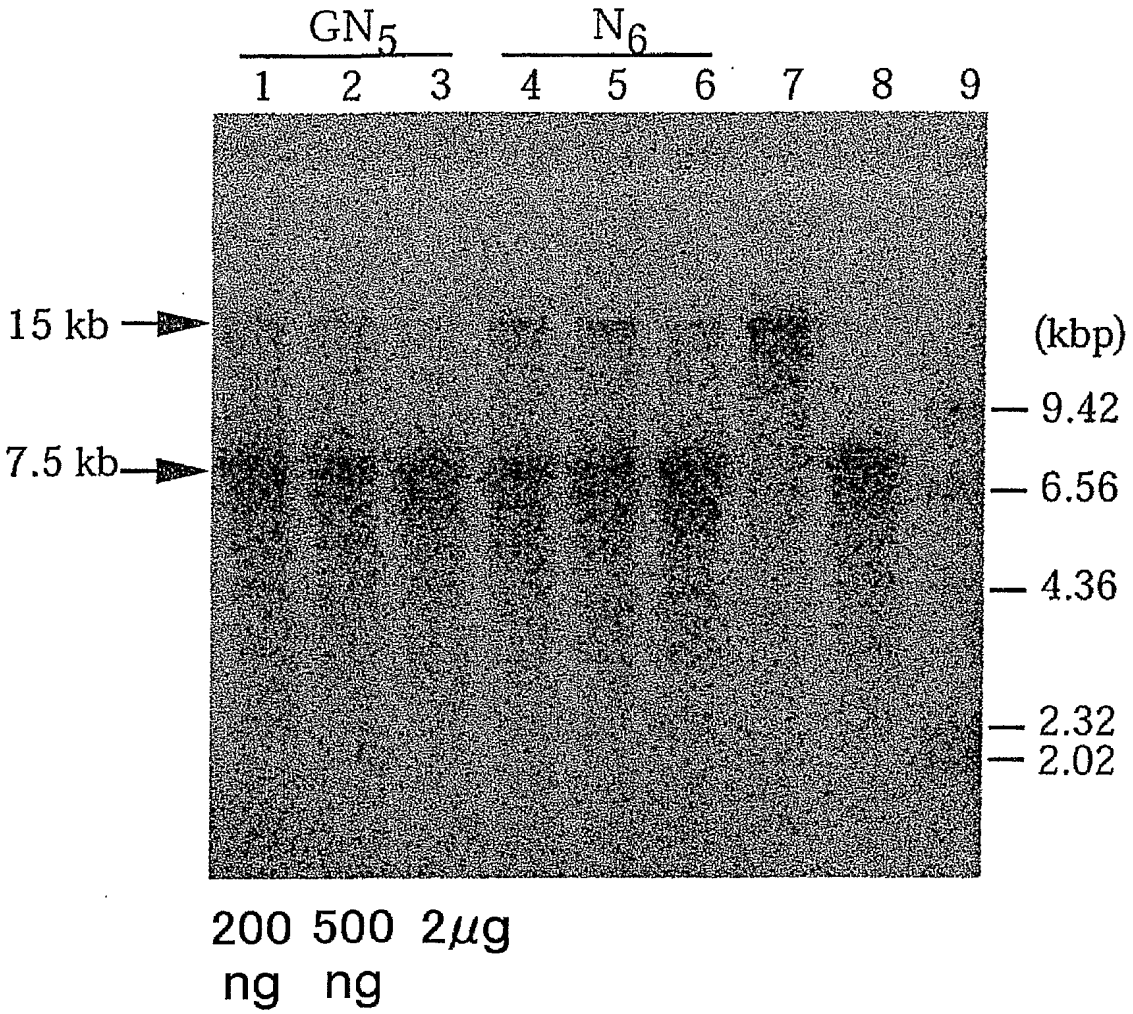
IDT Exhibit 1020.6860

Fig.2



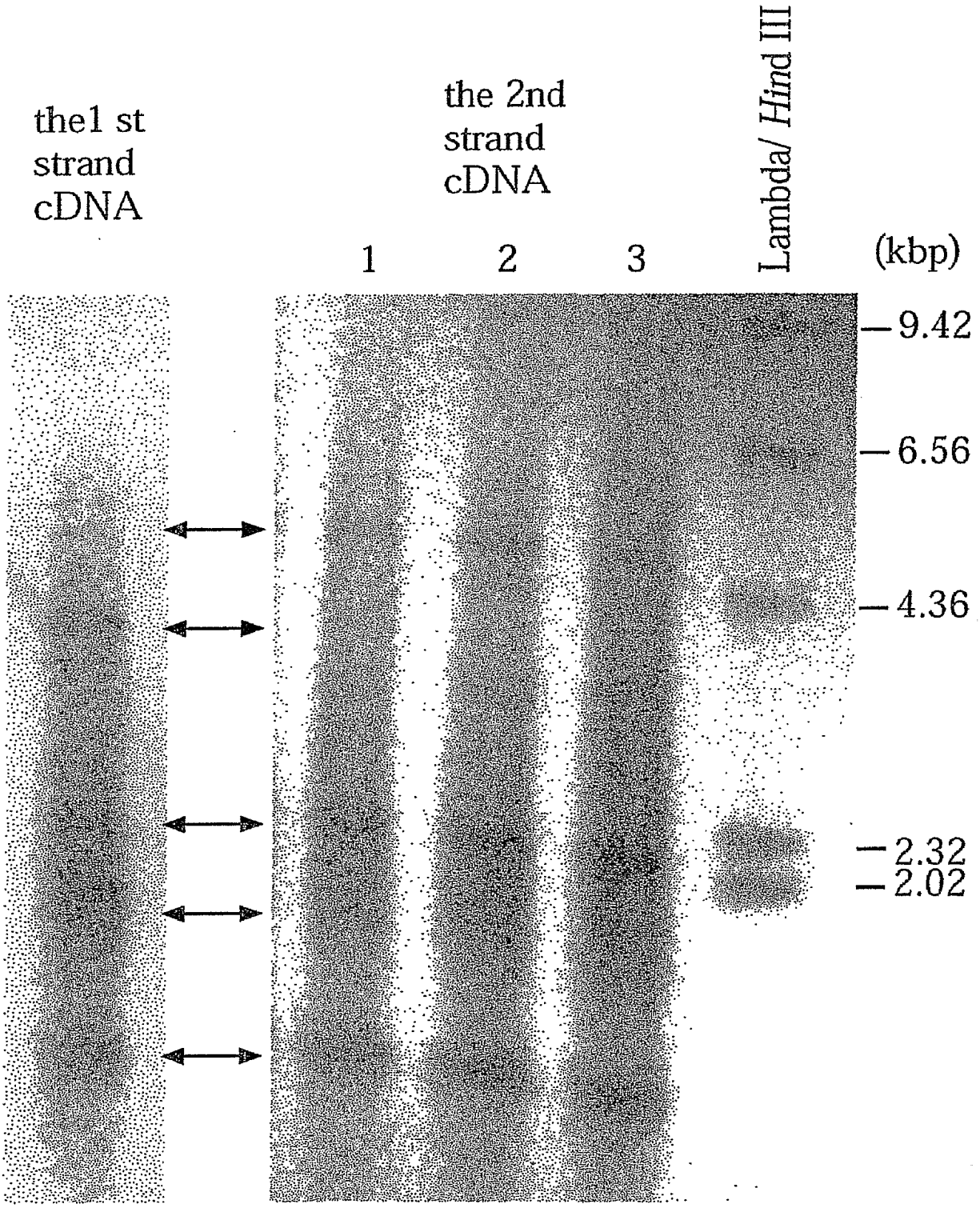
IDT Exhibit 1020.6861

Fig.3



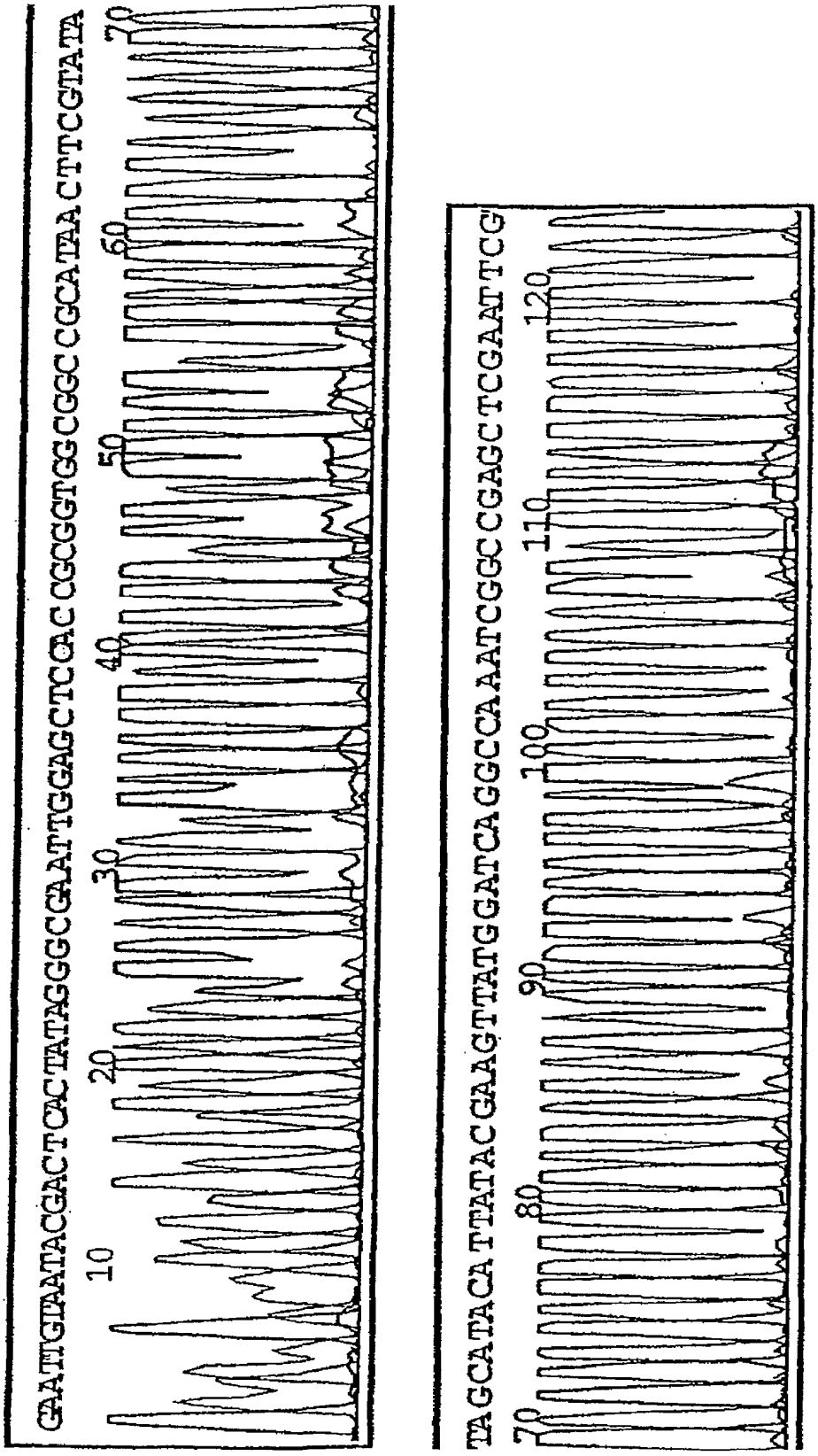
IDT Exhibit 1020.6862

Fig.4



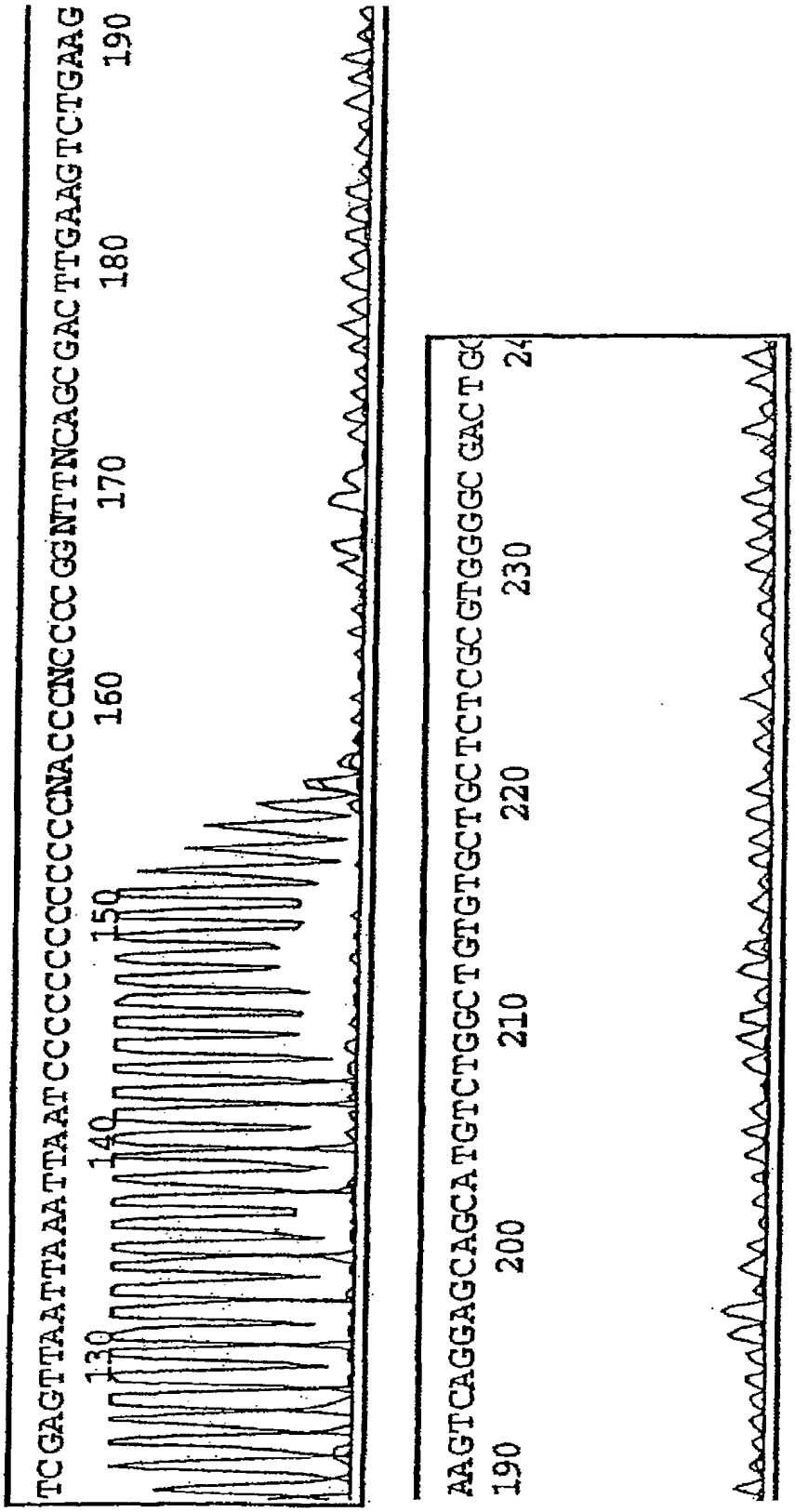
IDT Exhibit 1020.6863

Fig.5(a)



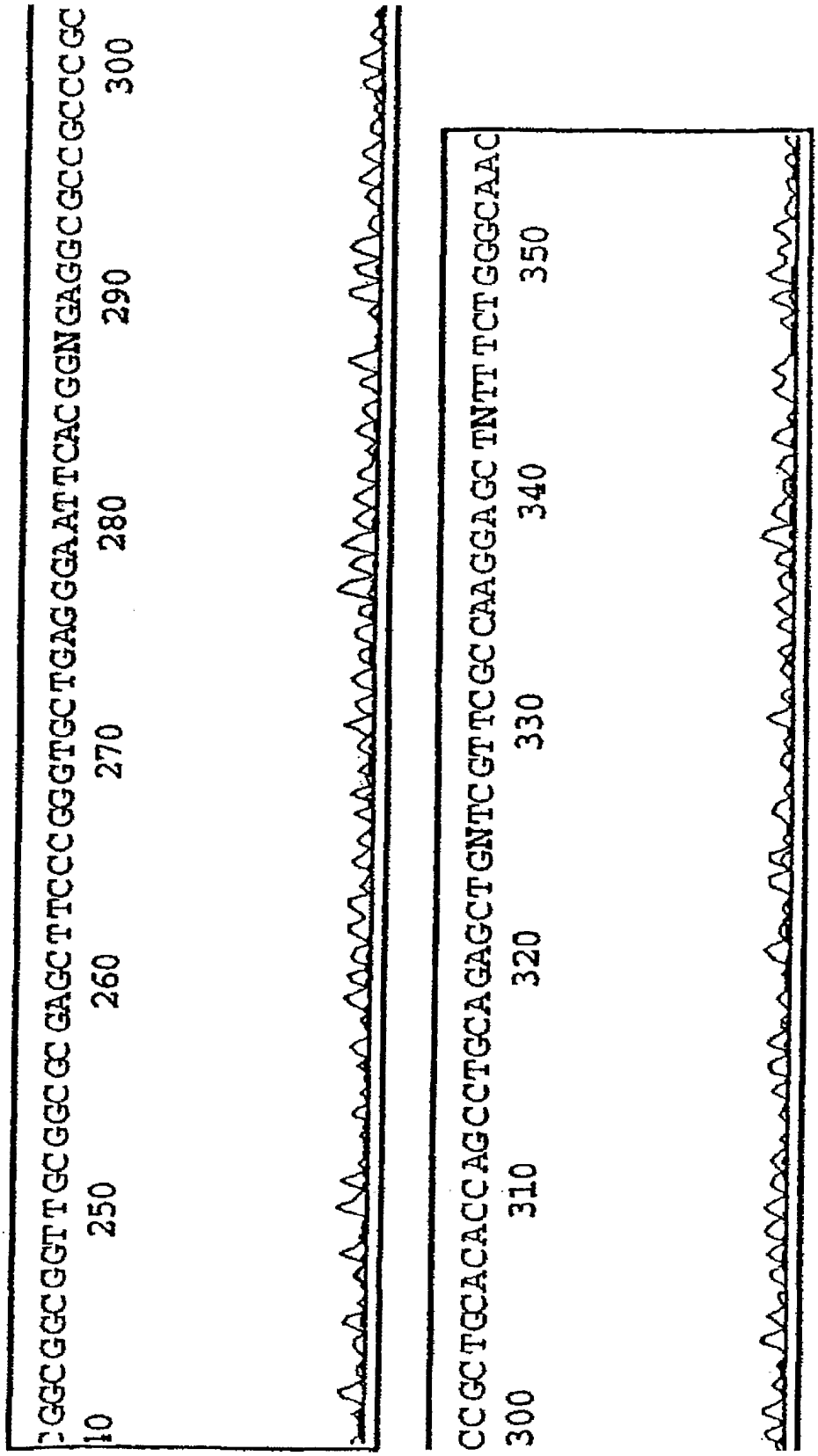
IDT Exhibit 1020.6864

Fig.5(b)



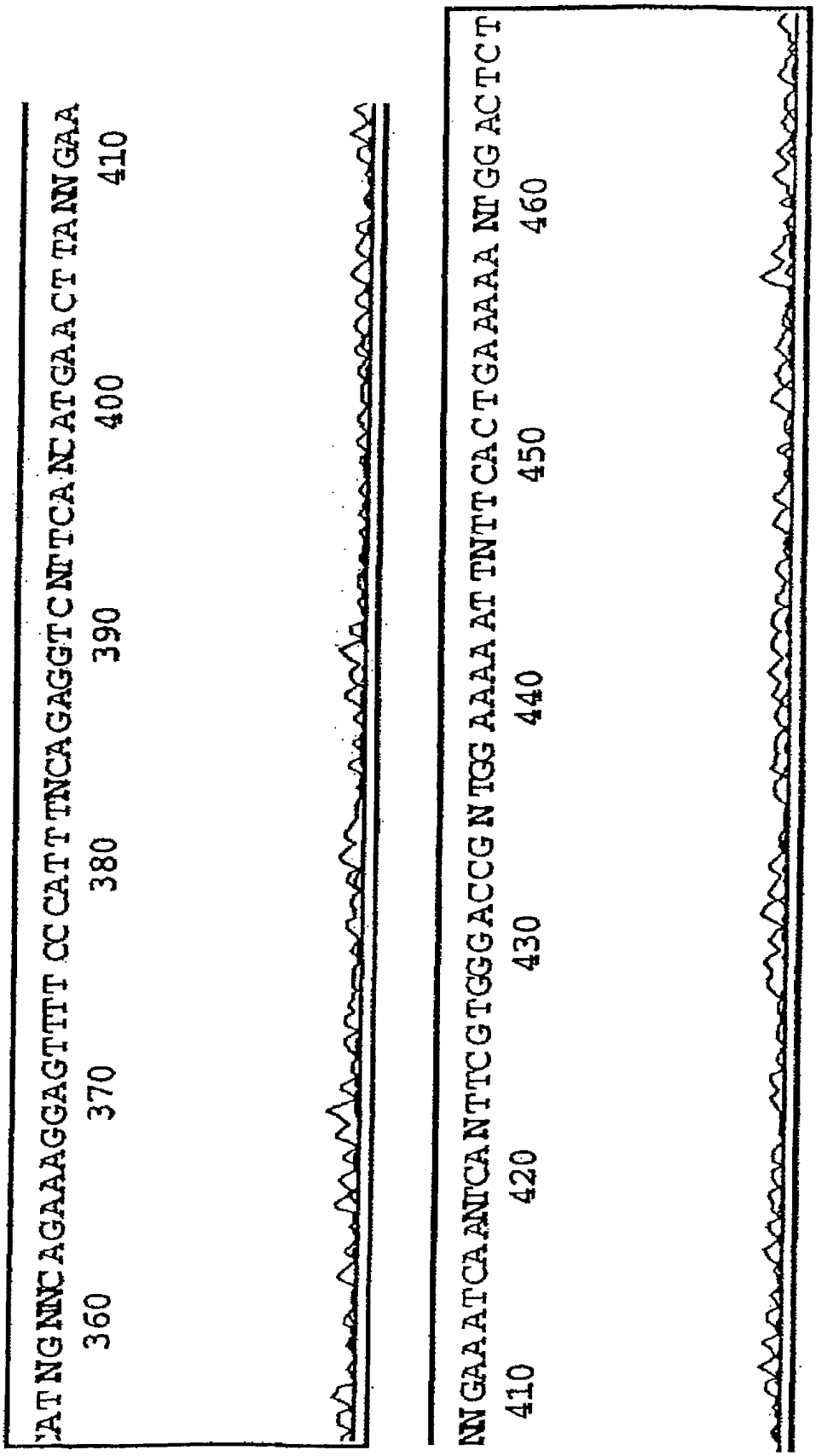
IDT Exhibit 1020.6865

Fig.5(c)



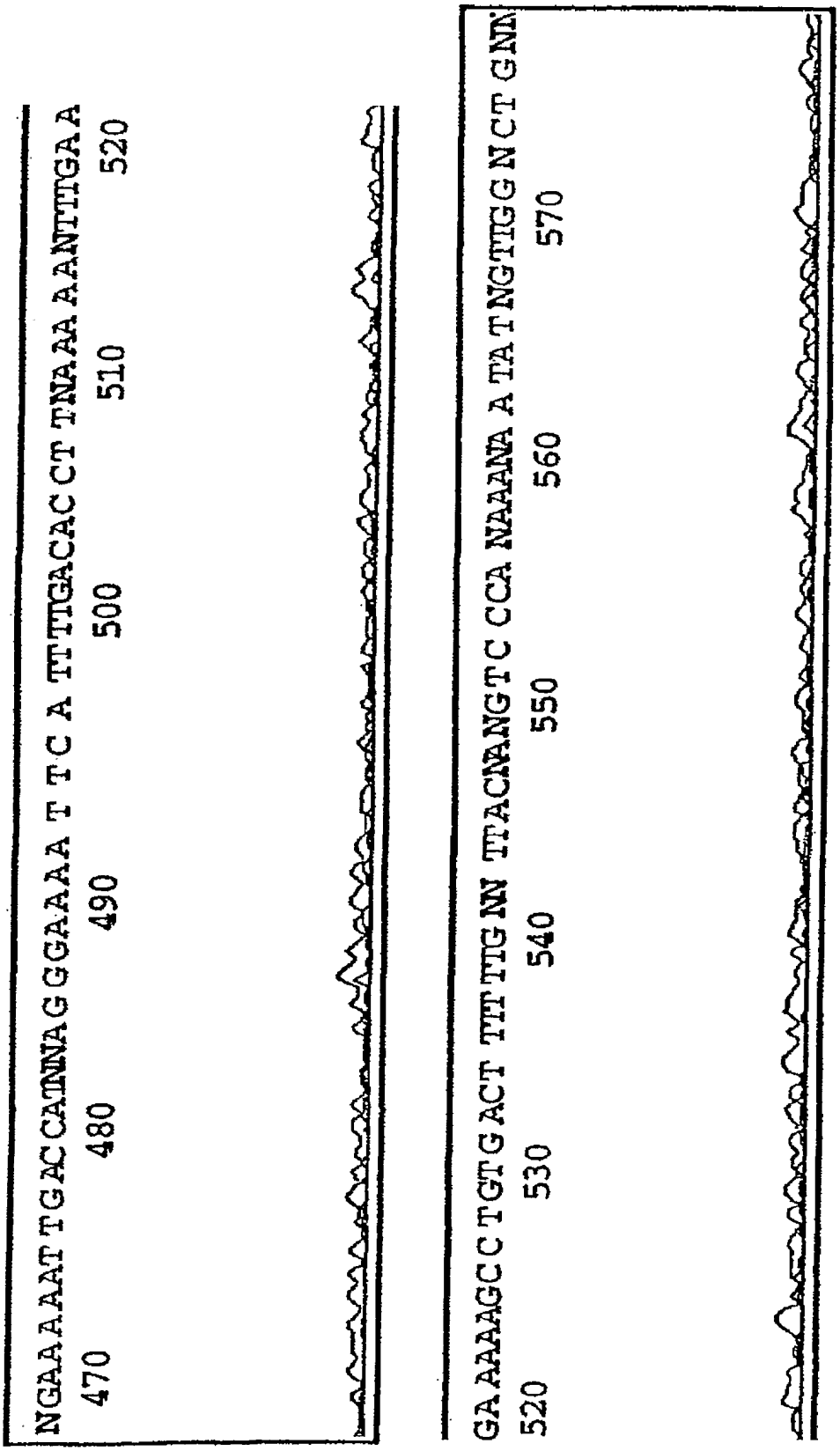
IDT Exhibit 1020.6866

Fig.5(d)



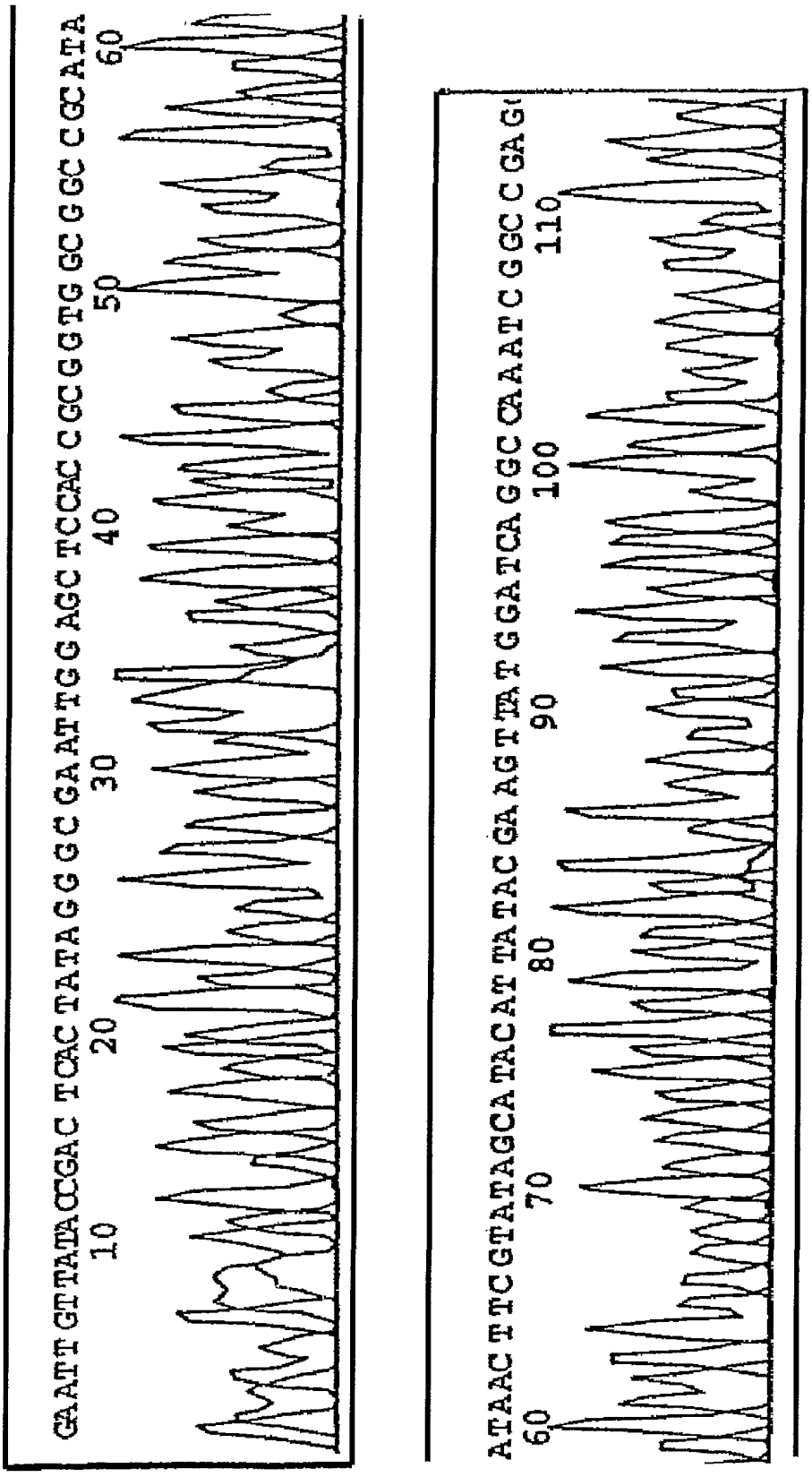
IDT Exhibit 1020.6867

Fig.5(e)



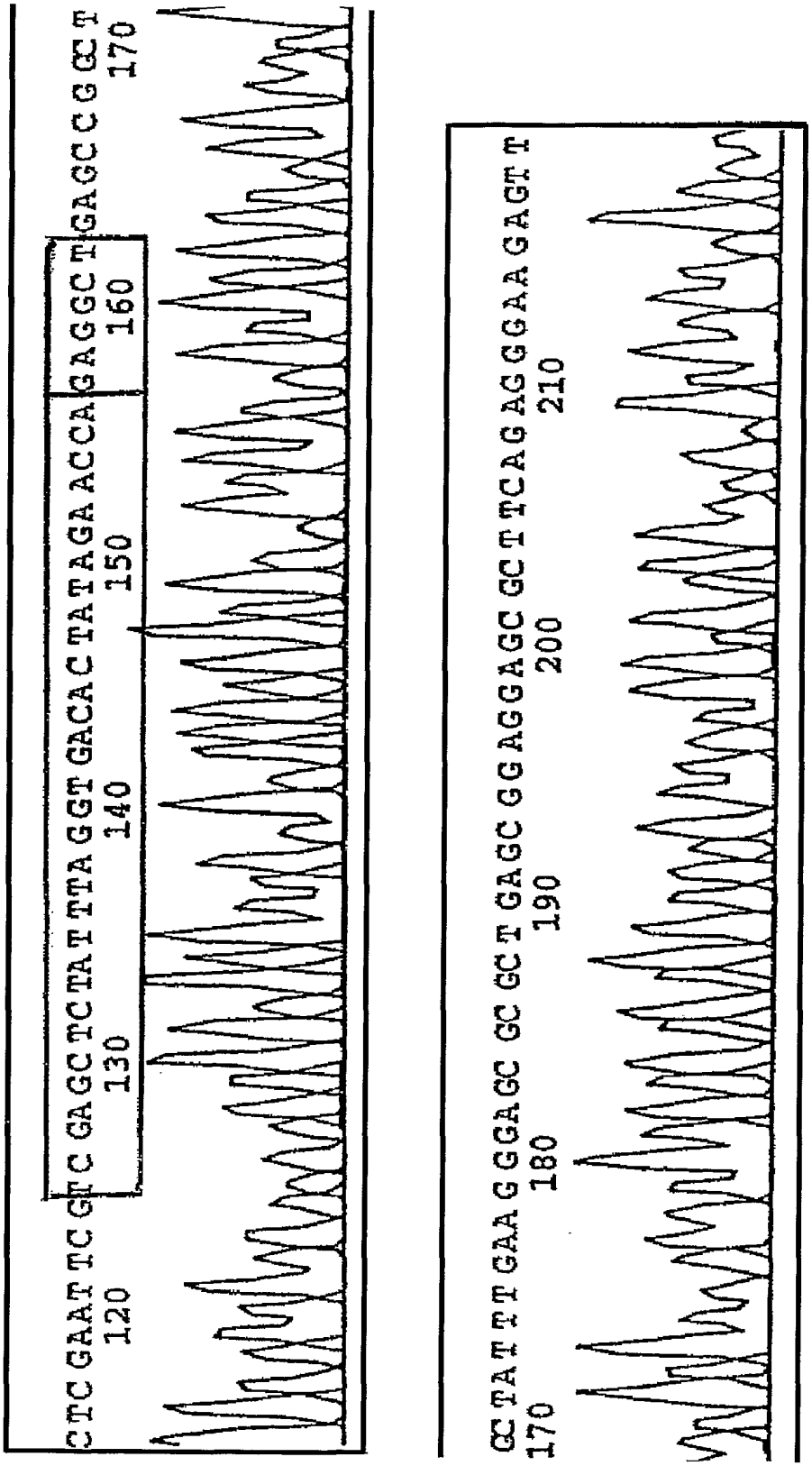
IDT Exhibit 1020.6868

Fig.6(a)



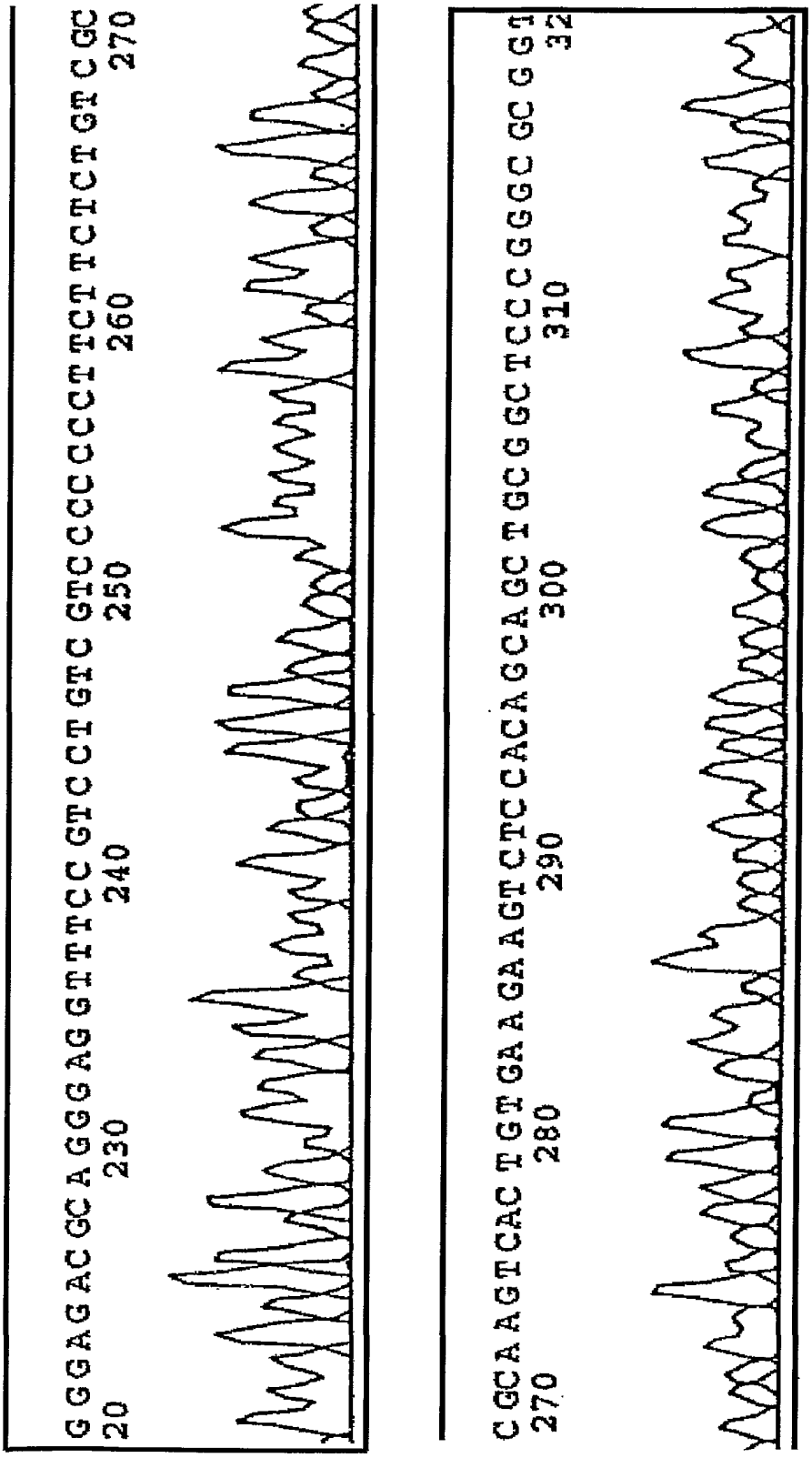
IDT Exhibit 1020.6869

Fig.6(b)



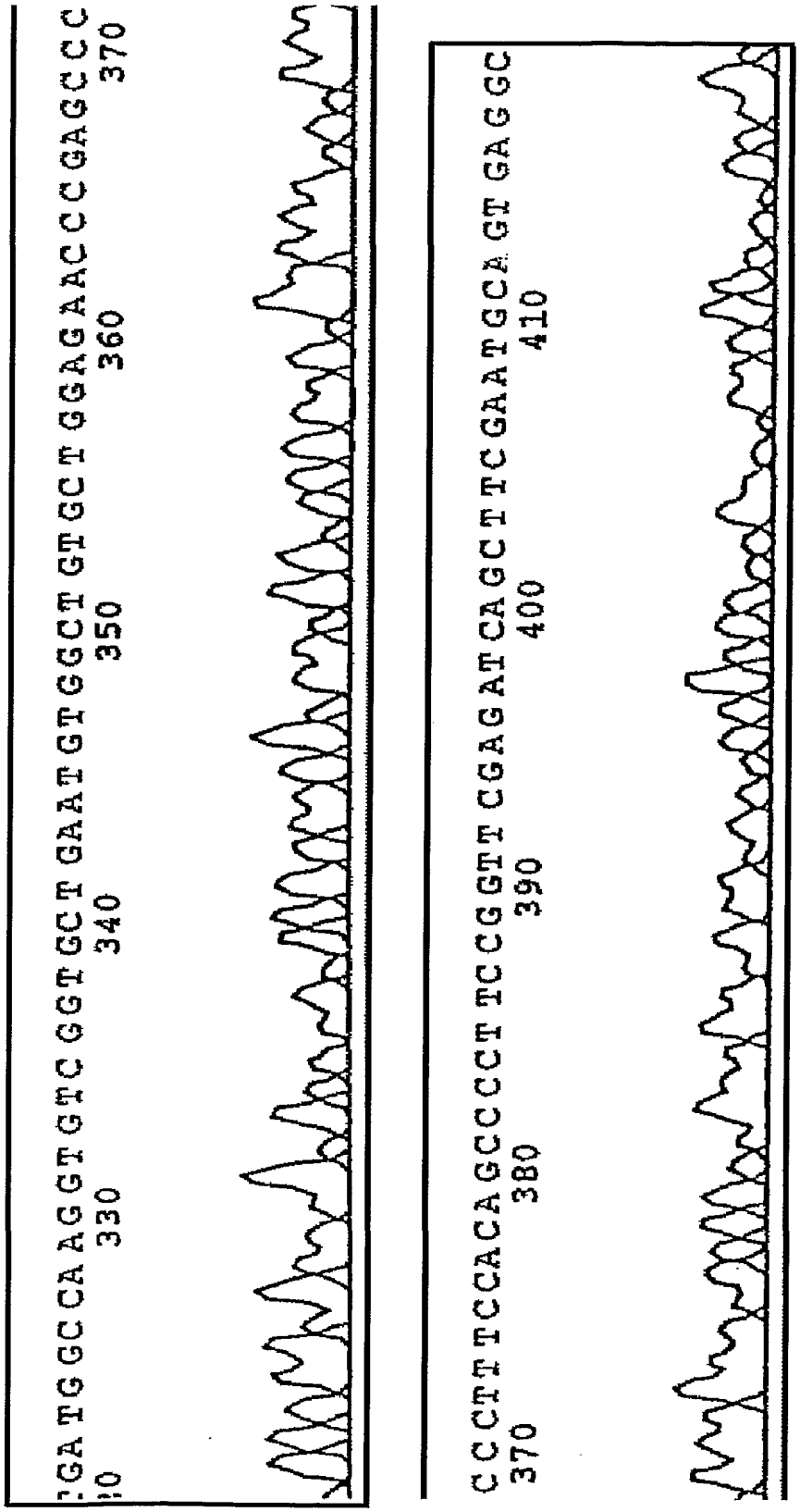
IDT Exhibit 1020.6870

Fig.6(c)



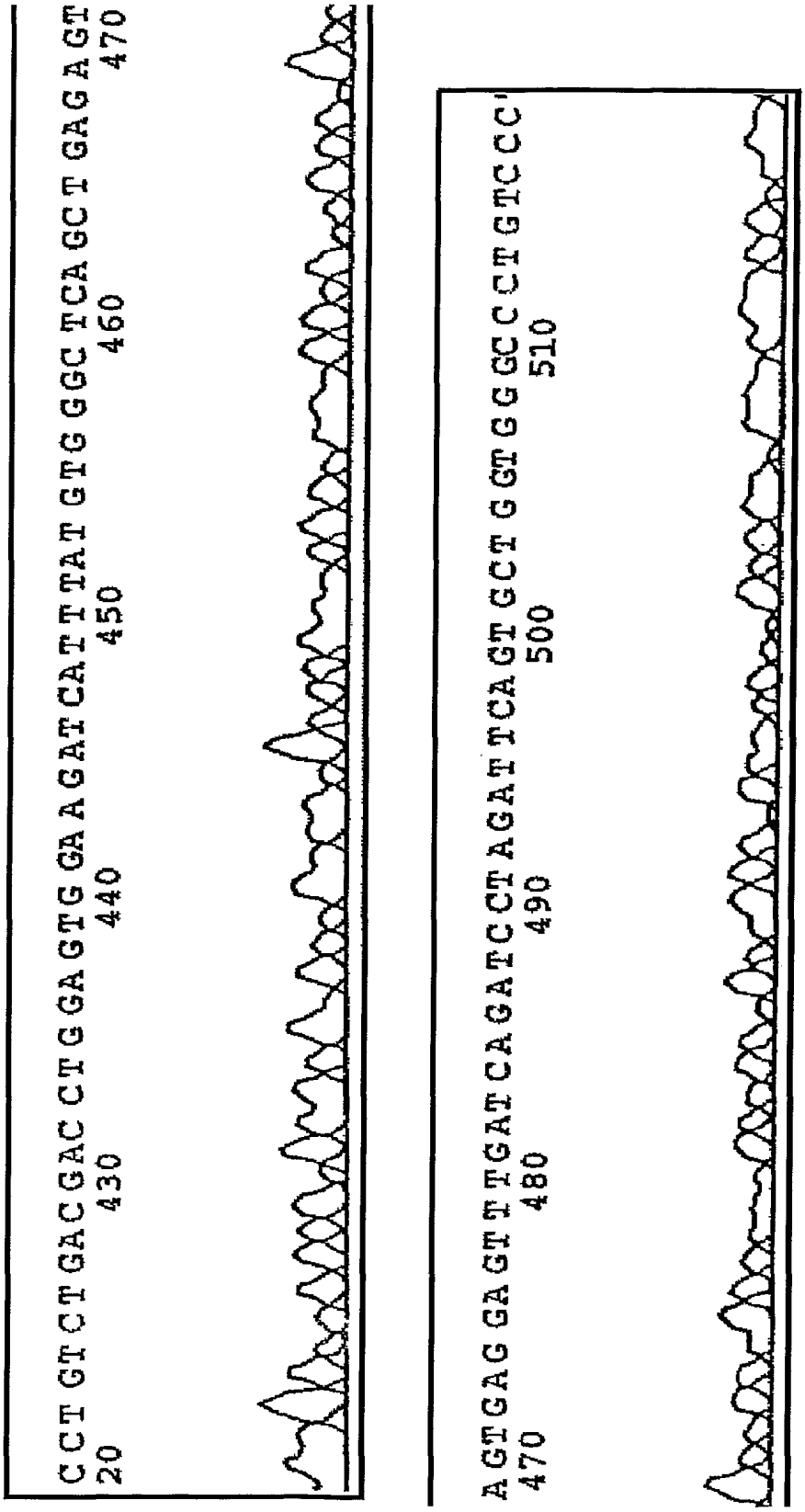
IDT Exhibit 1020.6871

Fig.6(d)



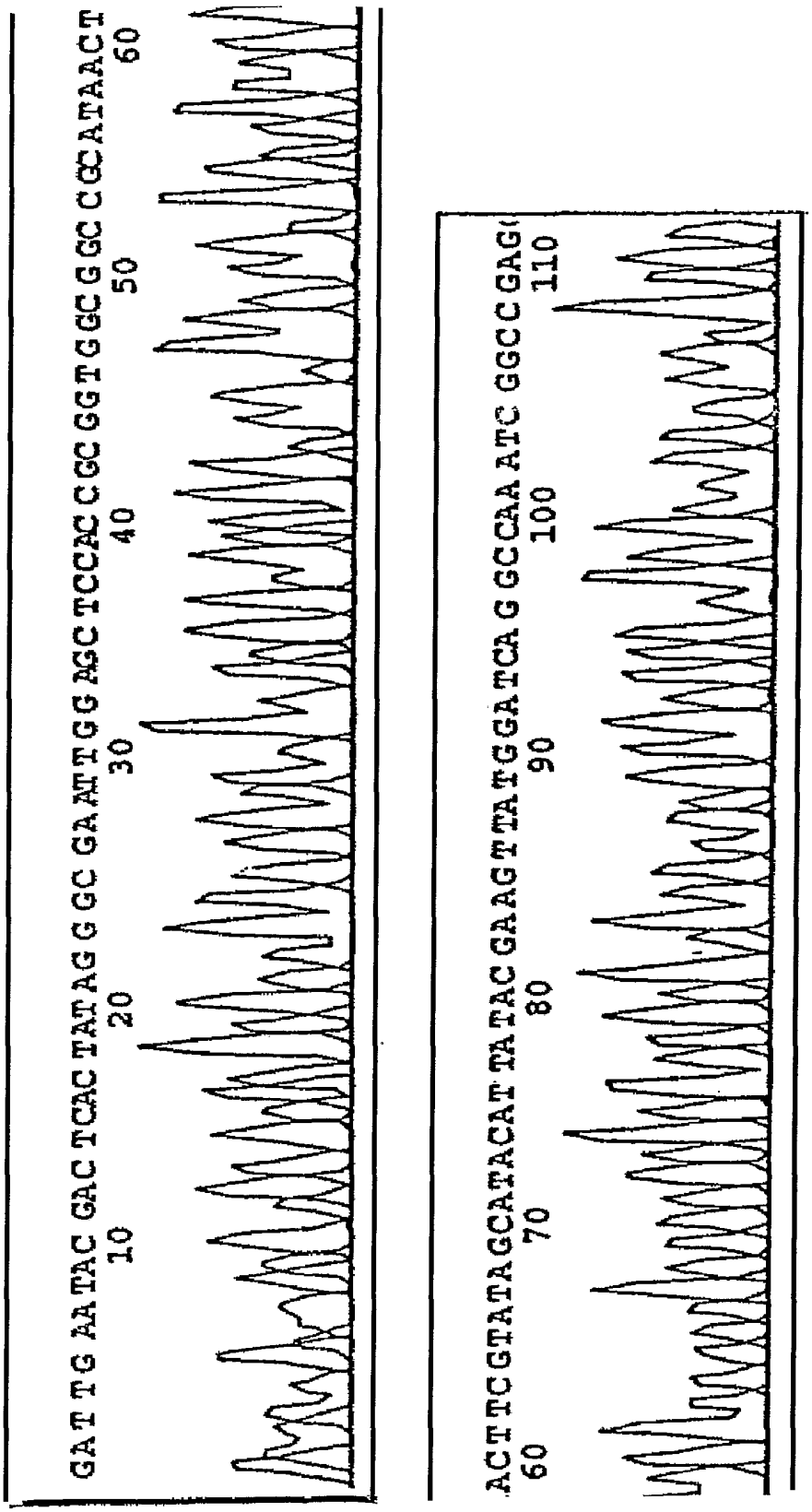
IDT Exhibit 1020.6872

Fig.6(e)



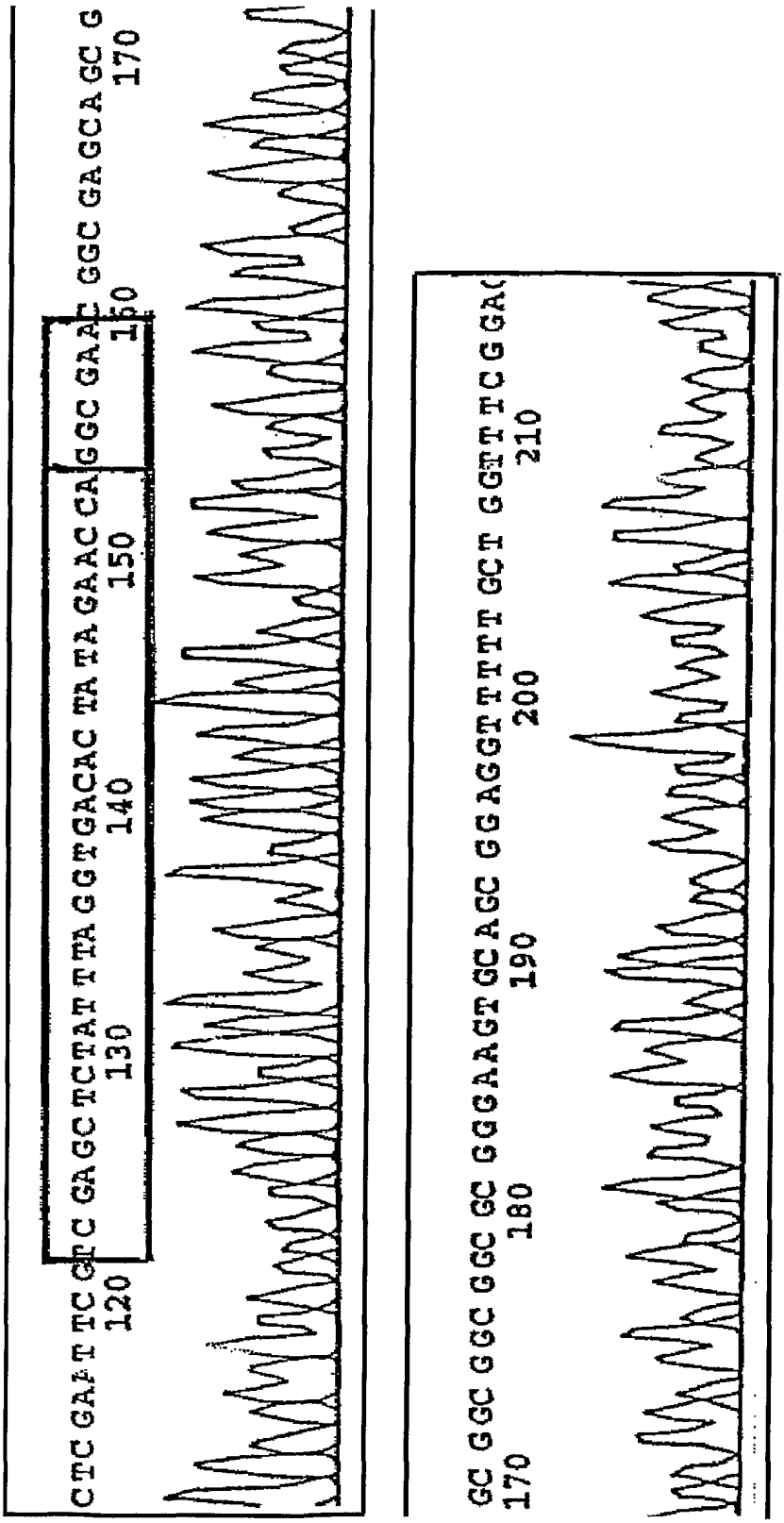
IDT Exhibit 1020.6873

Fig.7(a)



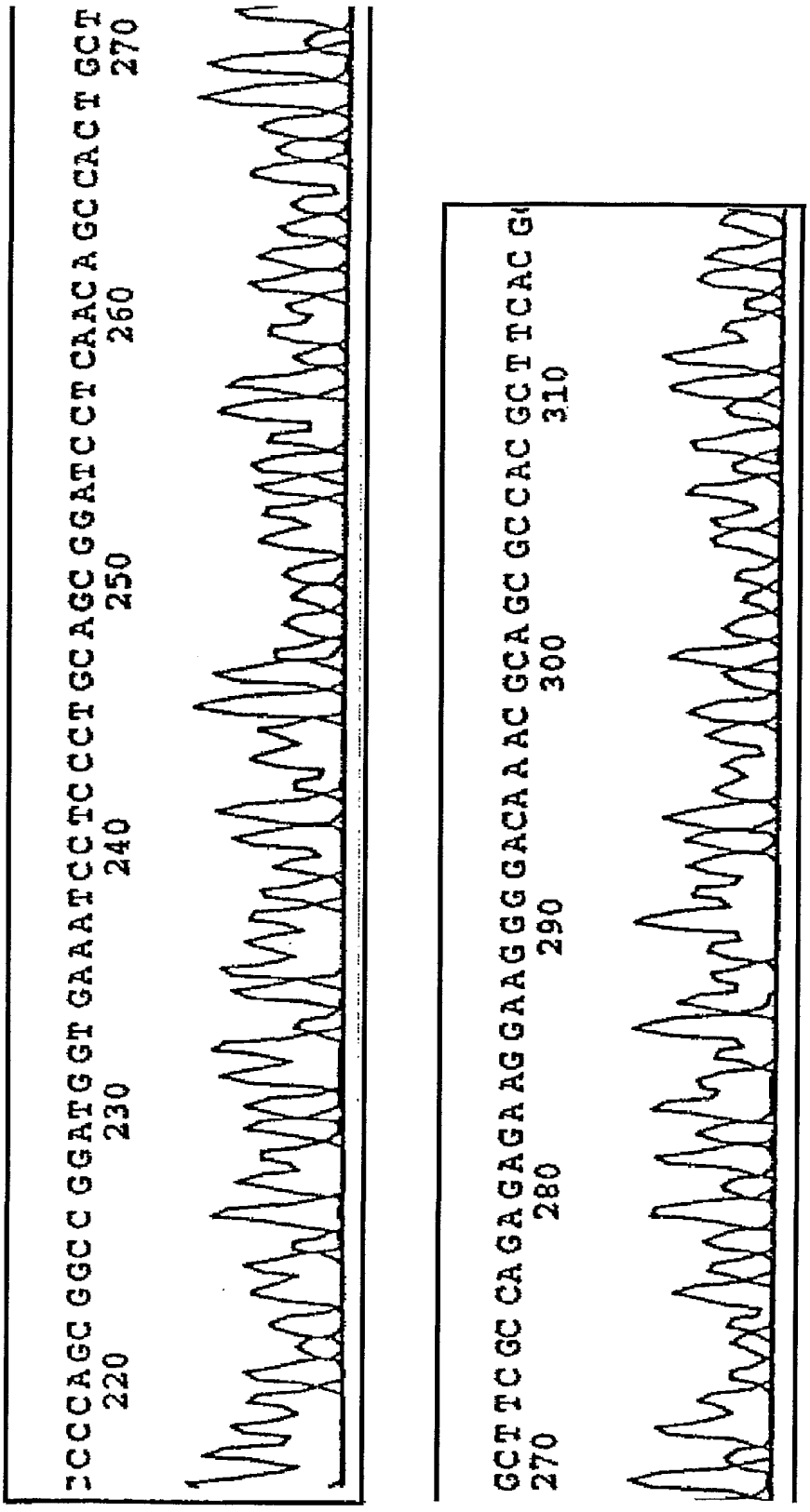
IDT Exhibit 1020.6874

Fig.7(b)



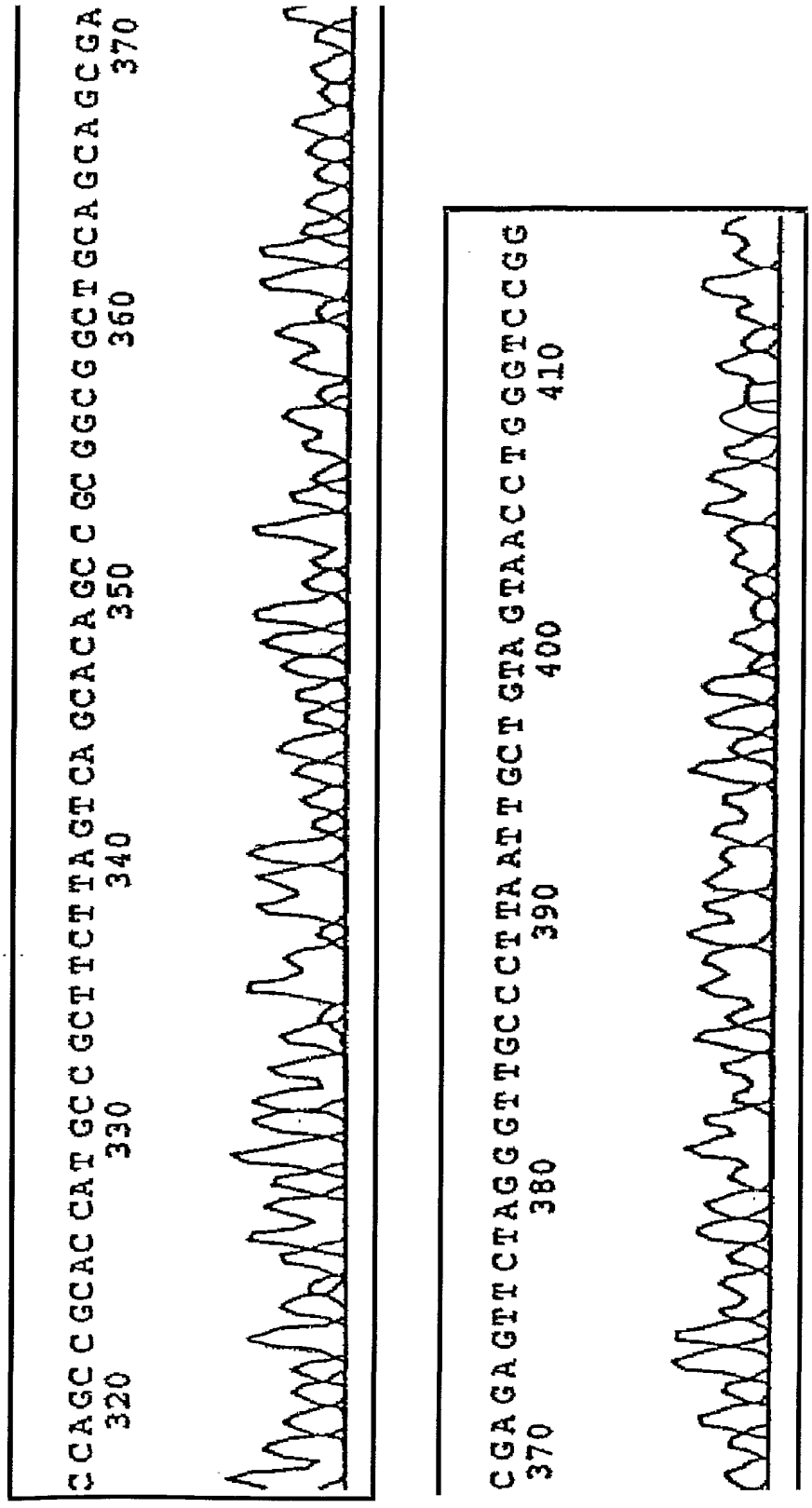
IDT Exhibit 1020.6875

Fig.7(c)



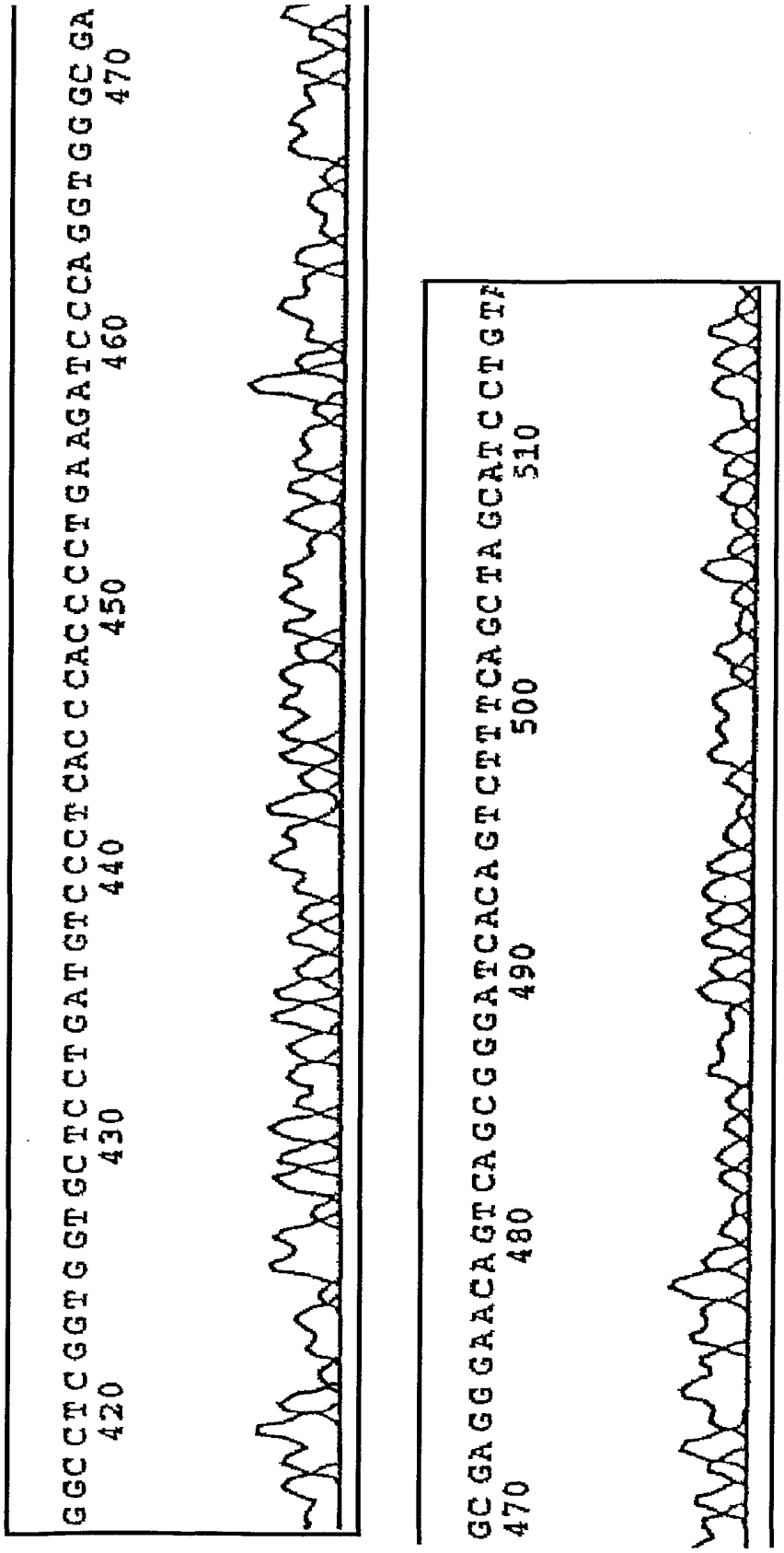
IDT Exhibit 1020.6876

Fig.7(d)



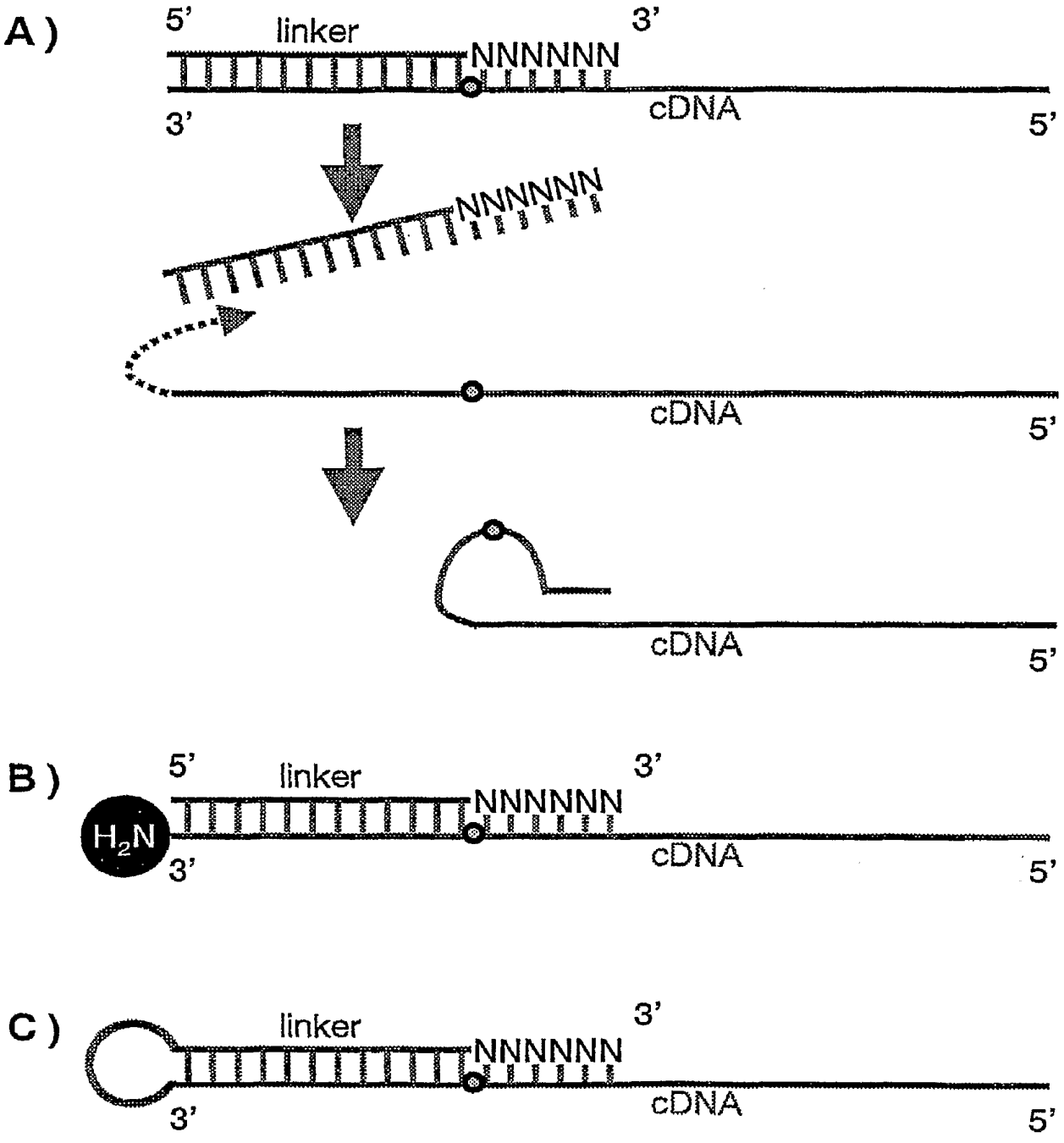
IDT Exhibit 1020.6877

Fig.7(e)



IDT Exhibit 1020.6878

Fig.8



IDT Exhibit 1020.6879

SEQUENCE LISTING

<110> The Institute of Physical and Chemical Research

Hayashizaki, Yoshihide

<120> Oligonucleotide linkers comprising a variable cohesive portion and method for the preparation of polynucleotide libraries by using said linkers

<130> Linker Ref:1-13

<140>

<141>

<160> 7

<170> PatentIn Ver. 2.1

<210> 1

<211> 49

<212> DNA

<213> Artificial Sequence

IDT Exhibit 1020.6880

<220>

<223> Description of Artificial Sequence: GN5 linker upper strand

<220>

<221> misc_feature

<222> (45-49)

<223> Nucleotides 45-49 are n wherein n = any nucleotide

<400> 1

agagagagag ctcgagctct atttaggtga cactatagaa ccagnnnnn 49

<210> 2

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: linker lower strand

<400> 2

IDT Exhibit 1020.6881

tggttctata gtgtcaccta aatagagctc gagctctctc tct 43

<210> 3

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: N6 linker

upper strand

<220>

<221> misc_feature

<222> (44-49)

<223> Nucleotides 44-49 are n wherein n = any nucleotide

<400> 3

agagagagag ctcgagctct atttaggtga cactatagaa ccannnnnn 49

<210> 4

<211> 43

IDT Exhibit 1020.6882

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer
comprising BamHI site

<220>

<221> misc_feature

<222> (42)

<223> Nucleotide 42 is v wherein v = g or c or a

<220>

<221> misc_feature

<222> (43)

<223> Nucleotide 43 is n wherein n = any nucleotide

<400> 4

gagagagaga aggatccaag agctcttttt ttttttttt tvn 43

<210> 5

<211> 18

IDT Exhibit 1020.6883

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: M13 forward primer

<400> 5

tgtaaaacga cggccagt 18

<210> 6

<211> 106

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligonucleotide

<400> 6

ggccgcataa cttcgtatag catacattat acgaagttat ggatcaggcc aaatcggccg 60

agctcgaatt cgtcgacgag agactgcagg agagaggatc cggtac 106

IDT Exhibit 1020.6884

<210> 7

<211> 98

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

Oligonucleotide

<400> 7

cggatcctct ctcctgcagt ctctcgtcga cgaattcgag ctcggccgat ttggcctgat 60

ccataacttc gtataatgta tgctatacga agttatgc 98

IDT Exhibit 1020.6885

Espacenet

Bibliographic data: WO0248402 (A2) — 2002-06-20

METHODS AND COMPOSITIONS FOR GENERATION OF MULTIPLE COPIES OF NUCLEIC ACID SEQUENCES AND METHODS OF DETECTION THEREOF

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

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

Classification: - international: ***C12M1/00; C12N15/09; C12Q1/68; G01N33/53; G01N33/566;*** (IPC1-7): C12Q1/68
- cooperative: **C12Q1/6823**; **C12Q1/6827**; **C12Q1/6853**; **C12Q1/6869** more

Application number: WO2001US47775 20011213 Global Dossier

Priority number (s): US20000255638P 20001213

Also published as: WO0248402 (A3) AT475720 (T) AU2897402 (A) CA2430329 (A1) EP1427847 (A2) more

Abstract of WO0248402 (A2)

The present invention provides novel isothermal methods of generating multiple copies of, detecting and/or quantifying nucleic acid sequences of interest based on limited primer extension or attachment of oligonucleotide pairs using composite RNA/DNA primers. Methods for generating multiple copies of and/ordetecting and/or quantifying nucleic acid sequences, wherein products of primer extension or attachment of oligonucleotide pairs comprising a cleavable portion are generated, and wherein cleavage of the products results in dissociation of cleaved products from target polynucleotides, are provided. The invention further provides compositions, kits and systems for practicing these methods.

IDT Exhibit 1020.6886

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 June 2002 (20.06.2002)

PCT

(10) International Publication Number
WO 02/48402 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US01/47775

(22) International Filing Date:
13 December 2001 (13.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/255,638 13 December 2000 (13.12.2000) US

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

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **KURN, Nurith** [IL/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, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, 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, 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 02/48402 A2

(54) Title: METHODS AND COMPOSITIONS FOR GENERATION OF MULTIPLE COPIES OF NUCLEIC ACID SEQUENCES AND METHODS OF DETECTION THEREOF

(57) Abstract: The present invention provides novel isothermal methods of generating multiple copies of, detecting and/or quantifying nucleic acid sequences of interest based on limited primer extension or attachment of oligonucleotide pairs using composite RNA/DNA primers. Methods for generating multiple copies of and/ordetecting and/or quantifying nucleic acid sequences, wherein products of primer extension or attachment of oligonucleotide pairs comprising a cleavable portion are generated, and wherein cleavage of the products results in dissociation of cleaved products from target polynucleotides, are provided. The invention further provides compositions, kits and systems for practicing these methods.

IDT Exhibit 1020.6887

METHODS AND COMPOSITIONS FOR GENERATION OF MULTIPLE COPIES OF NUCLEIC ACID SEQUENCES AND METHODS OF DETECTION THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the priority benefit of the provisional patent application U.S. Serial No. 60/255,638, filed December 13, 2000, which is incorporated by reference in its entirety.

TECHNICAL FIELD

The invention relates to the field of generating multiple copies, quantification and detection of specific nucleic acid sequences. More particularly, the invention provides methods, compositions and kits which employ composite primers for copying, quantifying and detecting nucleic acid sequences of interest.

BACKGROUND ART

The development of methods for nucleic acid amplification and detection of amplification products have advanced the detection, identification, quantification and sequence analyses of nucleic acid sequences in recent years.

Nucleic acid analysis is useful for detection and identification of pathogens, detection of gene alteration leading to defined phenotypes, diagnosis of genetic diseases or the susceptibility to a disease, assessment of gene expression in development, disease and in response to defined stimuli, as well as the various genome projects. Other applications of nucleic acid amplification methods are the detection of rare cells, detection of pathogens, and the detection of altered gene expression in malignancy, and the like. Nucleic acid amplification is potentially useful for both qualitative analysis, such as the detection of the presence of defined nucleic acid sequences, and quantification of defined gene sequences. The latter is useful for assessment of the amount of pathogenic sequences as well as the determination of gene multiplication or deletion, as often found in cell transformation from normal to malignant cell type.

The detection of sequence alterations, such as single nucleotide polymorphisms, in a nucleic acid sequence is important for the detection of mutant genotypes, as relevant for genetic analysis, the detection of mutations leading to drug resistance, pharmacogenomics, etc. Various methods for the detection of specific defined sequences include methods based on allele specific primer extension, allele specific probe ligation, differential probe hybridization, and limited primer extension. See, for example, U.S. Patent Nos. 5,888,819; 6,004,744; 5,882,867; 5, 854, 033; 5,710,028; 6,027,889; 6,004,745; 5,763,178; 5,011,769; 5,185,243; 4,876,187; 5,882,867; 5,731,146; WO US88/02746; WO 99/55912; WO92/15712; WO 00/09745; WO 97/32040; WO 00/56925.

Although detection of the presence of a defined nucleic acid sequence, and its sequence analysis, can be carried out by probe hybridization, the method generally lacks sensitivity when low amounts of the nucleic acid sequence is present in the test sample, such as a few molecules. One solution to this obstacle is the development of methods for generation of multiple copies of the defined nucleic acid sequence, which are suitable for further analysis. See, e.g., WO 01/20035. Other methods for increasing the sensitivity of detection of hybridization analysis are based on the generation of multiple products from the hybridized probe, or probes, for example cleavage of the hybridized probe to form multiple products or the ligation of adjacent probes to form a unique hybridization dependent product. Similarly, increased sensitivity of hybridization reaction was achieved by methods for amplification of signals generated by the hybridization event, such as the method based on hybridization of branched DNA probes.

Recent progress in the elucidation of nucleic acid sequences of various genomes has contributed to the identification of sequence alterations which define mutations leading to altered phenotypes and the uncovering of a vast number of single nucleotide polymorphisms (SNP) that are suspected of underlying important biological and clinical manifestations. Thus, there is a serious need for methods of generating multiple copies of specific nucleic acid sequences that provide a means for detection and quantification of specific nucleic acid sequences of interest. The invention provided herein fulfills this need and

provides additional benefits. Among other advantages, these methods can be performed isothermally, and detection and quantification of nucleic acid sequences can be achieved without the need for prior nucleic acid sequence amplification.

DISCLOSURE OF THE INVENTION

The invention provides methods and compositions for generating multiple copies of nucleic acid sequences of interest, and quantification and detection of these sequences.

Accordingly, in one aspect, the invention provides methods of generating multiple copies of and/or quanitfying a nucleic acid sequence of interest, said method comprising the steps of: (a) hybridizing a composite primer to a target polynucleotide, wherein the composite primer comprises an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion hybridizes from about 1 to about 10 nucleotides from the sequence of interest; (b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product is produced; and (c) cleaving the RNA portion of the primer extension product of (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved primer extension product dissociates from the target polynucleotide, wherein the primer extension product is of a size that when the RNA is cleaved the cleaved primer extension product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension, whereby multiple copies of the sequence of interest are produced.

In another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) hybridizing a composite primer to a single stranded target polynucleotide, wherein the composite primer comprises an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the

3' most nucleotide of the DNA portion of the composite primer hybridizes from about 1 to about 10 nucleotides from the sequence of interest; (b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product is produced; and (c) cleaving the RNA portion of the primer extension product of (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide, wherein the primer extension product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension. Cleavage of the primer extension product and dissociation of the cleaved product, followed by repetition of steps (a), (b) and (c) results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods for generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product, and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same

conditions as those for attachment of the oligonucleotides. Cleavage of the primer extension product and dissociation of the cleaved product, followed by repetition of steps (a)-(d) results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest. Attachment may or may not be covalent.

In another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest, said method comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product; and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved oligonucleotide combination product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved from the attached oligonucleotide combination product, the cleaved attached oligonucleotide product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, whereby multiple copies of the sequence of interest are produced.

In another aspect, the invention provides methods for generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of the target polynucleotide, wherein the portion of the

target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product, and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides. Cleavage of the primer extension product and dissociation of the cleaved product, followed by repetition of steps (a)-(d) results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods for generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product, and (d) cleaving the RNA portion of the attached

oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides. Cleavage of the primer extension product and dissociation of the cleaved product, followed by repetition of steps (a)-(d) results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods for generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of the target polynucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein each oligonucleotide is hybridizable to a portion of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product, and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides. Cleavage of the primer extension product and dissociation of the cleaved product, followed by repetition of steps (a)-(d) results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample, said method comprising the steps of: (a) hybridizing a composite primer to a target polynucleotide, wherein the composite primer comprises an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the DNA portion of the composite primer hybridizes from about 1 to about 10 nucleotides from the sequence of interest; (b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product comprising a detectable identifying characteristic is produced if the sequence of interest is present; and, (c) cleaving the RNA portion of the primer extension product of (b), if any, with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved primer extension product dissociates from the target polynucleotide, wherein the primer extension product is of a size that when the RNA is cleaved the cleaved primer extension product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates the presence of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) hybridizing a single stranded target polynucleotide with a composite primer that hybridizes to the target polynucleotide such that the 3' nucleotide of the primer is from about 1 nucleotide to about 10 nucleotides from the sequence of interest, said composite primer comprising an RNA portion and a 3' DNA portion; (b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product comprising a detectable identifying characteristic is produced; (c) cleaving the RNA portion of the primer extension product of (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide; wherein the primer extension product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide

under essentially the same conditions as those for primer extension, whereby detection of the primer extension product comprising the detectable identifying characteristic indicates the presence of the sequence of interest. Conversely, lack of detection of the cleaved primer extension product, or insignificant levels of cleaved primer extension product, indicates absence of the sequence of interest.

In yet another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample, said method comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic, whereby the attached oligonucleotide combination product is produced if the sequence of interest is present; and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c), if any, with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide, wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, whereby detection of the cleaved attached oligonucleotide combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest.

In yet another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest; and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide; wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, and whereby detection of the attached oligonucleotide combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest. Conversely, lack of detection of cleaved attached oligonucleotide combination product, or insignificant levels of cleaved attached oligonucleotide product, indicates absence of the sequence of interest.

In yet another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first

oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest; and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide; wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, and whereby detection of the attached oligonucleotide combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest.

In yet another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest; and (d) cleaving

the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide; wherein the cleaved attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, and whereby detection of the attached oligonucleotide combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest. Conversely, lack of detection of the attached oligonucleotide combination product, or insignificant levels of cleaved attached oligonucleotide combination product, indicates absence of the sequence of interest.

In yet another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the each oligonucleotide is hybridizable to a portion of the sequence of interest; (b) optionally extending the first oligonucleotide; (c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest; and (d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved product dissociates from the target polynucleotide; wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, and whereby detection of the attached oligonucleotide

combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest. Conversely, lack of detection of the attached oligonucleotide combination product, or insignificant levels of cleaved attached oligonucleotide combination product, indicates absence of the sequence of interest.

In yet another aspect, the invention provides method of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising: incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, primer extension and RNA cleavage, such that a primer extension product is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product from the target polynucleotide.

In still another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) combining a target polynucleotide; a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; and (b) incubating the mixture of step (a) under conditions that permit primer hybridization, primer extension and RNA cleavage, said conditions including the necessary substrates and buffer conditions for primer hybridization, primer extension such that a primer extension product comprising the RNA portion of the composite primer is produced, and RNA cleavage, wherein primer

extension product is of a size that when the RNA portion is cleaved the product dissociates from the target polynucleotide under said conditions. Cleavage of RNA from the RNA/DNA hybrid of the primer extension product results in dissociation of the product from the target polynucleotide. Dissociation of the cleaved product results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (c) optionally a DNA polymerase; (d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and (e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide, wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product is produced, and wherein the attached oligonucleotide combination product is of a size such that cleavage of RNA from the attached oligonucleotide combination product results in dissociation of the cleaved attached oligonucleotide product.

In yet another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping

portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said target polynucleotide; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide such that an attached oligonucleotide combination product comprising the RNA portion of the composite primer is produced, and RNA cleavage, wherein attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions. Cleavage of RNA from the RNA/DNA hybrid of the attached oligonucleotide combination product results in dissociation of the product from the target polynucleotide. Dissociation of the cleaved product results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In yet another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to

the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said target polynucleotide; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide such that an attached oligonucleotide combination product comprising the RNA portion of the composite primer is produced, and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and second oligonucleotide, wherein attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions. Cleavage of RNA from the RNA/DNA hybrid of the attached oligonucleotide combination product results in dissociation of the product from the target polynucleotide. Dissociation of the cleaved product results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In yet another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein the second

oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said target polynucleotide; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide such that an attached oligonucleotide combination product comprising the RNA portion of the composite primer is produced, and RNA cleavage, wherein attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions. Cleavage of RNA from the RNA/DNA hybrid of the attached oligonucleotide combination product results in dissociation of the product from the target polynucleotide. Dissociation of the cleaved product results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In yet another aspect, the invention provides methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, and wherein each oligonucleotide is hybridizable to a portion of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second

oligonucleotide to each other when said oligonucleotides are hybridized to said target polynucleotide; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide such that an attached oligonucleotide combination product comprising the RNA portion of the composite primer is produced, and RNA cleavage, wherein attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions. Cleavage of RNA from the RNA/DNA hybrid of the attached oligonucleotide combination product results in dissociation of the product from the target polynucleotide. Dissociation of the cleaved product results in cycling of the process and accumulation of multiple copies of polynucleotides comprising the complement of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer

extension product from the target polynucleotide, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) combining a target polynucleotide; a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; a DNA polymerase; and an enzyme that cleaves RNA from an RNA/DNA hybrid; (b) incubating the mixture of step (a) under conditions that permit primer hybridization, primer extension to generate a primer extension product comprising a detectable identifying characteristic that indicates presence of the sequence of interest, and RNA cleavage, said conditions including the necessary substrates and buffer conditions for primer hybridization, primer extension such that a primer extension product comprising the RNA portion of the composite primer is produced, and RNA cleavage, and wherein the primer extension product is of a size that when the RNA portion is cleaved the product dissociates from the target polynucleotide under said conditions, whereby the primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest. Conversely, lack of detection of the attached primer extension product, or insignificant levels of cleaved primer extension product, indicates absence of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide

is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (c) optionally a DNA polymerase; (d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and (e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide, wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product comprising a detectable identifying characteristic is produced, and wherein the attached oligonucleotide combination product is of a size such cleavage of the RNA from the attached oligonucleotide combination product results in dissociationg of the cleaved attached oligonucleotide combination product from the target polynucleotide, whereby detection of the cleaved attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first

oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said nucleic acid sequence of interest; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide and RNA cleavage, and wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions, whereby detection of the covalently attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleic acid sequence of interest. Conversely, lack of detection of the attached oligonucleotide combination product, or insignificant levels of cleaved attached oligonucleotide combination product, indicates absence of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said

nucleic acid sequence of interest; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide and RNA cleavage, and wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions, whereby detection of the covalently attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleic acid sequence of interest. Conversely, lack of detection of the attached oligonucleotide combination product, or insignificant levels of cleaved attached oligonucleotide combination product, indicates absence of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said nucleic acid sequence of interest; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide

extension, attachment of the first oligonucleotide and the second oligonucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic that indicates presence of the sequence of interest and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide and RNA cleavage, and wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions, whereby detection of the covalently attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleic acid sequence of interest. Conversely, lack of detection of the attached oligonucleotide combination product , or insignificant levels of cleaved attached oligonucleotide combination product, indicates absence of the sequence of interest.

In another aspect, the invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising the steps of: (a) combining a target polynucleotide; a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion, wherein each oligonucleotide is hybridizable to at least one nucleotide of the sequence of interest; optionally a DNA polymerase; an enzyme that cleaves RNA from an RNA/DNA hybrid; and an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to said nucleic acid sequence of interest; and (b) incubating the mixture of step (a) under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and the second oligonucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying

characteristic that indicates presence of the sequence of interest and RNA cleavage, said conditions including the necessary substrates and buffer conditions for oligonucleotide hybridization, optionally oligonucleotide extension, attachment of the first oligonucleotide and second oligonucleotide and RNA cleavage, and wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the product dissociates from the target polynucleotide under said conditions, whereby detection of the covalently attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleic acid sequence of interest.

In another aspect, the methods of determining whether a nucleic acid sequence of interest is present or absent in a sample further comprise determining whether two or more sequences of interest are present or absent in a sample, said method comprising using one or more different composite primers, wherein the detectable identifying characteristics of the cleaved primer extension products corresponding to two different sequences of interest are different from each other.

In another aspect, the methods of determining whether a nucleic acid sequence of interest is present or absent in a sample further comprise determining whether two or more different sequences of interest are present or absent in a sample, said method using a two or more sets of first and second oligonucleotides, wherein the detectable identifying characteristics of the cleaved oligonucleotide attachment products corresponding to two or more different sequences of interest are different from each other.

In another aspect, the invention provides methods for identifying an altered sequence of interest in a sample comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the altered sequence of interest; (c) a DNA polymerase; and (d) an enzyme

that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization and primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that a primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension product dissociates from the target polynucleotide, whereby the cleaved primer extension product is characterized to identify the altered sequence of interest.

In yet another aspect, the invention provides methods of identifying an altered sequence of interest in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the altered sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, and primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension product dissociates from the target polynucleotide, and wherein production of detectably fewer cleaved primer extension products from the target as compared to the amount of cleaved primer extension products produced from a reference template comprising the sequence of interest indicates that the target polynucleotide contains an altered sequence of interest.

In another aspect, the invention provides methods of identifying an altered sequence of interest in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a first

oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (c) optionally a DNA polymerase; (d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and (e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide, wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product comprising a detectable identifying characteristic is produced, and wherein the attached oligonucleotide combination product is of a size that when RNA is cleaved from the attached oligonucleotide combination product, the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide, wherein production of detectably fewer cleaved oligonucleotide attachment products from the target as compared to the amount of cleaved oligonucleotide attachment products produced from a reference template comprising the sequence of interest indicates that the target polynucleotide contains an altered sequence of interest. Various exemplary embodiments of conditions that permit primer extension to produce primer extension products of suitable sizes are described herein. For example, in some embodiments, conditions that permit primer extension comprise the presence of at least one terminator deoxyribonucleotide triphosphate or analog thereof. In other embodiments, conditions that permit primer extension comprise the absence of an essential nucleotide.

The enzymes and agents which may be used in the methods and compositions are described herein. For example, the enzyme that cleaves RNA

may be RNase H. In another example, the agent that effects attachment of hybridized oligonucleotides may be DNA ligase. In another example, the DNA polymerase may lack strand displacement activity.

Examples of detectable identifying characteristic of the primer extension product and attached oligonucleotide combination product are also described herein. Examples for primer extension products include size of the product, sequence of the non-primer portion of the product, and detectable signal associated with the product. Examples for attached oligonucleotide combination products include size of the attached oligonucleotide combination product, sequence of the attached oligonucleotide combination product, and detectable signal associated with the attached oligonucleotide combination product. Detectable signal may be associated with a label on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension. Detectable signal may also be associated with interaction of two labels. For example, in primer extension products, one label may be on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension and another label is on a deoxyribonucleotide triphosphate or analog thereof located in the primer portion of the primer extension product. In another example, in attached oligonucleotide combination products, one label may be on one oligonucleotide and another label on another oligonucleotide.

Various embodiments of the composite primer used in the methods of the invention are described herein. For example, in some embodiments, a composite primer comprises a 5'-RNA portion and a 3' DNA portion. In embodiments involving attachment of oligonucleotide pairs, the second oligonucleotide may be a composite primer comprising a 3'-RNA portion and a 5'-DNA portion.

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, for example, the invention provides compositions comprising a composite primer comprising a 5' RNA portion and a 3' DNA portion. In another aspect, the invention provides compositions comprising a composite primer comprising a 3'-

RNA portion and a 5'-DNA portion. In another aspect, for example, the invention provides compositions comprising two composite primers that are hybridizable to two non-overlapping sequences of a target polynucleotide. In yet another aspect, the invention provides compositions comprising a cleaved primer extension product comprising a unique identifying characteristic the detection of which indicates presence of a nucleic acid sequence in a sample. In still another aspect, the invention provides compositions comprising a cleaved attached oligonucleotide combination product comprising a unique identifying characteristic the detection of which indicates presence of a nucleic acid sequence in a sample.

In yet 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. For example, the invention provides compositions comprising a complex of (a) a target polynucleotide; and (b) a composite primer. In another example, the invention provides compositions comprising a complex of (a) a polynucleotide; and (b) two oligonucleotides hybridized to non-overlapping sequences of the polynucleotide, wherein at least one of the oligonucleotides is a composite primer.

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 methods. For example, the invention provides kits for generation of multiple copies, detection or quantification of a nucleic acid sequence of interest, comprising a composite primer comprising a 3' DNA portion and an RNA portion. The composite primer in the kits can be any described herein. The kits can contain further components, such as terminator deoxyribonucleotide triphosphates and/or at least one but not all four types of deoxyribonucleotide triphosphates. In another example, the invention provides kits for generation of multiple copies, detection or quantification of a nucleic acid sequence of interest, comprising a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, and wherein the two oligonucleotides

hybridize to non-overlapping portions of a target polynucleotide. These kits may further comprise an agent that effects attachment of said first oligonucleotide and second oligonucleotide when said oligonucleotides are hybridized to the target polynucleotides, and said agent may be DNA ligase. Any of the preceding kits can further comprise an enzyme that cleaves RNA from a RNA/DNA hybrid. The enzyme may be RNase H.

In another aspect, the invention provides kits for generation of multiple copies of a sequence of interest, comprising a composite primer comprising a 3' DNA portion and an RNA portion, and instructions for a method for generating multiple copies, said method comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, primer extension and RNA cleavage, such that a primer extension product is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product, whereby multiple copies are generated.

In yet another aspect, the invention provides kits for generation of multiple copies of a sequence of interest, comprising a composite primer comprising a 3' DNA portion and an RNA portion, and instructions for a method for generating multiple copies, said method comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an

RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, primer extension and RNA cleavage, such that a primer extension product is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product, whereby multiple copies are generated.

In another aspect, the invention provides kits for determining whether a sequence of interest is present or absent in a sample, comprising a composite primer comprising a 3' DNA portion and an RNA portion, and instructions for a method of determining whether a sequence of interest is present or absent in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target polynucleotide; (b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest; (c) a DNA polymerase; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid, wherein the incubation is under conditions that permit primer hybridization, primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension dissociates from the target polynucleotide, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

In another aspect, the invention provides kits for generation of multiple copies of a sequence of interest, comprising a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the

target nucleotide that is hybridizable to the second oligonucleotide, and wherein the oligonucleotides singly or in combination comprise at least one nucleotide of the sequence of interest; and an agent that effects attachment of said first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide.

In another aspect, the invention provides kits for determining whether a sequence of interest is present or absent in a sample, comprising a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, and wherein the oligonucleotides singly or in combination comprise at least one nucleotide of the sequence of interest; and an agent that effects attachment of said first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide.

In another aspect, the invention provides systems for generating multiple copies of, quantifying or determining whether a nucleic acid sequence of interest is present or absent in a sample, comprising (a) a composite primer comprising a 3' DNA portion and an RNA portion; (b) DNA polymerase; and (c) an enzyme that cleaves RNA from an RNA/DNA hybrid (such as RNase H). The composite primer may be any (one or more) described herein. The systems may further comprise terminator deoxyribonucletide triphosphosphates and/or at least one but not all four types of deoxyribonucleotide triphosphates.

In another aspect, the invention provides systems for generating multiple copies of, quantifying or determining whether a nucleic acid sequence of interest is present or absent in a sample comprising (a) a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, and wherein the oligonucleotides singly or in combination comprise the sequence of interest; (b) optionally DNA polymerase; (c) an agent

that effects attachment of said first oligonucleotide and second oligonucleotide when said oligonucleotides are hybridized to the target polynucleotides; and (d) an enzyme that cleaves RNA from an RNA/DNA hybrid (which may be RNase H).

In still another aspect, the invention provides reaction mixtures (or compositions comprising reaction mixtures) which contain various combinations of components described herein. For example, the invention provides reaction mixtures comprising (a) a target polynucleotide; (b) a composite primer comprising a 3' DNA portion and an RNA portion; and (c) DNA polymerase. A reaction mixture can further comprise terminator deoxyribonucletide triphosphates and/or at least one but not all four types of deoxyribonucleotide triphosphates. In another example, the invention provides reaction mixtures comprising (a) a target polynucleotide; (b) a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, and wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, and wherein the oligonucleotides singly or in combination comprise the sequence of interest; (c) optionally DNA polymerase; and (d) an agent (such as DNA ligase) that effects attachment of the first oligonucleotide and second oligonucleotide when the oligonucleotides are hybridized to the target polynucleotides.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 illustrates a first mechanism of dissociation of a primer extension product.

FIGURE 2 illustrates a second mechanism of dissociation of a primer extension product.

FIGURE 3 illustrates an oligonucleotide attachment-based method of generating multiple copies of a nucleic acid sequence.

FIGURE 4 shows results demonstrating that a single nucleotide sequence of interest in a template can be detected using limited primer extension.

FIGURE 5 shows electropherograms from limited primer extension reactions in which multiple differing sequences of interest contained in a single template polynucleotide are detected.

MODES FOR CARRYING OUT THE INVENTION

The invention provides methods, compositions and kits for generating multiple copies of and/or detecting and/or quantifying nucleic acid sequences of interest and/or alterations in nucleic acid sequences of interest. The methods generally comprise using RNA/DNA composite primers for limited primer extension or attachment, preferably covalent attachment, of pairs of oligonucleotides, wherein cleavage of the composite primer component in the primer extension or attached oligonucleotide combination product results in uninduced dissociation of the product from a target polynucleotide in the sense that an additional reaction step and/or condition generally is not performed (upon cleavage) in order to effect dissociation (such as increasing temperature). Thus, these methods provide for generation of multiple copies and quantification of a sequence of interest. Furthermore, these methods can be used for detecting the presence or absence of a sequence of interest (which can, for example, be a single nucleotide base or more than one nucleotide bases). Thus, in some embodiments, these methods can be used for detecting variant sequences such as single nucleotide polymorphisms. These methods can also be used to detect and/or identify alteration(s) in sequences of interest (or a portion of a sequence of interest).

In one aspect, as a general summary, the methods of generating multiple copies of and/or quantifying nucleic acid sequences of interest work as follows. A composite primer binds to a target polynucleotide at a position in close proximity to a sequence of interest. Primer extension is allowed to occur to include a copy of (generally a complementary sequence of) the sequence of interest. Cleavage of the composite primer that has been extended causes a significant shortening of the primer extension product, resulting in dissociation of the primer extension product from the target polynucleotide, thus allowing accumulation of primer

extension products in the reaction mixture. Quantification of accumulated primer extension product further permits quantification of the sequence of interest in a sample.

In another aspect, as a general summary, the detection methods work as follows. Accumulated cleaved primer extension product produced by the methods of the present invention are characterized to determine the presence or absence of the sequence of interest in a sample.

In another aspect, as a general summary, the methods of identifying or detecting an altered sequence of interest work as follows. Accumulated cleaved primer extension product produced by the methods of the present invention is characterized to determine the whether an altered sequence of interest is present or absent in a sample. In another aspect, lack of, or insignificant amounts of, accumulated cleaved primer extension product indicates the presence of an altered sequence of interest in a sample. Absence, or insignificant amounts of, accumulated cleaved primer extension product can also indicate the absence of the sequence of interest.

In yet another aspect, as a general summary, the methods of generating multiple copies of and/or quantifying nucleic acid sequences of interest work as follows. A first oligonucleotide and a second oligonucleotide are hybridized to two non-overlapping sequences of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion. At least one oligonucleotide comprises a sequence hybridizable to at least one nucleotide of a sequence of interest. At least one of the oligonucleotides is a composite primer. The first oligonucleotide is extended by polymerization if there is a gap of at least one nucleotide between the two oligonucleotides. The two oligonucleotides are then attached to each other if they are hybridized to the same target polynucleotide strand. Cleavage of the composite primer(s) in the covalently attached oligonucleotide combination product causes a shortening of the attached

oligonucleotide combination product, resulting in dissociation of the combination product from the target polynucleotide, thus allowing accumulation of the product in the reaction mixture. Quantification of accumulated oligonucleotide combination product further permits quantification of the sequence of interest in a sample.

In still another aspect, as a general summary, the detection methods work as follows. Accumulated cleaved attached oligonucleotide combination product generated by the methods of generating multiple copies of a nucleic acid sequence of interest of the present invention are characterized to determine the presence or absence of the sequence of interest in a sample. Conversely, lack of, or insignificant amounts of, accumulated cleaved attachment product indicates absence of the sequence of interest in a sample. In another aspect, absence or reduction of accumulated cleaved attached oligonucleotide combination product indicates the presence of an altered sequence of interest in a sample. The methods of the invention are further useful for copying and/or quantifying and/or determining whether a plurality of sequences of interest are present or absent in a sample. Sequence(s) of interest may be present in one target polynucleotide or in a plurality of target polynucleotides. In one aspect, a plurality of sequences of interest are present on a single target polynucleotide, for example, as when more than one locus on a single genomic polynucleotide are analyzed, or for example, as when a plurality of sequences of interest are present on a single cDNA polynucleotide. In another aspect, a plurality of sequences of interest are present on a single or a plurality of target polynucleotides, for example, as when a plurality of sequences of interest are present in a plurality of genomic polynucleotides or a plurality of cDNA polynucleotides. In another aspect, a plurality of sequences of interest present on a plurality of target polynucleotides comprise sequences of interest that are variant sequences to one another, as, for example, when the sequences of interest comprise different alleles, or a wildtype (native) and mutant (non-wildtype) sequences, or for example, different polymorphic sequences.

It is understood that, with respect to all embodiments described herein, as generally "comprising" components or aspects, the invention also includes embodiments which "consist essentially of" these components or aspects. The invention also includes embodiments which "consist of" these components or aspect. This applies to all embodiments described herein.

"Absent" or "absence" of product, and "lack of detection of product" as used herein includes insignificant, or *de minimus* levels, generally due to lack of significant accumulation of product due to cycling.

Advantages of the invention

The methods of copying and/or quantifying and/or determining (detecting) whether a sequence of interest is present (i.e. presence or absence of a sequence of interest) or absent and/or identifying an altered sequence of interest provide several significant advantages over other known methods.

The requirement of cleavage of the RNA portion of the composite primer in the hybridized and extended and/or attached form, by a ribonuclease such as RNaseH, whereby dissociation of the cleaved product occurs, results in the copying and/or quantifying and/or detection and/or identification of sequences of interest on DNA exclusively. Thus, it is possible to use the methods of the invention with DNA target polynucleotides, e.g., genomic DNA, in the presence of excess mRNA. This feature is useful for the accurate quantification of gene dosage. When the target polynucleotide is RNA, the target is first transcribed to produce cDNA which can serve as a target polynucleotide in the methods of the invention.

The methods of the invention do not require thermocycling in that extension, hybridization, and/or attachment can be performed isothermally, and the cleaved primer extension product and cleaved oligonucleotide-attachment product are of a size that when the RNA component is cleaved the product disassociates from the target polynucleotide generally under essentially the same conditions as those for primer extension, hybridization, and/or attachment. This feature provides numerous advantages, including facilitating automation and adaptation for high through-put procedures. For example, other methods that

have been reported that require thermal cycling for the separation of primer extension and/or attachment products from the target sequence. The isothermal reaction is faster than that afforded by thermal cycling and is suitable for performing the methods of the invention in miniaturized devices.

The amount of cleaved primer extension or attachment product produced is linearly related to the amount of nucleic acid sequence of interest in the sample. Thus, the methods of the invention are useful for quantification of the sequence of interest in a sample. For example, amount of cleaved primer extension or attachment product is compared to the amount of product obtained in a reference sample comprising a known amount of a sequence of interest. The method is also useful for quantification of the amounts of two sequences of interest in a sample, for example, determining the relative amount of two alleles in a sample.

In primer-extension-based methods of the invention, the composite primer binds to a target polynucleotide at a position in close proximity to a sequence of interest. This feature permits the use of a single composite primer with template polynucleotides comprising different sequences of interest, for example, that are variant sequences with respect to each other, for example, template polynucleotides comprising different alleles. This method is further useful for identifying alterations in a sequence of interest without previous knowledge of the identity of the altered sequences of interest. By contrast, other methods known in the art require primers that specifically hybridize to the sequence of interest, necessitating the use of sets of primers that are each specific to particular sequences of interest.

The cleaved primer extension product is of a size that when the RNA component is cleaved the product disassociates from the target polynucleotide under essentially the same conditions as those for primer extension. Thus, the primer extension-based methods do not require use of DNA polymerase possessing strand displacement activity. Thus, for example, polymerases lacking strand displacement activity, but specialized for increased incorporation of modified polynucleotides (analogs), including labeled polynucleotides, may be used in the primer extension-based methods of the invention.

In attachment-based methods, at least one oligonucleotide comprises a sequence hybridizable to a sequence of interest. This feature provides specificity because in the absence of the sequence of interest, substantial oligonucleotide hybridization is absent and accumulation of cleaved attachment product is absent. "Absent" or "absence" of product, and "lack of detection of product" as used herein includes insignificant, or *de minimus* levels, generally due to lack of significant accumulation of product due to cycling. Thus, attachment-based methods of the invention are useful for detection and/or copying and/or quantifying a sequence of interest, wherein accumulation of cleaved attachment product indicates presence of a sequence of interest. Attachment-based methods of the invention are also useful for identifying the presence of an altered sequence of interest, wherein absence or reduction of accumulated cleaved attachment product may indicate presence of an altered sequence of interest.

The methods of the invention are further useful for analysis of a plurality of sequences of interest in a given reaction mixture (multiplex analysis). The sequences of interest may be part of a single target polynucleotide, or may be part of different target polynucleotides, which may be present in a single test sample. For example, the methods of the invention are useful for the copying, quantification, and/or detection of multiple sequences of interest in a single genomic DNA sample in a single reaction. Similarly, a plurality of polymorphisms present at a single site, for example, a plurality of alleles or a heterozygous mixture of alleles, can be copied, quantified and/or detected in a single reaction.

It is understood that, generally, "detection" of a product (such as a cleaved product comprising a detectable identifying characteristic) means detection of significant amounts of product arising from cycling (.e., repeated cycles of a reaction). The cycling results in accumulated product. Lack of cycling (due to, or example, absence of sequence of interest) results in *de minimus*, or insignificant amount of product which, for purposes of the methods of the invention, is not "detected".

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).

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

Definitions

"Multiple copies,"as used herein, means at least 2 copies. A "copy" does not necessarily mean perfect sequence complementarity or identity to the template sequence. A "copy" includes a nucleic acid sequence that is hybridizable (preferably complementary) to the sequence of interest. Copies can include nucleotide analogs such as deoxyinosine, intentional sequence alterations (such as sequence alterations introduced through a primer comprising a sequence that is hybridizable, but not complementary, to the template), and/or sequence errors that occur during DNA polymerization.

A "sequence of interest," as used herein, is a nucleic acid sequence (including one or more nucleic acid sequences) the copying and/or quantification, and/or the detection of the presence or absence of which is desired. The identity of a sequence of interest is known. A sequence of interest may be, for example, a native (e.g. wildtype) or nonnative (e.g., mutant) nucleotide sequence. Examples of mutant sequence include, for example, a substitution, deletion, insertion, transversion of one or more nucleotides, or any combination thereof. As used herein, a sequence of interest can be a single nucleotide base or more than a single

nucleotide base. A sequence of interest can be a known polymorphic sequence, including, for example, single nucleotide polymorphism. A sequence of interest could be a variant sequence with respect to another sequence of interest (which in this context, could be considered a reference sequence), termed an "altered" sequence of interest. The differences between the sequence of interest (reference sequence) and the altered sequence of interest may comprise the addition, deletion, substitution, and/or transversion of one or more nucleotides, or any combination thereof. The sequence of an altered sequence of interest may or may not be known. A sequence of interest also includes the complementary sequence. For example, reference to generating multiple copies of a sequence of interest using the methods described herein refers to and includes making copies of the complementary sequence.

A "variant" sequence, as used herein, includes naturally or nonnaturally occurring variants of the polynucleotide sequence (e.g. degenerate variants, allelic variants, single nucleotide polymorphisms ("SNPs"), etc.). In general, allelic variants contain 15-25% base pair (bp) mismatches and can contain as little as 5-15%, or 2-5%, or 1-2% bp mismatch, as well as a single bp mismatch.

A "target polynucleotide," as used herein, is a polynucleotide known or suspected to comprise a sequence of interest. The terms "target sequence," "template", "template DNA," "template polynucleotide," "target nucleic acid," "target polynucleotide," and variations thereof, are used interchangeably.

In some embodiments, e.g., detection of an altered sequence of interest, 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 "sequence of interest" 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.

"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. As used herein, "DNA" and "RNA" include modified nucleotide or ribonucleotides, linakes, etc. as described herein. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and their analogs. If present, modification to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications include, for example, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, cabamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example, proteins (e.g., nucleases, toxins, antibodies, signal peptides, ply-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide(s). Further, any of the hydroxyl groups ordinarily present in the sugars may be replaced, for example, by phosphonate groups, phosphate groups, protected by standard protecting groups, or activated to prepare additional linkages to additional nucleotides, or may be conjugated to solid supports. The 5' and 3' terminal OH can be phosphorylated or substituted with amines or organic capping groups moieties of from 1 to 20 carbon atoms. Other hydroxyls may also be derivatized to standard protecting groups. Polynucleotides can also contain analogous forms of ribose or deoxyribose sugars that are generally known in the art, including, for example, 2'--O-methyl-, 2'-O-allyl, 2'-fluoro- or 2'-azido-ribose, carbocyclic sugar analogs, 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. Not all linkages in a polynucleotide need be identical. 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 and DNA.

A "primer" is generally a short single-stranded polynucleotide (often referred to as an oligonucleotide), generally with a free 3'-OH group, that binds to a target polynucleotide and/or a sequence of interest. It can be used to promote polymerization of a polynucleotide complementary to the target. It can also be used for attachment to another polynucleotide or oligonucleotide (such as another primer).

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

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. For example, an RNA portion

which is adjacent to a 5' DNA portion of a composite primer directly abuts that region.

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

"A", "an" and "the", and the like, unless otherwise indicated include plural forms. For example, a method of copying and/or quantifying and/or determining whether a sequence of interest is present or absent, as described herein includes copying and/or quantifying and/or determining one as well as more than one sequence of interest.

"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 conditions also depend on what event is desired, such as hybridization, cleavage, primer extension or ligation.

"Essentially the same conditions," as used herein, refer to conditions that are similar, but not necessarily identical (although the term does include identical conditions), in their chemical, physical, and/or physicochemical properties and characteristics. Conditions are essentially the same if variances between the conditions in question are within variance ranges that one of skill in the art would not expect to substantially alter the primary function or effect of the conditions with respect to the reactions for which the conditions are used. The reactions for which the conditions are used include primer extension (polymerization), oligonucleotide attachment (such as ligation by DNA ligase), primer or oligonucleotide hybridization, and dissociation of cleaved primer extension and attached oligonucleotide combination products.

"Conditions that permit limited primer extension," as used herein, are conditions that permit, enhance, facilitate, and/or are conducive to limited polymerization of deoxyribonucleotide triphosphates (dNTPs) and/or analogs thereof from a primer that is hybridized to a target polynucleotide. Such conditions are known in the art, and are discussed herein. Likewise, conditions that "permit" primer extension are conditions that are conducive to polymerization of deoxyribonucleotide triphosphates (dNTPs) and/or analogs thereof from a primer that is hybridized to a target polynucleotide.

"Microarray" and "array," as used interchangeably herein, refer to an arrangement of a collection of nucleotide sequences in a centralized location. Arrays can be on a surface, for example, a solid substrate, such as a glass slide, or on a semi-solid substrate, such as nitrocellulose membrane. The nucleotide sequences can be DNA, RNA, or any permutations thereof.

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. 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. 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.

"Non-primer portion of a primer extension product," as used herein, refers to the portion of a primer extension product that is the result of polymerization (extension) from a primer.

"Dissociation" of a product from a polynucleotide, as used herein, refers to dissociation of a product (generally cleaved primer extension or cleaved attached oligonucleotide combination product) from a target polynucleotide due to decreased affinity of hybridization to the target polynucleotide following cleavage of the cleavable component of the product.

As is well understood in the art, "affinity" or "binding affinity" means a measure of the strength of binding between two or more moieties; non-limiting examples of such bonding are hydrogen bonding, electrostatic interactions and hydrophobic interactions. In particular, "affinity" when used in reference to hybridized nucleic acids refers to hydrogen bonding between at least partly complementary nucleic acid strands 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 nucleic acid strands are in the double-stranded or hybridized form, under given hybridization conditions.

"Hybridizable," as used herein, refers to the capability and/or ability of two polynucleotide sequences to hybridize through complementary base pairing, under conditions used in any of the methods described herein; i.e., at the temperature, pH, ionic concentrations, and the like, used in carrying out the methods of the invention. As such, a sequence (such as a primer) which is

hybridizable to another sequence (such as a target polynucleotide) hybridizes to that sequence under suitable conditions.

"Uninduced dissociation" of a product from a polynucleotide, as used herein, refers to dissociation of a product (generally cleaved primer extension or cleaved attached oligonucleotide combination product) from a target polynucleotide due to decreased affinity of hybridization to the target polynucleotide following cleavage of the cleavable component of the product. Uninduced dissociation includes displacement of a cleaved product by the binding of an uncleaved primer (oligonucleotides) to the sequence of target polynucleotide that is complementary to the cleaved portion of the product. Uninduced dissociation is not caused by a separate reaction step that is aimed primarily at causing dissociation of the product. Such separate reaction steps are known in the art, and include, for example, increasing the temperature of reaction and adding material known to reduce nucleic acid binding affinity.

"Non-overlapping portions" of a sequence, as used herein, refers to two or more sequences of a sequence of interest or polynucleotide that are located in distinctly separate locations within the sequence of interest or polynucleotide. Non-overlapping portions of a sequence can be contiguous with each other or separated by one or more nucleotides.

"Covalent attachment" or "covalently-attached" as used herein refers to linkage of two nucleic acid molecules (generally oligonucleotides) by a covalent bond. "Attachment" or "attached" as used herein refers to linkage of two nucleic acid molecules (generally oligonucleotides), which may be covalent or non-covalent.

"Attached oligonucleotide combination product" refers to a product comprising two oligonucleotides, wherein the 3' end of a first oligonucleotide is attached to the 5' end of a second oligonucleotide. "Covalently-attached oligonucleotide combination product" refers to a product comprising two oligonucleotides, wherein the 3' end of a first oligonucleotide is covalently attached to the 5' end of a second oligonucleotide.

"Detectable identifying characteristic," as used herein, refers to one or more characteristics of a reaction product that indicates its identity, wherein the characteristic is detectable by methods known in the art.

"Terminator nucleotide," or "terminator dNTP," as used herein, refers to an enzymatically-incorporable nucleotide, or analog thereof, in which the sugar moiety does not support incorporation of subsequent nucleotides or analogs thereof. Examples of terminator nucleotides include those in which the nucleobase is a purine, a 7-deaza-purine, a pyrimidine, a normal nucleobase or a common analog thereof and the sugar moiety is a pentose which includes a 3'-substituent that blocks further synthesis, such as a dideoxyribonucleotide triphosphate (ddNTP). Substituents that block further synthesis include, but are not limited to, amino, deoxy, halogen, alkoxy and aryloxy groups. Exemplary terminators include, but are not limited to, those in which the sugar-phosphate ester moiety is 3'-(C_1-C_6) alkylribose-5'-triphosphate, 2'-deoxy-3'-(C_1-C_6)alkylribose-5'-triphosphate, 2'-deoxy-3'-(C_1-C_6) alkoxyribose-5'-triphosphate, 2'-deoxy-3'-(C_5-C_{14}) aryloxyribose-5'-triphosphate, 2'-deoxy-3'-(C_1-C_6) alkoxyribose-5'-triphosphate, 2'-deoxy-3'-(C_5-C_{14})aryloxyribose-5'-triphosphate, 2'-deoxy-3'-haloribose-5'-triphosphate, 2'-deoxy-3'-aminoribose-5'-triphosphate, 2',3'-dideoxyriose-5'-triphosphate or 2',3'-didehydroribose-5'-triphosphate.

"Essential nucleotide" or "essential dNTP," as used herein, refers to a dNTP, or analog thereof, that must be provided in a reaction mixture for primer extension to continue unlimited primer extension. Whether a nucleotide or dNTP is essential is determined within the context of the nucleic acid sequence that serves as the template for polymerization (primer extension).

"Analyte", as used herein, refers to a compound which is desired to be analyzed, for example, a compound for which the properties, location, quantity and/or identity of which is desired to be characterized.

Methods of the invention

Methods of generating multiple copies of and/or quantifying a nucleic acid sequence of interest

Primer extension-based methods of generating multiple copies of and/or quantifying nucleic acid sequences

The invention provides methods of generating multiple copies of and/or quantifying a sequence of interest in a target polynucleotide, comprising reacting the following combination: (a) a single-stranded target polynucleotide comprising a nucleic acid sequence of interest; (b) a composite primer comprising an RNA portion and a 3' DNA portion; (c) a DNA polymerase; (d) deoxyribonucleoside triphosphosphates (dNTPs), or suitable analogs that singly or in combination permits only limited primer extension; and (e) an enzyme, such as RNase H, which cleaves RNA from an RNA/DNA duplex. The combination is subjected to suitable conditions such that (a) the composite primer hybridizes to the target polynucleotide; (b) limited primer extension occurs from the composite primer, to form a duplex; (c) an enzyme such as RNase H cleaves RNA of the composite primer from the RNA/DNA duplex; (d) the cleaved primer extension product dissociates from the target polynucleotide; (e) another composite primer hybridizes to the target polynucleotide, and another round of primer extension occurs; and (f) cleaved primer extension product accumulates. This method is illustrated in Figures 1 and 2.

In one embodiment, the sequence of interest that is copied is a single nucleotide base. Examples of a sequence of interest that is a single nucleotide base is a single nucleotide polymorphic sequence, or a mutant sequence due to a mutation such as a nucleotide substitution, deletion or insertion. In another embodiment, the sequence of interest that is copied comprises more than one nucleotide base, which can be a mutant sequence due to nucleotide substitution, deletion, insertion, transversion, or any combination thereof. In embodiments wherein the sequence of interest comprises more than one nucleotide base, the

sequence of interest comprises preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 13 nucleotides, even more preferably from about 2 to about 11 nucleotides, and most preferably from about 2 to about 10 nucleotides.

The composite primers are designed to hybridize to a sequence on a target polynucleotide that is in close proximity to the sequence of interest. Accordingly, methods of the invention comprise hybridizing a composite primer to a single stranded target polynucleotide such that the 3' most nucleotide of the primer is from preferably about 1 nucleotide to about 10 nucleotides, more preferably about 1 nucleotide to about 8 nucleotides, even more preferably about 1 to about 6 nucleotides, and most preferably about 1 to about 4 nucleotides from the sequence of interest.

The hybridized primer is extended to include at least the sequence of interest. Generally, the amount of primer extension is limited to minimize the size of the primer extension product. The primer extension product should be of a size that when the RNA component is cleaved (either partially or comletely) the product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension. Thus, a preferred size of primer extension product is one that when its RNA component is cleaved results in a product that binds to the target polynucleotide with less affinity, such that the frequency of dissociation from the target polynucleotide is increased compared to the uncleaved primer extension product. The frequency of dissociation is increased if there is an increase of preferably at least about 20%, more preferably at least about 40%, even more preferably at least about 60%, and most preferably at least about 80% increase of uninduced dissociation of cleaved product relative to uncleaved product. It is to be understood that the dissociation frequency is generally measured under essentially the same conditions as those for primer extension. In one embodiment, the reaction is performed isothermally (at a single temperature or about the same temperature). In another embodiment, the temperature is a temperature that permits specific hybridization of the composite

primer but is too high for significant hybridization of the cleaved primer extension product.

Limited primer extension can be achieved by any of a variety of methods known in the art. Generally and preferably, the extent of primer extension is limited by providing at least one terminator dNTP and/or by not providing at least one essential dNTP. Terminator dNTPs are known in the art, and include, for example, dideoxyriboncleotide triphosphates. An essential deoxyribonucleotide is a deoxyribonucleotide the presence of which in a reaction is required for unlimited primer extension. Thus, a deoxyribonucleotide is essential within the context of a particular sequence identity of the target polynucleotide.

In some embodiments, limited primer extension comprises addition of only one nucleotide to the 3' end of the primer. In these embodiments, the sequence of interest is a single base nucleotide, which is located in the target polynucleotide immediately 5' and adjacent to the nucleotide of the target polynucleotide to which the 3'-most nucleotide of the primer is hybridized. Primer extension can be limited by not providing at least one essential dNTP, or by providing a suitable terminator dNTP the incorporation of which into the primer extension product results in termination of primer extension.

In other embodiments, limited primer extension comprises addition of more than one nucleotide. In these embodiments, primer extension comprises addition of preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 12 nucleotides, even more preferably from about 2 to about 10 nucleotides, and most preferably from about 2 to about 8 nucleotides. In these embodiments, primer extension can be limited by not providing at least one essential dNTP, by providing at least one terminator nucleotide, or by both not providing at least one essential dNTP and providing at least one terminator nucleotide in a single reaction mixture. One of skill in the art would be able to determine the necessary combination based on the identity of target polynucleotide sequence adjacent to the hybridization site of the primer.

A reaction mixture can comprise a variety of dNTP configurations, such as a single terminator dNTP, a combination of terminator dNTPs, at least one

terminator dNTP and at least one but not all of the four dNTPs (which can include non-terminator forms of a terminator dNTP that is included), or at least one but not all of the four dNTPs (without any terminator dNTPs). The choice of what types of dNTPs to provide in a reaction is governed by factors that are evident to one skilled in the art. Among these factors are the sequence of interest, and knowledge and identity of the sequence flanking the sequence of interest.

A DNA polymerase capable of effecting extension of the hybridized primer along the target polynucleotide is used to generate primer extension product. Suitable polymerases are known in the art, including, for example Bst DNA polymerase, Bca DNA polymerase, E. coli DNA polymerase I, the Klenow fragment of E. coli DNA polymerase I, Taq DNA polymerase, T7 DNA polymerase (Sequenase), or recombinant derivatives (or mutants) thereof, which are selected for improved utilization of modified dNTPs and their analog, for example, AcycloPol-DNA polymerase. DNA polymerases with or without strand displacement activity are suitable for the primer extension-based methods of the invention. Factors that determine suitability of a polymerase are known to one skilled in the art, and include the temperature used to effect primer hybridization. For example, the availability of thermostable DNA polymerases can be used where primer hybridization at elevated temperatures enables greater specificity of primer hybridization to the target polynucleotide.

The use of composite primers in primer extension renders at least a portion of the primer extension product cleavable when hybridized to a target polynucleotide. A suitable cleaving reagent is provided in the reaction mixture to effect cleavage of the primer extension product. Cleavage of the product permits accumulation of cleaved primer extension product. Generally and preferably, this reaction occurs under essentially the same conditions as those for primer extension. Thus, suitable cleaving agents are agents that are active in these conditions. In preferred embodiments, the cleavable portion is RNA, in which case suitable cleaving agents are those that specifically cleave RNA in a RNA/DNA duplex, such as RNase H.

Without intending to be bound by theory, it is expected that uninduced dissociation of cleaved primer extension product leading to accumulation of said product occurs due to the lower affinity (due to reduced melting temperature of hybrid formation) of the smaller (cleaved) primer extension product (see Figure 1). The skilled practitioner will appreciate, however, that some uninduced dissociation of the uncleaved portion of the cleaved primer extension product from the target polynucleotide may result from binding of an uncleaved primer to the target polynucleotide (see Figure 2).

In a preferred embodiment, the resultant cleaved primer extension products in a reaction mixture are separated for analysis on a suitable matrix. Any of a number of methods can be used to effect the separation, as described in, for example, McIntosh et al. (PCT Pub. No. WO98/59066). Such methods include, but are not limited to, oligonucleotide array hybridization, mass spectrometry, flow cytometry, HPLC, FPLC, size exclusion chromatography, affinity chromatography, and gel electrophoresis. The amount of the product(s) is linearly related to the amount of the target polynucleotide comprising a sequence of interest in the sample. Thus, comparison of the amount of cleaved primer extension product obtained in a test sample with the amount of product obtained in a reference sample comprising a known amount of the sequence of interest provides quantification of the sequence of interest in the test sample. Methods of making such comparisons are known in the art.

It is appreciated that the methods of the invention are suitable for generating copies and/or quantifying more than one sequence of interest in a sample. That is to say, more than one sequence of interest may be copied and/or quantified simultaneously in a single reaction mixture. The different sequences of interest may be on the same or different target polynucleotides.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Oligonucleotide attachment-based methods of generating multiple copies of and/or quantifying nucleic acid sequences

The invention provides methods of generating multiple copies of and/or quantifying a sequence of interest in a target polynucleotide, comprising reacting the following combination: (a) a single-stranded target polynucleotide comprising a nucleic acid sequence of interest; (b) a first oligonucleotide and a second oligonucleotide that are hybridizable to non-overlapping sequences of a target polynucleotide, wherein at least one of the oligonucleotide comprises an RNA portion, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target polynucleotide that is hybridizable to the second oligonucleotide, and wherein at least one oligonucleotide comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest; (c) optionally a DNA polymerase; (d) an agent capable of effecting attachment, preferably covalent attachment, of the two oligonucleotides when the oligonucleotides are hybridized to the same target polynucleotide strand; (e) optionally dNTPs, or suitable analogs; and (f) an enzyme, such as RNase H, which cleaves RNA from an RNA/DNA duplex. The combination is subjected to suitable conditions such that (a) the oligonucleotides (at least one of which is a composite primer) hybridize to a target polynucleotide comprising the sequence of interest; (b) oligonucleotide extension from the first oligonucleotide by DNA polymerase occurs if there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and 5' end of the second oligonucleotide; (c) attachment of hybridized first oligonucleotide (or oligonucleotide extension product thereof) and second oligonucleotide occurs to form an attached oligonucleotide combination product which comprises an RNA portion (from a composite primer); (d) an enzyme such as RNaseH cleaves RNA of the composite primer(s) from the RNA/DNA duplex of the hybridized product; (e) the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide; (f) another pair of oligonucleotides hybridize to the target polynucleotide, optionally another round of polymerization occurs, and attachment and cleavage occurs; and (g) cleaved attached oligonucleotide combination product accumulates. This method is illustrated in Figure 3.

In this and other sections describing embodiments using oligonucleotide attachment, generally covalent attachment is exemplified. However, non-covalent methods are contemplated and included in this invention.

In one embodiment, the sequence of interest that can be copied and/or quantified by these methods is a single nucleotide base. Examples of a sequence of interest that is a single nucleotide base is a single nucleotide polymorphic sequence. In this embodiment, either the 3' most nucleotide of the first oligonucleotide or the 5' most nucleotide of the second oligonucleotide is capable of hybridizing to the sequence (single nucleotide base) of interest. Amplification of the sequence of interest is achieved as follows. If the 3' most nucleotide of the first oligonucleotide is designed to hybridize to the sequence of interest, hybridization of the 3' most nucleotide of the first oligonucleotide to a target nucleic acid is required for primer extension(if there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and 5' end of the second oligonucleotide)) or attachment, such as with DNA ligase, of the extended or unextended product of the first oligonucleotide. If the 5' most nucleotide of the second oligonucleotide is designed to hybridize to the sequence of interest, hybridization of the 5' most nucleotide of the second oligonucleotide to a target nucleic acid is required for attachment, such as with DNA ligase, of the extended or unextended product. In both cases, generation of attached oligonucleotide combination product reflects generation of multiple copies of the sequence of interest.

In another embodiment, the sequence of interest comprises more than 1 nucleotide base. A sequence of interest that comprises more than 1 nucleotide base comprises preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 13 nucleotides, even more preferably from about 2 to about 11 nucleotides, and most preferably from about 2 to about 10 nucleotides.

In embodiments wherein the sequence of interest comprises more than 1 nucleotide base,.at least one oligonucleotide of the oligonucleotide pair (first oligonucleotide and second oligonucleotide) comprises a sequence that does not permit substantial hybridization to a target polynucleotide in the absence of a

substantial portion of a sequence of interest in the target polynucleotide. Substantial hybridization of an oligonucleotide to a target polynucleotide can be at least preferably 50%, more preferably 60%, even more preferably 75%, and most preferably 90% of the efficiency of hybridization of the oligonucleotide to a target polynucleotide in the presence of a substantial portion of a sequence of interest in the target polynucleotide. Absence or presence of a substantial portion of a sequence of interest refers to absence or presence, respectively, of at least preferably 50%, more preferably 60%, even more preferably 75% and most preferably 90% of the sequence of interest in a target polynucleotide.

In one embodiment, the RNA portion(s) of the first and/or second oligonucleotide are designed to be substantially hybridizable to at least one nucleotide of the sequence of interest. Generally, a suitable oligonucleotide sequence is designed such that the oligonucleotide preferentially hybridizes to at least one nucleotide of a sequence of interest.

At least one of the oligonucleotides in the oligonucleotide pair (first oligonucleotide and second oligonucleotide) is a composite primer. Thus, in one embodiment, the first oligonucleotide is a composite primer. In another embodiment, the second oligonucleotide is a composite primer. In yet another embodiment, both oligonucleotides are composite primers.

The overall size of oligonucleotides, and the size of the cleavable portions, are selected to ensure that the attached oligonucleotide combination product is of a size that when the cleavable component is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment. Suitable sizes are determined according to the specific reaction conditions, and can be determined empirically by one skilled in the art. Generally, the cleaved product is substantially smaller than an uncleaved oligonucleotide. In one embodiment, the reaction is performed isothermally (at a single temperature or about the same temperature). In another embodiment, the temperature is a temperature that permits specific hybridization of the composite primer but is too high for significant hybridization of the cleaved primer extension product.

At least one of the oligonucleotides in an oligonucleotide pair (first oligonucleotide and second oligonucleotide) comprises a sequence hybridizable to the sequence of interest, or at least one nucleotide thereof. Thus, in one embodiment, the first oligonucleotide comprises a 3' most nucleotide which is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base. In another embodiment, the first oligonucleotide comprises a sequence hybridizable to the sequence of interest, or to at least one nucleotide of the sequence of interest. In another embodiment, the sequence hybridizable to at least one nucleotide of the sequence of interest further comprises a 3' most nucleotide of the first oligonucleotide. In yet another embodiment, the second oligonucleotide comprises a 5' most nucleotide which is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base. In yet another embodiment, the second oligonucleotide comprises a sequence hybridizable to the sequence of interest, or to at least one nucleotide of the sequence of interest. In another embodiment, the sequence hybridizable to at least one nucleotide of the sequence of interest further comprises a 5' most nucleotide of the second oligonucleotide. In still another embodiment, each oligonucleotide is hybridizable to a portion (or at least one polynucleotide) of the sequence of interest. In another embodiment, the oligonucleotides in combination further comprise the entire sequence of interest.

If there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and the 5' end of the second oligonucleotide, the gap is filled by oligonucleotide extension, which can be achieved using polymerization methods known in the art. Thus, for example, a DNA polymerase and the necessary dNTPs (corresponding to the gap sequence) would be provided under conditions that permit DNA polymerization.

Attachment of hybridized first oligonucleotide and second oligonucleotide can occur only if the 3' most nucleotide of the first oligonucleotide and the 5' most nucleotide of the second oligonucleotide are located adjacent to each other on a target polynucleotide. Thus, suitable agents or means for effecting attachment of the oligonucleotide pair are those that function only when this

requirement is met. An example of such an agent is DNA ligase. Other examples of agents and means for effecting covalent attachment include those described in U.S. Pat. No. 5,185,243.

At least one of the two oligonucleotides (first oligonucleotide and second oligonucleotide) is a composite primer. A first oligonucleotide that is a composite primer comprises an RNA portion and a DNA portion. A second oligonucleotide that is a composite primer comprises a DNA portion and an RNA portion. In one embodiment, both the first oligonucleotide and second oligonucleotide are composite primers. In another embodiment, only one of the two oligonucleotides is a composite primer.

The use of composite primers as described herein renders at least a portion of the attached oligonucleotide combination product cleavable when hybridized to a target polynucleotide. Thus, in a reaction mixture comprising a suitable cleaving agent, cleavage of the product results in accumulation of cleaved covalently-attached oligonucleotide combination product. Generally and preferably, this reaction occurs under essentially the same conditions as those for oligonucleotide attachment. Suitable cleaving agents are agents that are active in these conditions. Preferably, suitable cleaving agents are those that specifically cleave RNA in a RNA/DNA duplex, such as RNase H.

Without intending to be bound by theory, it is expected that dissociation of cleaved attached oligonucleotide combination product leading to accumulation of said product occurs due to the lower affinity (due to reduced melting temperature of hybrid formation) of the smaller (cleaved) attached oligonucleotide combination product. It is appreciated, however, that some uninduced dissociation of the uncleaved portion of the cleaved primer extension product from the target polynucleotide may result from binding of an uncleaved primer to the target polynucleotide.

In a preferred embodiment, the resultant cleaved attached oligonucleotide combination products in a reaction mixture are separated for analysis on a suitable matrix. Any of a number of methods can be used to effect the separation, as described in, for example, McIntosh et al., *supra*. Such methods include, but are

not limited to, oligonucleotide array hybridization, mass spectrometry, flow cytometry, HPLC, FPLC, size exclusion chromatography, affinity chromatography, and gel electrophoresis. The amount of the products is linearly related to the amount of the nucleic acid sequence of interest in the sample. Thus, comparison of the amount of cleaved attached oligonucleotide combination product obtained in a test sample with the amount of product obtained in a reference sample comprising a known amount of the sequence of interest provides quantification of the sequence of interest in the test sample. Methods of making such comparisons are known in the art.

It is appreciated that the oligonucleotide attachment-based methods of the invention are also suitable for generating copies of a plurality of sequences of interest in a sample. That is to say, a plurality of target sequences may be copied and/or quantified in a single reaction mixture. Different sequences of interest may be present on a single target polynucleotide or on different target polynucleotides. These methods employ two or more sets of first and second oligonucleotides, with each set of oligonucleotides being substantially hybridizable to each target nucleic acid sequence.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Methods of determining whether a nucleic sequence of interest is present or absent in a sample

Primer extension-based detection of nucleic acid sequence

The invention provides methods of determining whether a nucleic acid sequence of interest is present or absent in a sample, comprising reacting the following combination: (a) a single-stranded target polynucleotide suspected of comprising a nucleic acid sequence of interest; (b) a composite primer comprising an RNA portion and a 3' DNA portion; (c) a DNA polymerase; (d) labeled or unlabeled deoxyribonucleoside triphosphosphates, or suitable analogs that singly or in combination permit only limited primer extension; and (e) an enzyme, such

as RNase H, which cleaves RNA from an RNA/DNA duplex. The combination is subjected to suitable conditions such that (a) the composite primer hybridizes to the target polynucleotide; (b) limited primer extension occurs from the composite primer, to form a duplex; (c) RNase H cleaves RNA of the composite primer from the RNA/DNA duplex; (d) the cleaved primer extension product comprising a detectable identifying characteristic dissociates from the target polynucleotide; (e) another composite primer hybridizes to the target polynucleotide, and another round of primer extension occurs; and (f) cleaved primer extension product accumulates, whereby detection of a detectable identifying characteristic of the primer extension product (and thus, detection of the primer extension product comprising a detectable identifying characteristic) indicates the presence of the sequence of interest. Conversely, lack of detection of the cleaved primer extension product is indicates absence of the sequence of interest.

In one embodiment, the sequence of interest that is detected is a single nucleotide base. Examples of a sequence of interest that is a single nucleotide base is a single nucleotide polymorphic sequence, or a mutation such as a nucleotide substitution, deletion or insertion. In another embodiment, the sequence of interest that is detected comprises more than one nucleotide base, which can be a mutant sequence due to nucleotide substitution, deletion or insertion, or any combination thereof. In embodiments wherein the sequence of interest comprises more than one nucleotide base, the sequence of interest comprises preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 13 nucleotides, even more preferably from about 2 to about 11 nucleotides, and most preferably from about 2 to about 10 nucleotides. It is understood that a sequence of interest comprising more than one base can also be detected by detection of a single base (or less than the total number of bases of the sequence of interest), provided that the detection of a single base (or less than the total number of bases) indicates that the larger sequence of interest is present.

The composite primers are designed to hybridize to a sequence on a target polynucleotide that is in close proximity to the sequence of interest. Accordingly, methods of the invention comprise hybridizing a composite primer to a single

stranded target polynucleotide such that the 3' most nucleotide of the primer is from preferably about 1 nucleotide to about 10 nucleotides, more preferably about 1 nucleotide to about 8 nucleotides, even more preferably about 1 to about 6 nucleotides, and most preferably about 1 to about 4 nucleotides from the sequence of interest.

The hybridized primer is extended to include at least the sequence of interest. Generally, the amount of primer extension is limited to minimize the size of the primer extension product. The primer extension product should be of a size that when the RNA component is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension. Thus, a preferred size of primer extension product is one that when its RNA component is cleaved results in a product that binds to the target polynucleotide with less affinity, such that the frequency of dissociation from the target polynucleotide is increased compared to the uncleaved primer extension product. The frequency of dissociation is increased if there is an increase of preferably at least about 20%, more preferably at least about 40%, even more preferably at least about 60%, and most preferably at least about 80% increase of dissociation of cleaved product relative to uncleaved product. It is to be understood that the dissociation frequency is generally measured under essentially the same condition as those for primer extension. In one embodiment, the reaction is performed isothermally (at a single temperature or about the same temperature). In another embodiment, the temperature is a temperature that permits specific hybridization of the composite primer but is too high for significant hybridization of the cleaved primer extension product.

Limited primer extension can be achieved by any of a variety of methods known in the art. Generally and preferably, the extent of primer extension is limited by providing at least one terminator dNTP or by not providing at least one essential dNTP. Terminator dNTPs are known in the art, and include, for example, dideoxyribonucleotide triphosphates. An essential dNTP is a deoxyribonucleotide the presence of which in a reaction is required for unlimited

primer extension. Thus, a deoxyribonucleotide is essential within the context of a particular sequence identity of the target polynucleotide.

In some embodiments, limited primer extension comprises addition of only one nucleotide to the 3' end of the primer. In these embodiments, the sequence of interest is a single base nucleotide, which is located in the target polynucleotide immediately 5' and adjacent to the nucleotide of the target polynucleotide to which the 3'-most nucleotide of the primer is hybridized. Primer extension can be limited to only one nucleotide by providing in the reaction mixture terminator deoxyribonucleotide(s) that corresponds to the next nucleotide that can be added to the primer based on the identity of the sequence of interest. Thus, for example, if the sequence of interest is an adenine, providing only terminator dTTPs in the reaction mixture would result in a single nucleotide primer extension only if the sequence of interest is present. Thus, generation of a primer extension product that incorporates the terminator dTTP would be indicative of the presence of the sequence of interest. On the other hand, lack of expected primer extension product would indicate the absence of the sequence of interest.

In another embodiment, the sequence of interest is the deletion of a single nucleotide base which is located in the target polynucleotide immediately 5' and adjacent to the nucleotide of the target polynucleotide to which the 3'-most nucleotide of the primer is hybridized. Primer extension can be limited to only one nucleotide by providing in the reaction mixture terminator deoxyribonucleotide(s) that corresponds to the next nucleotide that can be added to the primer based on the identity of the nucleotide 3' of the sequence of interest (which is a deleted base). Thus, for example, if the nucleotide expected to be deleted is an adenine, and the nucleotide 3' of the sequence of interest is a cytosine, providing only terminator dGTPs in the reaction mixture would result in a single nucleotide primer extension only if the adenine is absent. Thus, generation of a primer extension product that incorporates the terminator dGTP would be indicative of the presence of the sequence of interest. On the other

hand, lack of expected primer extension product would indicate the absence of the sequence of interest.

In other embodiments, limited primer extension comprises addition of more than one nucleotide. In these embodiments, primer extension comprises addition of preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 12 nucleotides, even more preferably from about 2 to about 10 nucleotides, and most preferably from about 2 to about 8 nucleotides. In these embodiments, primer extension can be limited by not providing at least one essential dNTP, by providing at least one terminator nucleotide, or by both not providing at least one essential dNTP and providing at least one terminator nucleotide in a single reaction mixture. One of skill in the art would be able to determine the necessary combination based on the identity of target polynucleotide sequence adjacent to the hybridization site of the primer.

It is understood that the methods of the invention may be used to determine whether multiple sequences of interest are present or absent in a sample, as further described herein and in Examples 1 and 6-9. That is to say, multiple of sequences of interest may be detected simultaneously in a single reaction mixture. The different sequences of interest may be on the same or different target polynucleotides. In one aspect, multiple of sequences of interest are present on different target polynucleotides, for example, as when a plurality of sequences of interest are present in different genomic polynucleotides or different cDNA polynucleotides. In another aspect, multiple sequences of interest are present on multiple target polynucleotides, for example, different alleles, or a wildtype (native) and mutant (non-wildtype) sequences, or polymorphic sequences. In another aspect, multiple sequences of interest are present on a single target polynucleotide, for example, as when more than one locus on a single polynucleotide are analyzed, or for example, as when multiple interest are present on a single cDNA polynucleotide. Detection of accumulated primer extension products comprising multiple sequence of interest is described below.

A reaction mixture can comprise a variety of dNTP configurations, such as a single terminator dNTP, a combination of terminator dNTPs, at least one

terminator dNTP and at least one but not all of the four dNTPs (which can include non-terminator forms of a terminator dNTP that is included), or at least one but not all of the four dNTPs (without any terminator dNTPs). These dNTPs can be labeled or unlabeled. The choice of what form (e.g., labeled, unlabeled, type of label) and type of dNTPs is governed by factors that are evident to one skilled in the art. Among these factors are the sequence of interest, knowledge and identity of the sequence flanking the sequence of interest, and the choice of detectable identifying characteristic(s) (described below) that the practitioner wishes to incorporate in the primer extension product.

A DNA polymerase capable of effecting extension of the hybridized primer along the target polynucleotide is used to generate primer extension product. Suitable polymerases are known in the art, including, for example Bst DNA polymerase, Bca DNA polymerase, E. coli DNA polymerase I, the Klenow fragment of E. coli DNA polymerase I, Taq DNA polymerase, T7 DNA polymerase (Sequenase), or recombinant derivatives thereof, which are selected for improved utilization of modified dNTPs and their analog. DNA polymerases with or without strand displacement activity are suitable for the primer extension-based methods of the invention. Factors that determine suitability of a polymerase are known to one skilled in the art, and include the temperature used to effect primer hybridization. For example, the availability of thermostable DNA polymerases can be used where primer hybridization at elevated temperatures enables greater specificity of primer hybridization to the target polynucleotide.

The use of composite primers in primer extension renders at least a portion of the primer extension product cleavable when hybridized to a target polynucleotide. A suitable cleaving reagent is provided in the reaction mixture to effect cleavage of the primer extension product. Cleavage of the product permits accumulation of cleaved primer extension product. Generally and preferably, this reaction occurs under essentially the same conditions as those for primer extension. Thus, suitable cleaving agents are agents that are active in these conditions. In preferred embodiments, the cleavable portion is RNA, in which

case suitable cleaving agents are those that specifically cleave RNA in a RNA/DNA duplex, such as RNase H.

Preferred detectable identifying characteristics include size of the primer extension product, sequence of the non-primer portion of the primer extension product (which may or may not be considered in the context of the sequence of the primer portion of the primer extension product), detectable signal associated with the primer extension product. It is understood by one skilled in the art that while the preceding discussion addresses detection of accumulated cleaved primer extension products comprising a detectable identifying characteristic(s), absence of accumulated primer extension product comprising a detectable identifying characteristic(s) is also informative. For example, if the sequence of interest is a native (e.g., wildtype) sequence, primer extension product comprising a detectable identifying characteristic corresponding to the native sequence will not accumulate in the presence of a nonnative (e.g., mutant) sequence of interest. In one embodiment, absence of primer extension product specific to the wildtype sequence of interest indicates presence of a mutant sequence of interest. In another embodiment, absence of primer extension product indicates presence of an altered sequence of interest, as discussed below.

Appropriate detectable identifying characteristics to be incorporated in a primer extension product can be determined by one skilled in the art, in view of and based on the context of the components (such as type and/or form of the dNTP(s) provided, and/or type of label associated with the dNTPs provided and the primer), the identity of the sequence of interest, and the sequences that flank the sequence of interest in a target polynucleotide. It is appreciated that one or more detectable identifying characteristics may be used to characterize a primer extension product, and that characterization may be performed iteratively.

For example, characterization of primer extension product based on size of the product is generally employed when the dNTPs are unlabeled. In one illustrative scenario, primer extension is terminated due to the presence of a sequence of interest and a terminator dNTP corresponding to that sequence. Thus, in the absence of that sequence and any other dNTP other than the terminator

dNTP corresponding to that sequence in a sample, the product would be smaller than if the sequence is present. In another illustrative scenario where (i) the sequence of interest comprises a single nucleotide, (ii) the same nucleotide base is present in the intervening sequence of the target polynucleotide between the 3' end of the primer hybridization site and the sequence of interest, (iii) the only terminator dNTP type provided is that which corresponds to the sequence of interest, and (iv) a non-terminator form of the terminator dNTP is also provided, presence of the sequence of interest is indicated by detection of products of 3 distinct sizes: 1 size resulting from incorporation of terminator dNTP in the intervening site, 1 size resulting from incorporation of terminator dNTP at the single nucleotide site (sequence) of interest, and 1 size due to conditions that permit only limited extension resulting in termination at a site downstream from the sequence of interest.

In another example, characterization of primer extension products based on size may be used when determining whether a plurality of sequences of interest are present or absent in a sample (for example, multiplexing). Depending on the sequence of interest, the sequence of the template nucleic acid and the mixture of dNTPs and/or terminator dNTPs included in the mixture, limited primer extension can results in the creation of different sized products (or patterns of products as described above) for each sequence of interest.

In another example, characterization of primer extension product can be based on determining the sequence of the primer extension product. Each primer extension product generally has a unique sequence. Methods of determining sequence are well known in the art. Sequence-based detection methods include hybridization of the primer extension product to specific oligonucleotides, for example, immobilized on an array. This method is particularly useful when a plurality of primer extension products are produced.

In yet another example, characterization of primer extension product can be based on detection of a signal, or lack thereof, associated with incorporation of a labeled dNTP, or an analog thereof into the product. For example, in the scenarios illustrated above, a labeled dNTP corresponding to a sequence of

interest could be incorporated into a primer extension product. Detection of product with the signal generated by the label indicates the presence of that sequence.

In another illustrative scenario, (i) each sequence of interest comprises at least a single nucleotide, for example, target polynucleotides possessing two genotypes (alleles) of a defined sequence, or target polynucleotides possessing both a wildtype and a mutant allele of a defined sequence, and (ii) two target polynucleotides comprise sequence of interest that are variant sequences with respect to each other. Limited primer extension is conducted in the presence of two or more distinguishably labeled terminator dNTPs corresponding to the two sequences of interest. Detection of product with the signal generated by each label associated with each dNTP indicates the presence of each variant sequences of interest.

Labels suitable for use in the methods of this invention are known in the art, and include, for example, fluorescent dye labels and isotopic labels. Homogeneous detection of the primer extension product can also be employed. For example, the optical properties of a label associated with a dNTP can be altered subsequent to incorporation of the labeled dNTP into a primer extension product. Such a label includes fluorescent dyes that undergo fluorescence polarization between being attached to free dNTPs and being incorporated in a polynucleotide. See, e.g. U.S. patent No. 6,326,142, and references cited therein.

Another example of homogeneous detection is based on alteration of spectral properties of a label by means of energy transfer. When a primer is labeled by a donor or acceptor dye, for example, and the dNTPs, or their analogs, are labeled with acceptor or donor dyes, respectively, incorporation of the labeled dNTPs into the product (i.e., linkage to the primer) enables energy transfer between the donor-acceptor dyes, thus resulting in specific detectable spectral properties of the attached dyes. These dyes are known in the art, as described in, for example, U.S. Pat. No. 4,996,143 (e.g., fluorescein and Texas Red donor acceptor dye pair), and U.S. Pat. No. 5,688,648. Other label combinations are also possible. For example, two ligands (such as digoxigenin and biotin) each

attached to different parts of a primer extension product (generally, primer and non-primer portion) can be brought into close proximity in the context of a primer extension product. Binding of the two ligands with their corresponding antibodies which are differentially labeled can be detected due to the interaction of the labels. For instance, if the two different labels are a photosensitizer and a chemiluminescent acceptor dye, the interaction of the labels can be detected by the luminescent oxygen channeling assay as described in U.S. Pat. No. 5,340,716. Other interacting label pairs useful in the present invention are known in the art, see, e.g., 5,340,716; 3,999,345; 4,174,384; and 4,261,968 (Ullman et al.); and 5,565,322 (Heller et al.); 5,709,994 (Pease et al.); and 5,925,517 (Tyagi et al.). Examples of ligands 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 ligands produce little or no signal when in close proximity, and a greater signal when separated. In other embodiments, the ligands may generate a signal when in close proximity and generate less or no signal when separated. Examples of the latter ligands 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.

The detectable identifying characteristics described above can be detected by methods known in the art. Size of a polynucleotide (primer extension product) can be determined by, for example, gel electrophoresis sizing and mass spectrometry (see, for example, Monforte et al., U.S. Pat. Nos. 5,830,655 and 5,700,642). Methods of sequencing a polynucleotide (primer extension product) are well-known in the art. Methods of specific hybridization of a polynucleotide (primer extension product) are well known in the art, and include, for example, hybridization to polynucleotides immobilized on an array. As is known in the art, a microarray refers to an assembly of distinct polynucleotides or oligonucleotides immobilized at defined positions on a substrate (surface). Arrays are formed on substrates fabricated with materials such as paper, glass, plastic (e.g., polypropylene, nylon), polyacrylamide, nitrocellulose, silicon, optical fiber, polystyrene, or any other suitable solid or semi-solid support, and configured in a

planar (e.g., glass plates, silicon chips) or three-dimensional (e.g., pins, fibers, beads, particles, microtiter wells, capillaries) configuration. Polynucleotides or oligonucleotides forming arrays may be attached to the substrate by any number of ways including (i) *in situ* synthesis (e.g., high-density oligonucleotide arrays) using photolithographic techniques (see, Fodor et al., *Science* (1991), 251:767-773; Pease et al., *Proc. Natl. Acad. Sci. U.S.A.* (1994), 91:5022-5026; Lockhart et al., *Nature Biotechnology* (1996), 14:1675; U.S. Pat. Nos. 5,578,832; 5,556,752; and 5,510,270); (ii) spotting/printing at medium to low-density (e.g., cDNA probes) on glass, nylon or nitrocellulose (Schena et al, *Science* (1995), 270:467-470, DeRisi et al, *Nature Genetics* (1996), 14:457-460; Shalon et al., *Genome Res.* (1996), 6:639-645; and Schena et al., *Proc. Natl. Acad. Sci. U.S.A.* (1995), 93:10539-11286); (iii) by masking (Maskos and Southern, *Nuc. Acids. Res.* (1992), 20:1679-1684) and (iv) by dot-blotting on a nylon or nitrocellulose hybridization membrane (see, e.g., Sambrook et al., Eds., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., Vol. 1-3, Cold Spring Harbor Laboratory (Cold Spring Harbor, N.Y.)). Polynucleotides or oligonucleotides may also be noncovalently immobilized on the substrate by hybridization to anchors, by means of magnetic beads, or in a fluid phase such as in microtiter wells or capillaries. Arrays or microarrays of polynucleotides are generally nucleic acids such as DNA, RNA, PNA, and cDNA but may also include proteins, polypeptides, oligosaccharides, cells, tissues and any permutations thereof which can specifically bind the target molecules.

Methods of detecting detectable signals are known in the art. Signal detection may be visual or utilize a suitable instrument appropriate to the particular label used, such as a spectrometer, fluorimeter, or microscope. For example, where the label is a radioisotope, detection can be achieved using, for example, a scintillation counter, or photographic film as in autoradiography. Where a fluorescent label is used, detection may be by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence, such as by microscopy, visual inspection or photographic film. Where enzymatic labels are used, detection may be by providing appropriate substrates for the

enzyme and detecting the resulting reaction product. Simple colorimetric labels can usually be detected by visual observation of the color associated with the label; for example, conjugated colloidal gold is often pink to reddish, and beads appear the color of the bead.

In a preferred embodiment, the resultant cleaved primer extension products in a reaction mixture are separated for analysis on a suitable matrix. Any of a number of methods can be used to effect the separation, as described in, for example, McIntosh et al. (PCT Pub. No. WO98/59066). Such methods include, but are not limited to, oligonucleotide array hybridization, mass spectrometry, flow cytometry, HPLC, FPLC, size exclusion chromatography, affinity chromatography, and gel electrophoresis.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Primer extension-based methods of quantifying a sequence of interest in a sample, if present

It is apparent that the primer extension-based methods described herein may also be used to quantify a sequence of interest in a sample. The amount of product produced is linearly related to the amount of nucleic acid of interest in the sample. Thus, in some embodiments, comparison of amount of cleaved primer extension product obtained in a test sample with the amount of product obtained in a reference sample comprising a known amount of a sequence of interest provides quantification of the sequence of interest in the test sample. Methods of making such comparisons are known in the art. As is evident, as used herein, "quantification" refers to the determination of an absolute level of a sequence of interest (for example, amount of a sequence of interest in a sample as measured by number of copies or weight), as well as a relative levels of a sequence of interest in a sample. In one embodiment, amount of a sequence of interest is compared to amount of another sequence of interest. In another embodiment, amount of a sequence of interest is compared to amount of target polynucleotide. Thus, quantification of a sequence of interest also includes the determination of the

relative level of two or more sequences of interest, for example, of two polymorphic alleles, or a wildtype verses a polymorphic (mutant) allele. Comparison of the amount of cleaved primer extension product containing a first detectable identifying characteristic obtained in a test sample and the amount of cleaved primer extension product containing a second detectable identifying characteristic obtained in the same test sample permits quantification of the relative amounts of each sequence of interest. Methods of making such comparisons are known in the art and further described in Example 8. It is further appreciated that a reference label is desirably used, for example, to normalize signal intensity for incorporated labeled dNTPs (or ddNTPs) and to control for variation in experimental and/or detection conditions. A non-limiting example of a reference dye includes LIZ (ABI).

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Primer extension-based methods of identifying an altered sequence of interest and/or determining whether an altered sequence of interest is present in a sample

It is to be understood that the primer extension based methods of the invention are also useful for identifying altered sequences of interest, including, for example, a deletion, insertion, substitution, transversion, or any combination thereof, of one or more nucleotides of a sequence of interest. In general summary, a composite primer suitable for determining whether a sequence of interest is present or absent in a target polynucleotide is used in the primer extension-based methods of the invention with a target polynucleotide suspected of comprising an altered sequence of interest.

For example, in one illustrative scenario, a composite primer are designed to hybridize in close proximity to a sequence of interest. The composite primer binds to a target polynucleotide suspected of comprising an altered sequence of interest, in close proximity to the suspected altered sequence of interest. Primer extension is allowed to occur. Cleavage of the composite primer results in dissociation of the primer extension product from the target polynucleotide, thus

allowing accumulation of primer extension products. Accumulated cleaved primer extension products are characterized to determine the presence or absence of the sequence of interest in a sample. In one embodiment, absence of primer extension product comprising the sequence of interest indicates the presence of an altered sequence of interest (the sequence of which may or may not be previously unknown). The altered sequence of interest can be further characterized as described herein. In another embodiment, absence of primer extension product comprising the sequence of interest indicates the presence of a previously characterized altered sequence of interest.

In one illustrative scenario, a wildtype polynucleotide sequence is known, for example, in a gene of interest or in a portion of a gene known to possess a "hot spot" for mutation. The target polynucleotide can be screened for alteration in the sequence as follows: composite primers are designed that hybridize to the target polynucleotide, such that primer extension is occurs creating a primer extension product complementary to the portions of the target polynucleotide suspected of comprising an altered sequence of interest. Primer extension is allowed to occur, and cleaved primer extension product accumulates as described herein. Analysis (characterization) of cleaved primer extension product reveals the presence or absence of an altered sequence of interest.

It is also to be understood that the primer extension-based methods of the invention are useful for identifying whether an altered sequence of interest is present or absent in a sample, as described herein. Thus, the methods described herein are applicable to determining whether a sequence of interest or an altered sequence of interest is present in a sample. Presence of an altered sequence of interest may be indicated, for example, by absence of a sequence of interest. Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Oligonucleotide attachment-based detection of nucleic acid sequence

The invention provides methods for detecting whether a sequence of interest is present or absent in a target polynucleotide, comprising reacting the

following combination: (a) a single-stranded target polynucleotide suspected of comprising a nucleic acid sequence of interest; (b) a first oligonucleotide and a second oligonucleotide that are hybridizable to non-overlapping sequences of a target polynucleotide, wherein at least one of the oligonucleotides comprises an RNA portion, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target polynucleotide that is hybridizable to the second oligonucleotide, and wherein at least one oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest; (c) optionally a DNA polymerase; (d) an agent, such as DNA ligase, capable of effecting attachment, preferably covalent attachment, of the two oligonucleotides when the oligonucleotides are hybridized to the same target polynucleotide strand; (e) optionally dNTPs, or suitable analogs; and (f) an enzyme, such as RNase H, which cleaves RNA from an RNA/DNA duplex. The combination is subjected to suitable conditions such that (a) the oligonucleotides (at least one of which is a composite primer) hybridize to a target polynucleotide comprising the sequence of interest; (b) oligonucleotide extension from the first oligonucleotide by DNA polymerase occurs if there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and 5' end of the second oligonucleotide; (c) attachment, preferably covalent attachment, of hybridized first oligonucleotide (or oligonucleotide extension product thereof) and second oligonucleotide occurs to form an attached oligonucleotide combination product that comprises an RNA portion (from a composite primer); (d) an enzyme such as RNaseH cleaves RNA of the composite primer(s) from the RNA/DNA duplex of the hybridized product; (e) the cleaved attached oligonucleotide combination product comprising a detectable identifying characteristic dissociates from the target polynucleotide; (f) another pair of oligonucleotides hybridize to the target polynucleotide, optionally another round of polymerization occurs, and attachment and cleavage occurs; and (g) cleaved attached oligonucleotide combination product accumulates, whereby detection of a detectable identifying characteristic of the oligonucleotide attachment product indicates the presence of the sequence of interest. Conversely, lack of detection

of the attached oligonucleotide combination product is indicates absence of the sequence of interest.

In this and other sections describing embodiments using oligonucleotide attachment, generally covalent attachment is exemplified. However, non-covalent attachment methods are contemplated and included in this invention.

In one embodiment, the sequence of interest that can be detected by these methods is a single nucleotide base. Examples of a sequence of interest that is a single nucleotide base is a single nucleotide polymorphic sequence, or a mutation such as nucleotide substitution, deletion or insertion. In this embodiment, either the 3' most nucleotide of the first oligonucleotide or the 5' most nucleotide of the second oligonucleotide is capable of hybridizing to the sequence (single nucleotide base) of interest. The presence of the sequence of interest is indicated as follows. If the 3' most nucleotide of the first oligonucleotide is designed to hybridize to the sequence of interest, hybridization of the 3' most nucleotide of the first oligonucleotide to a target nucleic acid is required for primer extension (if there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and 5' end of the second oligonucleotide) or attachment, such as with DNA ligase, of the extended or unextended product. Therefore, production of attachment product indicates that the target nucleic acid contains the sequence defined by the 3' most nucleotide of the first oligonucleotide. Reduction or absence of product amplification, on the other hand, indicates that the target nucleic acid contains a sequence alteration that includes at least the base complementary to the 3' most nucleotide of the composite primer. If the 5' most nucleotide of the second oligonucleotide is designed to hybridize to the sequence of interest, hybridization of the 5' most nucleotide of the second oligonucleotide to a target nucleic acid is required for attachment, such as with DNA ligase, of the extended or unextended product. Therefore, production of cleaved attachment product indicates that the target nucleic acid contains the sequence defined by the 5' most nucleotide of the second oligonucleotide. Reduction or absence of product amplification, on the other hand, indicates that the target nucleic acid

contains a sequence alteration that includes at least the base complementary to the 5' most nucleotide of the second oligonucleotide.

In another embodiment, the sequence of interest comprises more than 1 nucleotide base. A sequence of interest can be a mutant sequence due to nucleotide substitution, deletion or insertion. The sequence of interest that can be detected by these methods comprises preferably from about 2 to about 15 nucleotides, more preferably from about 2 to about 13 nucleotides, even more preferably from about 2 to about 11 nucleotides, and most preferably from about 2 to about 10 nucleotides. It is understood that a sequence of interest comprising more than one base can also be detected by detection of a single base (or less than the total number of bases of the sequence of interest), provided that the detection of a single base (or less than the total number of bases) indicates that the larger sequence of interest is present.

In embodiments wherein the sequence of interest comprises more than 1 nucleotide base, at least one of the oligonucleotide pair (first oligonucleotide and second oligonucleotide) comprises a sequence that does not permit substantial hybridization to a target polynucleotide in the absence of a substantial portion of a sequence of interest in the target polynucleotide. Substantial hybridization of a oligonucleotide to a target polynucleotide can be at least preferably 50%, more preferably 60%, even more preferably 75%, and most preferably 90% of the efficiency of hybridization of the oligonucleotide to a target polynucleotide in the presence of a substantial portion of a sequence of interest in the target polynucleotide. Absence or presence of a substantial portion of a sequence of interest refers to absence or presence, respectively, of at least preferably 50%, more preferably 60%, even more preferably 75% and most preferably 90% of the sequence of interest in a target polynucleotide. Hybridization of the first oligonucleotide and the second oligonucleotide to a target polynucleotide, thus allowing attachment (after oligonucleotide extension to fill any gap as necessary) of the two bound oligonucleotides, would be indicative of the presence of a substantial portion of the sequence of interest. In one embodiment, the reaction is performed isothermally (at a single temperature or about the same temperature).

In another embodiment, the temperature is a temperature that permits specific hybridization of the composite primer but is too high for significant hybridization of the cleaved primer extension product. In one embodiment, the RNA portion(s) of the first and/or second oligonucleotide are designed to be substantially hybridizable to the sequence of interest. Generally, a suitable oligonucleotide sequence is designed such that the oligonucleotide preferentially hybridizes to a sequence of interest. In the presence of an altered sequence of interest (which may be a previously unidentified sequence of interest, or an allele, or wildtype sequence if the oligonucleotide is designed to hybridize to a mutant sequence), a mismatch is formed which results in reduced efficiency of cleavage of the RNA portion of the oligonucleotide primer, in which case formation of cleaved attachment product will be inhibited.

At least one of the oligonucleotides in the oligonucleotide pair (first oligonucleotide and second oligonucleotide) is a composite primer. Thus, in one embodiment, the first oligonucleotide is a composite primer. In another embodiment, the second oligonucleotide is a composite primer. In yet another embodiment, both oligonucleotides are composite primers.

The overall size of oligonucleotides, and the size of the cleavable portions, are selected to ensure that the attached oligonucleotide combination product is of a size that when the cleavable component is cleaved the product dissociates from the target polynucleotide under essentially the same conditions as those for attachment. Suitable sizes are determined according to the specific reaction conditions, and can be determined empirically by one skilled in the art. Generally, the cleaved product is substantially smaller than an uncleaved oligonucleotide.

At least one of the oligonucleotides in an oligonucleotide pair (first oligonucleotide and second oligonucleotide) comprises a sequence hybridizable to the sequence of interest, or to at least one nucleotide thereof. Thus, in one embodiment, the first oligonucleotide comprises a 3' most nucleotide which is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base. In another embodiment, the first oligonucleotide

comprises a sequence hybridizable to the sequence of interest, or to at least one nucleotide of the sequence of interest. In another embodiment, the sequence hybridizable to a portion of the sequence of interest further comprises a 3' most nucleotide of the first oligonucleotide. In yet another embodiment, the second oligonucleotide comprises a 5' most oligonucleotide which is hybridizable to the sequence of interest, and wherein the sequence of interest is a single nucleotide base. In yet another embodiment, the second oligonucleotide comprises a sequence hybridizable to the sequence of interest or at least one nucleotide of the sequence of interest. In another embodiment, the sequence hybridizable to a portion of the sequence of interest further comprises a 5' most nucleotide of the second oligonucleotide. In still another embodiment, each oligonucleotide is hybridizable to a portion (or to at least one nucleotide) of the sequence of interest. In another embodiment, the oligonucleotides in combination further comprise the entire sequence of interest.

Another embodiment of the invention provides methods of determining whether a plurality of target nucleic acid sequences are present or absent in a sample. These methods employ two or more sets of first and second oligonucleotides, with each set of oligonucleotides being specific to each target nucleic acid sequence of interest.

If there is a gap of at least one nucleotide between the 3' end of the first oligonucleotide and the 5' end of the second oligonucleotide, the gap is filled by oligonucleotide extension, which can be achieved using polymerization methods known in the art. Thus, for example, a DNA polymerase and the necessary dNTPs (corresponding to the gap sequence) would be provided under conditions that permit DNA polymerization. It is appreciated that the sequence added during extension can comprise a detectable identifying characteristic, including, for example, its size, sequence or incorporation of suitably labeled nucleotides.

Attachment of hybridized first oligonucleotide and second oligonucleotide can occur only if the 3' most nucleotide of the first oligonucleotide and the 5' most nucleotide of the second oligonucleotide are located adjacent to each other on a target polynucleotide. Thus, suitable agents or means for effecting

attachment of the oligonucleotide pair are those that function only when this requirement is met. An example of such an agent is DNA ligase. Other examples of agents or means for effecting covalent attachment include those described in U.S. Pat. No. 5,185,243.

Although the preceding discussion refers to the attachment as covalent, it is appreciated that methods of non-covalent attachment are suitable for use with the oligonucleotide attachment-based methods of the invention.

At least one of the two oligonucleotides (first oligonucleotide and second oligonucleotide) is a composite primer. A first oligonucleotide that is a composite primer comprises an RNA portion and a DNA portion. A second oligonucleotide that is a composite primer comprises a DNA portion and an RNA portion. In one embodiment, both the first oligonucleotide and second oligonucleotide are composite primers. In another embodiment, only one of the two oligonucleotides is a composite primer.

The use of composite primers as described herein renders at least a portion of the attached oligonucleotide combination product cleavable when hybridized to a target polynucleotide. Thus, in a reaction mixture comprising a suitable cleaving agent, cleavage of the product results in accumulation of cleaved attached oligonucleotide combination product. Generally and preferably, this reaction occurs under essentially the same conditions as those for oligonucleotide attachment. Suitable cleaving agents are agents that are active in these conditions. Preferably, suitable cleaving agents are those that specifically cleave RNA in a RNA/DNA duplex, such as RNase H.

The attached oligonucleotide combination products generated in the methods of this invention possess one or more detectable identifying characteristics. The formation of the product is characterized by the joining of the DNA sequences of the two oligonucleotides. Thus, detectable identifying characteristics of these products include, for example, the sequence of the product (since the sequences of the oligonucleotides are known), the size of the product (since the size of the oligonucleotides are known), and detectable signal on the product due to the presence of label(s) in the oligonucleotide(s), or presence of

label on polynucleotides incorporated during primer extension (if there is a gap of at least one nucleotide between the first and second primer).

For example, in the case of detectable signal as a detectable identifying characteristic, the combination of the two oligonucleotides in a single polynucleotide product can be detected by virtue of alteration of spectral properties of a label by means of energy transfer. In one illustration, one oligonucleotide could be labeled by a donor or acceptor dye, and the other oligonucleotide could be labeled with acceptor or donor dye, respectively. The attachment of the two oligonucleotides would bring the two labels into close proximity to enable energy transfer between the dyes, thus resulting in specific detectable spectral properties of the attached dyes. These dyes are known in the art, as described in, for example U.S. Pat. No. 4,996,143 (e.g., fluorescein and Texas Red donor acceptor dye pair), and U.S. Pat. No. 5,688,648. Other label combinations are also possible. For example, two ligands (such as digoxigenin and biotin) each attached to one of the oligonucleotides in the oligonucleotide pair can be brought into close proximity following attachment of the oligonucleotide pair. Binding of the two ligands with their corresponding antibodies which are differentially labeled can be detected due to the interaction of the labels. For instance, if the two different labels are a photosensitizer and a chemiluminescent acceptor dye, the interaction of the labels can be detected by the luminescent oxygen channeling assay as described in U.S. Pat. No. 5,340,716. Other interacting label pairs useful in the present invention are known in the art, see, e.g., U.S. Patents Nos. 5,340,716; 3,999,345; 4,174,384; and 4,261,968 (Ullman et al.); and 5,565,322 (Heller et al.); 5,709,994 (Pease et al.); and 5,925,517 (Tyagi et al.). Examples of ligands 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 ligands produce little or no signal when in close proximity, and a greater signal when separated. In other embodiments, the ligands may generate a signal when in close proximity and generate less or no signal when separated. Examples of the latter such ligands 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. Incorporation of label during primer extension is additionally detectable as described above.

In addition, the product can be detected by hybridizing it to a labeled oligonucleotide under conditions which are not suitable for hybridization of the either the first or the second oligonucleotide individually (i.e., when not ligated to each other). In another embodiment, one member of an interacting label pair is used in the attachment product and the other member is used in the labeled (third) olignucleotide. The third oligonucleotide may be immobilized on a solid surface and the hybridization of the labeled ligation product to the oligonucleotide results in an altered detectable signal due to the interaction of the labels on the ligation product and the third (immobilized) oligonucleotide. The immobilized oligonucleotide may be a part of an array of a plurality of oligonucleotides, each oligonucleotide being specific for a ligation product comprising a distinct sequence.

Detection of these detectable identifying characteristics can be achieved by a variety of methods evident to one skilled in the art. Methods of determining size and sequencing of polynucleotides (such as an attached oligonucleotide combination product) are known in the art. Methods of detecting detectable signals are known in the art, and are described above.

In a preferred embodiment, the resultant cleaved attached oligonucleotide combination products in a reaction mixture are separated for analysis on a suitable matrix. Any of a number of methods can be used to effect the separation, as described in, for example, McIntosh et al., *supra*. Such methods include, but are not limited to, oligonucleotide array hybridization, mass spectrometry, flow cytometry, HPLC, FPLC, size exclusion chromatography, affinity chromatography, and gel electrophoresis.

It is appreciated that while the preceding discussion describes detection of attachment product comprising a detectable identifying characteristic(s), the absence of accumulation of a attachment product comprising a detectable identifying characteristic(s) is also informative. For example, if the sequence of

interest is a native (wildtype sequence), attachment product comprising a detectable identifying characteristic corresponding to the native sequence will not accumulate in the presence of a nonnative (mutant) sequence of interest. Absence of attachment product can be a detectable identifying characteristic for a mutant sequence of interest. Absence of attachment product comprising a detectable identifying characteristic can also be indicative of the presence of an altered sequence of interest, as discussed below.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Oligonucleotide attachment-based methods of identifying an altered sequence of interest and/or determining whether an altered sequence of interest is present in a sample

It is apparent that the oligonucleotide attachment-based methods of the invention are also useful for identifying altered sequences of interest, including, for example, a deletion, insertion, substitution, transversion, or any combination thereof, of one or more nucleotides of a sequence of interest. In general summary, first and/or second oligonucleotides (or sets of first and second oligonucleotides) suitable for determining whether a sequence of interest is present in a target polynucleotide is used in the oligonucleotide attachment-based methods of the invention with a target polynucleotide suspected of comprising an altered sequence of interest. For example, in one illustrative scenario, first and second oligonucleotides are designed, wherein at least one oligonucleotide hybridizes to the sequence of interest. The first and second oligonucleotides bind to a target polynucleotide suspected of comprising an altered sequence of interest. Optional primer extension is allowed to occur (if a gap of one or more nucleotide exists between the first and second oligonucleotides). Attachment is effected by an agent capable of attachment. Cleavage of the covalently attached oligonucleotide product results in dissociation of the product from the target polynucleotide, thus allowing accumulation of primer extension products. Presence or absence of accumulated cleaved oligonucleotide attachment product indicates presence or

absence of the sequence of interest in a sample. In one embodiment, absence of cleaved attachment product indicates the presence of an altered sequence of interest (the sequence of which was previously unknown). In another embodiment, absence of cleaved attachment product indicates the presence of a previously identified altered sequence of interest. In another embodiment, primer extension adds sequence comprising an altered sequence of interest (or portion of an altered sequence of interest). Detection of an altered sequence of interest added by primer extension is described herein.

It is also understood that the oligonucleotide attachment-based methods of the invention are useful for identifying whether an altered sequence of interest is present or absent in a sample, as described herein. Thus, the methods described here are applicable to determining whether a sequence or interest or an altered sequence of interest is present in a sample. Presence of an altered sequence of interest may be indicated, for example, by absence of a sequence of interest.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Components and reaction conditions used in the methods of the invention

Target polynucleotide

A target polynucleotide includes nucleic acids from any source in purified or unpurified form, which can be DNA (dsDNA and ssDNA), mitochondrial DNA, chloroplast DNA, DNA-RNA hybrids, or mixtures thereof, genes, chromosomes, plasmids, the genomes of biological material such as microorganisms, e.g., bacteria, yeasts, viruses, viroids, molds, fungi, plants, animals, humans, and fragments thereof. Obtaining and purifying nucleic acids use standard techniques in the art. If a sample contains RNA which comprises a sequence of interest, cDNA can be generated from it by cDNA synthesis, as known in the art. The use of a DNA-RNA hybrid would require denaturation of the hybrid to obtain a ssDNA, or denaturation followed by reverse transcription to obtain a cDNA. The target nucleic acid can be only a minor fraction of a complex

mixture such as a biological sample and can be obtained from, for example, various biological material by procedures well known in the art.

The initial step of the methods of the invention is rendering the target polynucleotide single stranded. If the target polynucleotide is a double stranded (ds) DNA, the initial step is target denaturation. The denaturation step may be thermal denaturation or any other method known in the art, such as alkali treatment. The target polynucleotide may be a copy of another polynucleotide, for example, copy(ies) produced using nucleic acid amplification techniques which are known in the art. 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. Example 9 shows the use of a primer extension-based method of the invention in conjunction with single primer isothermal linear amplification of target polynucleotide sequences.

The target nucleic acid may be free in solution, or, in other embodiments, immobilized on a surface, e.g., as part of an array as discussed herein. Target nucleic acid sequence(s) may be immobilized on a surface (substrate) fabricated from a material such as paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, polystyrene, silicon, and optical fiber. Alternatively, the target nucleic acid sequence may be immobilized on the surface (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 another aspect of the invention, the target nucleic acid sequence is associated with an analyte. In one embodiment, the target nucleic acid sequence may be attached to an analyte, e.g., an antibody, polypeptide, or chemical compound, the presence, location, or quantity of which is desired to be known. It is understood that an analyte may be a member(s) of a binding pair. Non-limiting examples of a binding pair include a protein:protein binding pair, and a protein: antibody binding pair. In another embodiment, a target polynucleotides are attached to (tag) a molecular library of analytes, e.g., a molecular library of chemical compounds, a phage peptide display library, or a library of antibodies. Methods of attaching a target polynucleotide (tagging or addressing) to analytes are known in the art. See, e.g. U.S. Patent Nos. 6,309,843; 6,306,365; 6,280,935; 6,087,103 (and methods discussed therein).

Composite primer

In preferred embodiments of the methods of the invention, composite primers composed of RNA and DNA portions are used. The composite design of the primer is critical for subsequent dissociation of the primer extension product or attached oligonucleotide combination product from a target polynucleotide. An essential feature of the composite primers used in the methods of the invention is that cleavage of the cleavable portion of the primer in a primer extension or attached oligonucleotide combination product results in a marked decrease in the size of the product compared to the size of intact primers. A marked decrease in the product compared to primer size is defined as a cleaved product that is preferably at least about 2 nucleotides, more preferably at least about 4 nucleotides, even more preferably at least about 6 nucleotides, and most preferably at least about 8 nucleotides shorter than an uncleaved composite primer.

Composite primers for use in the methods and compositions of the present invention comprise at least one RNA portion that is (a) capable of being cleaved

with a ribonuclease when hybridized to the target DNA; and (b) preferably capable of binding (hybridizing) to a sequence on the target polynucleotide independent of hybridization of the DNA portion(s) to a sequence on the target polynucleotide. The composite primers bind to the target polynucleotide to form a partial heteroduplex in which only the RNA portion of the primer is cleaved upon contact with a ribonuclease such as RNase H, while the target strand remains intact, thus enabling annealing of another composite primer.

To achieve hybridization (which, as is well known and understood in the art, depends on other factors such as, for example, ionic strength and temperature), composite primers for use in the methods and compositions of the present invention are preferably of at least about 60%, more preferably at least about 75%, even more preferably at least about 90%, and most preferably at least about 95% complementarity to the target nucleic acid. The individual DNA and RNA portions of the composite primers are preferably of at least about 60%, more preferably at least about 75%, even more preferably at least about 90%, and most preferably at least about 95% complementarity to the target polynucleotide.

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 primer 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. For a given set of reaction conditions, 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.

One factor in designing and constructing primers 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, primer sequences for which the required free energy changes for formation of various complexes will be met.

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 primer-target and primer-primer duplexes, minor groove binders, O-methylation or other modification of nucleotides, temperature, and salt are potentially important factors in constructing primers with the requisite differences in binding energies.

As described herein, one or more composite primers may be used in a reaction.

Composite primers for use in primer extension-based methods and as first oligonucleotides in oligonucleotide attachment-based methods

The following is a description of some of the embodiments of composite primers suitable for use in primer extension-based methods and as first oligonucleotides in attachment-based methods of the invention.

The composite primers comprise a 3' DNA portion that is preferably capable of hybridization to a sequence on the target polynucleotide such that its hybridization to the target polynucleotide is favored over that of the product that dissociates from the target polynucleotide. Such primers can be rationally

designed based on well known factors that influence nucleic acid binding affinity, such as sequence length and/or identity, as well as hybridization conditions. In a preferred embodiment, hybridization of the 3' DNA portion of the composite primer to its complementary sequence in the target polynucleotide is favored over the hybridization of the homologous sequence in the 5' end of the dissociated product to the target polynucleotide.

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

The length of an RNA portion in a composite primer comprising a 3'-DNA portion and an RNA portion can be preferably from about 1 to about 30, more preferably from about 3 to about 20, even more preferably from about 4 to about 15, and most preferably from about 5 to about 10 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and an RNA

portion, an RNA portion can be at least about any of 1, 3, 4, 5 nucleotides, with an upper limit of about any of 10, 15, 20, 30, 35 nucleotides.

The length of the 5'-RNA portion in a composite primer comprising a 5'-RNA portion and a 3'-DNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most preferably from about 10 to about 15 nucleotides. In other embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, the 5'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30, 35 nucleotides.

In embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion further comprising non-5'-RNA portion(s), a non-5'-RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about 2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In certain embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion further comprising non-5'-RNA portion(s), a non-5'-RNA portion can be at least about any of 1, 2, 3, 5, with an upper limit of about any of 5, 6, 7, 10 nucleotides.

In embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, in which the 5'-RNA portion is adjacent to the 3'-DNA portion, the length of the 5'-RNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most preferably from about 10 to about 15 nucleotides. In certain embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, in which the 5'-RNA portion is adjacent to the 3'-DNA portion, the 5'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30 or 35 nucleotides.

The length of an intervening RNA portion in a composite primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about

2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In some embodiments of a composite primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion, an intervening RNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 5, 6, 7, 10 nucleotides. The length of an intervening RNA portion in a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about 2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion, an intervening RNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 5, 6, 7, 10 nucleotides. In a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion, further comprising a 5'-RNA portion, the 5'-RNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most preferably from about 10 to about 15 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion, further comprising a 5'-RNA portion, the 5'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30, 35 nucleotides.

The length of the 3'-DNA portion in a composite primer comprising a 3'-DNA portion and an RNA portion can be preferably from about 1 to about 30, more preferably from about 3 to about 18, even more preferably from about 5 to about 15, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and an RNA portion, the 3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 30, 35 nucleotides.

The length of the 3'-DNA portion in a composite primer comprising a 5'-RNA portion and a 3'-DNA portion can be preferably from about 1 to about 30 nucleotides, more preferably from about 3 to about 18 nucleotides, even more

preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, the 3' DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 30, 35 nucleotides.

In embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, further comprising non-3'-DNA portion(s), a non-3'-DNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion, further comprising non-3'-DNA portion(s), a non-3'-DNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides.

In embodiments of a composite primer comprising a 5'-RNA portion and a 3'-DNA portion in which the 5'-RNA portion is adjacent to the 3'-DNA portion, the length of the 3'-DNA portion can be preferably from about 1 to about 30 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In certain embodiments of the primer comprising a 5'-RNA portion and a 3'-DNA portion in which the 5'-RNA portion is adjacent to the 3'-DNA portion, the 3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 30, 35 nucleotides.

The length of a non-3'-DNA portion in a composite primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion, a non-3'-DNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides.

The length of the 3'-DNA portion in a composite primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion can be preferably

from about 1 to about 30 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising 5'- and 3'-DNA portions with at least one intervening RNA portion, the 3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 30, 35 nucleotides.

The length of a non-3'-DNA portion (i.e., any DNA portion other than the 3'-DNA portion) in a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion, a non-3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides. The length of the 3'-DNA portion in a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion can be preferably from about 1 to about 30 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 3'-DNA portion and at least one intervening RNA portion, the 3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 30, 35 nucleotides. It is understood that the lengths for the various portions can be greater or less, as appropriate under the reaction conditions of the methods of this invention.

In some embodiments, the 5'-DNA portion of a composite primer includes the 5'-most nucleotide of the primer. In some embodiments, the 5'-RNA portion of a composite primer includes the 5' most nucleotide of the primer. In other embodiments, the 3'-DNA portion of a composite primer includes the 3' most nucleotide of the primer. In other embodiments, the 3'-DNA portion is adjacent

to the 5'-RNA portion and includes the 3' most nucleotide of the primer (and the 5'-RNA portion includes the 5' most nucleotide of the primer).

The total length of the composite primer can be preferably from about 10 to about 50 nucleotides, more preferably from about 15 to about 40 nucleotides, and most preferably from about 20 to about 25 nucleotides. In some embodiments, the length can be at least about any of 10, 15, 20, 25 nucleotides, with an upper limit of about any of 25, 40, 50, 60 nucleotides. It is understood that the length can be greater or less, as appropriate under the reaction conditions of the methods of this invention.

Composite primers for use as second oligonucleotide in attachment-based methods

The following is a description of some of the embodiments of composite primers suitable for use as second oligonucleotides in attachment-based methods of the invention.

The composite primers comprise a 5' DNA portion that is preferably capable of hybridization to a sequence on the target polynucleotide such that its hybridization to the target polynucleotide is favored over that of the product that dissociates from the target polynucleotide. Such primers can be rationally designed based on well known factors that influence nucleic acid binding affinity, such as sequence length and/or identity, as well as hybridization conditions. In a preferred embodiment, hybridization of the 5' DNA portion of the composite primer to its complementary sequence in the target polynucleotide is favored over the hybridization of the homologous sequence in the 3'-end of the dissociated product to the target polynucleotide.

The composite primer comprises a combination of RNA and DNA (see definition above), with the 5'-end nucleotide being a nucleotide suitable for attachment to another nucleotide. The 5'-end nucleotide can be any nucleotide or analog that when present in a primer, is capable of being covalently attached to another nucleotide. Generally, the 5'-end nucleotide has a phosphate group. Suitable primers include those that comprise at least one portion of RNA and at

least one portion of DNA. Composite primers can comprise a 3'-RNA portion and a 5'-DNA portion (in which the RNA portion is adjacent to the 5'-DNA portion); or 3'- and 5'-DNA portions with an intervening RNA portion. Accordingly, in one embodiment, the composite primer comprises a 3' RNA portion and a 5'-DNA portion, preferably wherein the RNA portion is adjacent to the 5'-DNA portion. In another embodiment, the composite primer comprises 3'- and 5'-DNA portions with at least one intervening RNA portion (i.e., an RNA portion between the two DNA portions). In yet another embodiment, the composite primer of the present invention comprises a 5'-DNA portion and at least one intervening RNA portion (i.e., an RNA portion between DNA portions).

The length of an RNA portion in a composite primer comprising a 5'-DNA portion and an RNA portion can be preferably from about 1 to about 30, more preferably from about 3 to about 20, even more preferably from about 4 to about 15, and most preferably from about 5 to about 10 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and an RNA portion, an RNA portion can be at least about any of 1, 3, 4, 5 nucleotides, with an upper limit of about any of 10, 15, 20, 30, 35 nucleotides.

The length of the 3'-RNA portion in a composite primer comprising a 3'-RNA portion and a 5'-DNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most preferably from about 10 to about 15 nucleotides. In other embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, the 3'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30, 35 nucleotides.

In embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion further comprising non-3'-RNA portion(s), a non-3'-RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about 2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In certain embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion further comprising non-3'-RNA portion(s), a non-

3'-RNA portion can be at least about any of 1, 2, 3, 5, with an upper limit of about any of 5, 6, 7, 10 nucleotides.

In embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, in which the 3'-RNA portion is adjacent to the 5'-DNA portion, the length of the 3'-RNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most preferably from about 10 to about 15 nucleotides. In certain embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, in which the 3'-RNA portion is adjacent to the 5'-DNA portion, the 3'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30 or 35 nucleotides.

The length of an intervening RNA portion in a composite primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about 2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In some embodiments of a composite primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion, an intervening RNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 5, 6, 7, 10 nucleotides. The length of an intervening RNA portion in a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion can be preferably from about 1 to about 7 nucleotides, more preferably from about 2 to about 6 nucleotides, and most preferably from about 3 to about 5 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion, an intervening RNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 5, 6, 7, 10 nucleotides. In a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion, further comprising a 3'-RNA portion, the 3'-RNA portion can be preferably from about 3 to about 30 nucleotides, more preferably from about 5 to about 20 nucleotides, even more preferably from about 7 to about 18 nucleotides, preferably from about 8 to about 17 nucleotides, and most

preferably from about 10 to about 15 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion, further comprising a 3'-RNA portion, the 3'-RNA portion can be at least about any of 3, 5, 7, 8, 10 nucleotides, with an upper limit of about any of 15, 17, 18, 20, 30, 35 nucleotides.

The length of the 5'-DNA portion in a composite primer comprising a 5'-DNA portion and an RNA portion 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, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and an RNA portion, the 5'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 20, 22 nucleotides.

The length of the 5'-DNA portion in a composite primer comprising a 3'-RNA portion and a 5'-DNA portion can be preferably from about 1 to about 20 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, the 5' DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 20, 22 nucleotides.

In embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, further comprising non-5'-DNA portion(s), a non-5'-DNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion, further comprising non-5'-DNA portion(s), a non-5'-DNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides.

In embodiments of a composite primer comprising a 3'-RNA portion and a 5'-DNA portion in which the 3'-RNA portion is adjacent to the 5'-DNA portion, the length of the 5'-DNA portion can be preferably from about 1 to about 20

nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In certain embodiments of the primer comprising a 3'-RNA portion and a 5'-DNA portion in which the 3'-RNA portion is adjacent to the 5'-DNA portion, the 3'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 20, 22 nucleotides.

The length of a non-5'-DNA portion in a composite primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion, a non-5'-DNA portion can be at least about any of 1, 2, 3, 5 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides.

The length of the 5'-DNA portion in a composite primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion can be preferably from about 1 to about 20 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising 3'- and 5'-DNA portions with at least one intervening RNA portion, the 5'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 20, 22 nucleotides.

The length of a non-5'-DNA portion (i.e., any DNA portion other than the 5'-DNA portion) in a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion can be preferably from about 1 to about 10 nucleotides, more preferably from about 2 to about 8 nucleotides, and most preferably from about 3 to about 6 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion, a non-5'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 6, 8, 10, 12 nucleotides. The length of the 5'-DNA portion in a composite primer comprising a 5'-DNA portion

and at least one intervening RNA portion can be preferably from about 1 to about 20 nucleotides, more preferably from about 3 to about 18 nucleotides, even more preferably from about 5 to about 15 nucleotides, and most preferably from about 7 to about 12 nucleotides. In some embodiments of a composite primer comprising a 5'-DNA portion and at least one intervening RNA portion, the 5'-DNA portion can be at least about any of 1, 3, 5, 7, 10 nucleotides, with an upper limit of about any of 10, 12, 15, 18, 20, 22 nucleotides. It is understood that the lengths for the various portions can be greater or less, as appropriate under the reaction conditions of the methods of this invention.

In some embodiments, the 3'-DNA portion of a composite primer includes the 3'-most nucleotide of the primer. In some embodiments, the 3'-RNA portion of a composite primer includes the 3' most nucleotide of the primer. In other embodiments, the 5'-DNA portion of a composite primer includes the 5' most nucleotide of the primer. In other embodiments, the 5'-DNA portion is adjacent to the 3'-RNA portion and includes the 5' most nucleotide of the primer (and the 3'-RNA portion includes the 3' most nucleotide of the primer).

The total length of the composite primer can be preferably from about 10 to about 50 nucleotides, more preferably from about 15 to about 40 nucleotides, and most preferably from about 20 to about 25 nucleotides. In some embodiments, the length can be at least about any of 10, 15, 20, 25 nucleotides, with an upper limit of about any of 25, 40, 50, 60 nucleotides. It is understood that the length can be greater or less, as appropriate under the reaction conditions of the methods of this invention.

Other oligomers for use in the methods of the invention

The following is a description of oligomers and primers suitable for use as a first or second oligonucleotide (in combination with a composite oligomer as described herein) in the oligonucleotide-attachment based methods of the invention. Characteristics of suitable oligomers are evident to one skilled in the art. A primer suitable for use as the first oligonucleotide has a 3'-end nucleotide suitable for nucleic acid extension and/or attachment to another nucleotide. The

3'-end nucleotide can be any nucleotide or analog that when present in a primer, is extendable by DNA polymerase or capable of being covalently attached to another nucleotide. Generally, the 3'-end nucleotide has a 3'-OH. A primer suitable for use as the second oligonucleotide has a 5'-end nucleotide being a nucleotide suitable for attachment to another nucleotide, preferably covalent attachment. The 5'-end nucleotide can be any nucleotide or analog that when present in a primer, is capable of being covalently attached to another nucleotide. Generally, the 5'-end nucleotide has a phosphate group. In another embodiment, the 3'-end nucleotide of the second oligonucleotide is not capable of being extended by a polymerase. Such nucleotides are known in the art and include, for example, a nucleotide or analog thereof lacking a -OH group.

Labeled or unlabeled primers (oligonucleotides) 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), 111:6966-6976. A review of synthesis methods is provided in, for example, Goodchild, Bioconjugate Chemistry (1990), 1(3):165-187.

DNA polymerase, ribonuclease and agent that cleaves RNA from a RNA/DNA hybrid

DNA polymerases for use in the methods and compositions of the present invention are capable of effecting extension of the composite primer and oligonucleotide according to the methods of the present invention. Accordingly, a preferred polymerase is one that is capable of extending a nucleic acid primer along a nucleic acid template that is comprised at least predominantly of deoxynucleotides. Preferably, the DNA polymerase has high affinity for binding at the 3'-end of an oligonucleotide hybridized to a nucleic acid strand. Preferably, the DNA polymerase does not possess substantial nicking activity. Preferably, the polymerase has little or no 5'->3' exonuclease activity so as to minimize degradation of primer, or primer extension polynucleotides. Generally, this exonuclease activity is dependent on factors such as pH, salt concentration, whether the template is double stranded or single stranded, and so forth, all of which are familiar to one skilled in the art. Mutant DNA polymerases in which the 5'->3' exonuclease activity has been deleted, are known in the art and are suitable for the amplification methods described herein. Suitable DNA polymerases for use in the methods and compositions of the present invention include those disclosed in U.S. Pat. Nos. 5648211 and 5744312, which include exo$^-$ Vent (New England Biolabs), exo$^-$ Deep Vent (New England Biolabs), Bst (BioRad), exo$^-$ Pfu (Stratagene), Bca (Panvera), sequencing grade Taq (Promega), thermostable DNA polymerases from thermoanaerobacter thermohydrosulfuricus, or recombinant derivatives (or mutants) thereof, which are selected for improved utilization of modified dNTPs and their analogs, for example, AcycloPol-DNA polymerase. DNA polymerases with or without strand displacement activity are suitable for the primer extension-based methods of the invention. Preferably, the DNA polymerase has little to no proofreading activity.

The agent for cleaving RNA from a RNA/DNA hybrid may be an enzyme, for example, a ribonuclease. Preferably, the ribonuclease cleaves ribonucleotides regardless of the identity and type of nucleotides adjacent to the ribonucleotide to be cleaved. It is preferred that the ribonuclease cleaves independent of sequence

identity. Examples of suitable ribonucleases for the methods and compositions of the present invention are well known in the art, including RNase H.

The ligase for use in the methods and compositions of the present invention is capable of effecting attachment of two polynucleotides when the polynucleotides are hybridized immediately adjacent to each other on a target polynucleotide, preferably, capable of effecting attachment of two pieces of DNA when the DNA pieces are hybridized immediately adjacent to each other on a target polynucleotide. A variety of ligases can be used including T4 DNA ligase, E. coli DNA ligase, and thermostable DNA ligase.

In general, the enzymes used in the methods and compositions of the present invention should not produce substantial degradation of the nucleic acid components of said methods and compositions.

Single-stranded nucleic acid or DNA binding protein ("SSB") can be used to enhance the efficiency of the hybridization and the denaturation of the primer and the target nucleic acid molecule. Examples of SSBs suitable for use in the present invention include E. coli SSB ("EcoSSB"), T4 gene 32 protein, T7 SSB, Coliphage N4 SSB, calf thymus unwinding protein and adenovirus DNA binding protein. SSBs may decrease or remove secondary structure in ssDNA. EcoSSB is stable up to 100°C, and appears to be less sensitive to salt concentrations than SSB32. EcoSSB also has a lower tendency to aggregate than SSB32. Generally, EcoSSB, SSB32 and phage T7 SSB may improve hybridization of polynucleotides with complementary nucleic acid sequences. SSB32 may be useful for improving the specificity of hybridization, and can be used with mispairing for the identification of point mutations. EcoSSB also may reduce artifacts during deletion mutagenesis when used with Taq DNA polymerase.

Reaction conditions and detection

Appropriate reaction media and conditions for carrying out the methods of the present invention are those that permit nucleic acid copying and/or generation of products comprising unique identifying characteristics according to the methods of the present invention. Such media and conditions are known to

persons of skill in the art, and are described in various publications, such as U.S. Pat. Nos. 5,679,512, 6,107,061, 5,185,243, 6,004,744, 5,679,524, 6,013,431 & 5,888,819, and PCT Pub. Nos. WO99/42618, WO95/21271. For example, a buffer may be Tris buffer, although other buffers can also be used as long as the buffer components are non-inhibitory to enzyme components of the methods of the invention. The pH is preferably from about 5 to about 11, more preferably from about 6 to about 10, even more preferably from about 7 to about 9, and most preferably from about 7.5 to about 8.5. The reaction medium can also include bivalent metal ions such as Mg^{2+} or Mn^{2+}, at a final concentration of free ions that is within the range of from about 0.01 to about 10 mM, and most preferably from about 1 to 5 mM. The reaction medium can also include other salts, such as KCl, that contribute to the total ionic strength of the medium. For example, the range of a salt such as KCl is preferably from about 0 to about 100 mM, more preferably from about 0 to about 75 mM, and most preferably from about 0 to about 50 mM. The reaction medium can further include additives that could affect performance of the polymerization and/or ligation reactions, and/or generation of products with unique identifying characteristics, but that are not integral to the activity of the enzyme components of the methods. Such additives include proteins such as BSA, and non-ionic detergents such as NP40 or Triton. Reagents, such as DTT, that are capable of maintaining enzyme activities can also be included. Such reagents are known in the art. Where appropriate, an RNase inhibitor (such as Rnasine) that does not inhibit the activity of the RNase employed in the method can also be included. Any aspect of the methods of the present invention can occur at the same or varying temperatures. Preferably, the reactions are performed isothermally, which avoids the cumbersome thermocycling process. The reaction is carried out at a temperature that permits hybridization of the primers of the present invention to the target polynucleotide and that does not substantially inhibit the activity of the enzymes employed. The temperature can be in the range of preferably about 25°C to about 85°C, more preferably about 30°C to about 75°C, and most preferably about 37°C to about 70°C.

Nucleotide and/or nucleotide analogs, such as deoxyribonucleoside triphosphates, that can be employed for synthesis of the primer (or oligonucleotide) extension products in the methods of the invention are provided in the amount of from preferably about 50 to about 2500 μM, more preferably about 50 to about 2000 μM, even more preferably about 50 to about 1700 μM, and most preferably about 50 to about 1000 μM. In some embodiments, terminator and non-terminator forms of the same dNTP type are included in the same reaction.

The primer or oligonucleotide components of the reactions of the invention are generally in excess of the amount of nucleic acid sequence of interest that is to be copied, quantified and/or detected. They can be provided at about or at least about any of the following: 10, 10^2, 10^4, 10^6, 10^8, 10^{10}, 10^{12} times the amount of target polynucleotide. Composite primers and oligonucleotides can be provided at about or at least about any of the following concentrations: 50 nM, 100 nM, 500 nM, 1000 nM, 2500 nM, 5000 nM.

In one embodiment, the foregoing components are added simultaneously at the initiation of the copying, quantification and/or detection process. In another embodiment, components are added in any order prior to or after appropriate timepoints during the process, as required and/or permitted by the reaction. Such timepoints, some of which are noted below, can be readily identified by a person of skill in the art. For example, in some cases the target oligonucleotide may be denatured first prior to the addition of the one or more primers of the invention, the primer or primers of the invention may be added to the sample following denaturation and the enzymes may be added with or following the addition of the primer or primers of the invention. In addition, or alternatively, the polynucleotides may be added following the addition of the enzymes. The enzymes used in the methods of the present invention can be added to the reaction mixture either prior to the nucleic acid denaturation step, following the denaturation step, if any, or following hybridization of the primer or oligonucleotide to the target polynucleotide, as determined by their thermal stability and/or other considerations known to the person of skill in the art.

The reactions can be stopped at various timepoints, and resumed at a later time. Said timepoints can be readily identified by a person of skill in the art. Methods for stopping the reactions are known in the art, including, for example, cooling the reaction mixture to a temperature that inhibits enzyme activity. Methods for resuming the reactions are also known in the art, including, for example, raising the temperature of the reaction mixture to a temperature that permits enzyme activity. In some embodiments, one or more of the components of the reactions is replenished prior to, at, or following the resumption of the reactions. Alternatively, the reaction can be allowed to proceed (i.e., from start to finish) without interruption.

Compositions and kits of the invention

The invention also provides compositions and kits used in the methods described herein. The compositions may be any component(s), reaction mixture and/or intermediate described herein, as well as any combination. In one embodiment, the invention provides a composition comprising a composite primer, wherein the composite primer comprises a 5' RNA portion and a 3' DNA portion. In another embodiment, the invention provides a composition comprising a composite primer, wherein the composite primer comprises a 3'-RNA portion and a 5'-DNA portion. In one embodiment, the RNA portion is adjacent to the DNA portion. In another embodiment, the invention provides a composition comprising a composite primer, wherein the composite primer comprises 5'- and 3'-DNA portions with at least one intervening RNA portion. In another embodiment, the invention provides a composition comprising two composite primers that are hybridizable to two non-overlapping sequences of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA and a DNA portion. In other embodiments, the invention provides a composition comprising a composite primer that is further derivatized by attachment of a

moiety capable of effecting attachment of a polynucleotide comprising the composite primer to a solid substrate used in preparing nucleic acid microarrays. In some embodiments, the composite primer is further derivatized by attachment of a positively charged moiety such as an amine. In some embodiments, the compositions comprise two composite primers that are hybridizable to two non-overlapping sequences of a target polynucleotide, wherein the primers singly comprise at least one nucleotide of a sequence of interest, or in combination comprise a sequence of interest (or a portion of a sequence of interest). The compositions are generally in aqueous form, preferably in a suitable buffer.

The invention also provides compositions comprising the products described herein. Accordingly, the invention provides a population of DNA (anti-sense) copies of a sequence of interest, which are produced by any of the methods described herein. The DNA copies can be primer extension products or attached oligonucleotide combination products. In some embodiments, the DNA copies comprise a unique identifying characteristic the detection of which indicates the presence or absence of a sequence of interest. For example, the invention provides compositions comprising a cleaved primer extension product comprising a identifying characteristic the detection of which indicates presence of a nucleic acid sequence in a sample. In another embodiment, the cleaved primer extension product is of a size that the product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension. The invention also provides compositions comprising a cleaved attached oligonucleotide combination product, preferably covalently attached, comprising an identifying characteristic the detection of which indicates presence of a nucleic acid sequence in a sample. In one embodiment, the cleaved attached oligonucleotide combination product is of a size that the product dissociates from the target polynucleotide under essentially the same conditions as used for hybridization of the oligonucleotides, or under essentially the same conditions used for primer extension.

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).

The invention provides kits for carrying out the methods of the invention. Accordingly, a variety of kits are provided in suitable packaging. 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 methods of the invention described herein, including: copying a nucleotide sequence; quantifying a nucleotide sequence; detection of whether a sequence of interest is present or absent in a sample; and/or identifying whether an altered sequence of interest is present or absent in a sample. As such, these instructions may describe any one or more steps of the methods, the reaction conditions, and complexes, intermediates and/or products formed, and interpretation of results.

The kits of the invention comprise one or more containers comprising any combination of the components described herein, and the following are examples of such kits. A kit may comprise any of the composite primers described herein. In some embodiments, a kit comprises two or more composite primers, which may or may not be separately packaged. The composite primer may be labeled or unlabeled. Kits may also optionally include any of one or more of the enzymes described herein, as well as dNTPs, and/or analogs (such as terminator dNTPs) thereof. Kits may also include an enzyme that cleaves RNA from a RNA/DNA hybrid, such as RNase H. In some embodiments, kits include an agent (such as DNA ligase) that effects attachment of oligonucleotides that are hybridized to a nucleic acid sequence of interest. In another embodiment, the kits comprise an agent that effects attachment, preferably covalent attachment, of oligonucleotides that are hybridized to a nucleic acid sequence of interest, a polymerase, and an enzyme that cleaves RNA from a RNA/DNA hybrid, such as RNase H. Kits may also include one or more suitable buffers (as described herein). Kits may include labeled or unlabelled terminator dNTPs that upon incorporation into a primer extension product effect termination of nucleotide polymerization. One or more reagents in the kit can be provided as a dry powder, usually lyophilized, including

excipients, which on dissolution will provide for a reagent solution having the appropriate concentrations for performing any of the methods described herein. Each component can be packaged in separate containers or some components can be combined in one container where cross-reactivity and shelf life permit.

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 copying, quantification, determination whether a sequence of interest is present or absent in a sample and/or identifying whether an altered sequence of interest is present or absent in a sample. 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.

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

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.

The invention also provides systems for effecting the methods described herein. These systems comprise various combinations of the components discussed above. For example, in some embodiments, the invention provides systems suitable for generating multiple copies of, quantifying, determining whether a sequence of interest is present or absent in a sample and/or identifying whether an altered sequence of interest is present or absent in a sample by primer extension comprising (a) a composite primer (any of those described herein), (b) DNA polymerase; and (c) an enzyme (such as RNase H) that cleaves RNA from an RNA/DNA hybrid. These systems may include terminator dNTPs and/or at least one but not all four dNTP types. In yet another embodiment, the invention

provides systems suitable for generating multiple copies of, quantifying, determining whether a sequence of interest is present or absent in a sample and/or identifying whether an altered sequence of interest is present or absent in a sample by oligonucleotide attachment comprising (a) a first and a second oligonucleotide (at least one of these oligonucleotides is a composite primer); (b) optionally DNA polymerase; (c) an agent (such as DNA ligase) that effects attachment of the first oligonucleotide and second oligonucleotide when the oligonucleotides are hybridized to the nucleic acid sequence of interest; and (d) an enzyme (such as RNase H) that cleaves RNA from an RNA/DNA hybrid. These systems may also include dNTPs. Any of the systems embodiments may also comprise a target polynucleotide known to comprise or suspected of comprising a sequence of interest, as described herein. Systems may also comprise devices, apparatus and/or equipment such as, for example, a waterbath or heat block, robots and other automated equipment for dispensing template, reaction components, as well as performing reactions, devices for characterizing reaction products, and computer systems for collecting/analyzing results.

The invention also provides reaction mixtures (or compositions comprising reaction mixtures) which contain various combinations of components described herein. In some embodiments, the invention provides reaction mixtures comprising (a) a target polynucleotide known to comprise or suspected of comprising a sequence of interest; (b) a composite primer comprising a 3' DNA portion and an RNA portion; and (c) DNA polymerase. These reaction mixtures may also include terminator dNTPs and/or at least one but not all four types of dNTPs. In other embodiments, the invention provides a reaction mixture comprising (a) a target polynucleotide template; (b) a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, and wherein the two oligonucleotides hybridize to non-overlapping portions of the target polynucleotide; (c) optionally DNA polymerase; and (d) an agent (such as DNA ligase) that effects attachment, preferably covalent attachment, of the first oligonucleotide and second oligonucleotide when the oligonucleotides are hybridized to the target polynucleotide. As described herein,

any of the composite primers may be in the reaction mixture (or a plurality of composite primers). The reaction mixtures could further comprise an enzyme which cleaves RNA from an RNA/DNA hybrid, such as RNase H. Other reaction mixtures are described herein and are encompassed by the invention.

The invention also includes compositions comprising any of the complexes (which are intermediates in the methods described herein) described herein. As an example, one complex of the invention is a complex comprising: (a) a target polynucleotide known to comprise or suspected of comprising a nucleic sequence of interest; and (b) a composite primer. As another example, one complex of the invention is a complex comprising: (a) a target polynucleotide known to comprise or suspected of comprising a nucleic acid sequence of interest; and (b) two oligonucleotides (a first oligonucleotide and a second oligonucleotide) hybridized to non-overlapping sequences of the polynucleotide strand, wherein at least one of the oligonucleotides is a composite primer.

Reaction conditions, components, and other experimental parameters as well as illustrative embodiments in this section are generally as described herein.

Examples of methods of generating multiple copies, quantifying and detecting nucleic acid sequences of the present invention.

The following are examples of the primer extension-based and oligonucleotide attachment-based methods of the invention. It is understood that various other embodiments may be practiced, given the general description provided above. For example, reference to using a composite primer means that any of the composite primers described herein may be used.

Primer extension-based methods

A sample comprising target polynucleotide suspected to comprise a nucleic acid of interest is combined with a composite primer, a DNA polymerase and RNase H, in a reaction mixture containing 1, 2, or 3 dNTPs and one of four possible terminators (dideoxynucleotides), and buffer suitable for polymerase and RNase H activity. The combination is heated to a temperature suitable for

denaturation of the target polynucleotide (if in double stranded form), and cooled to a temperature which permits specific hybridization of the composite primer to the target polynucleotide, and is further incubated at this temperature for carrying out the method of the invention.

The composite primer is composed of a 3'-DNA portion, and a 5'-RNA portion. The sequence of the composite primer is complementary to a sequence of the target polynucleotide which is immediately adjacent to the sequence of interest (i.e. the sequence which varies between the various genotypes to be determined).

The 3'-end of the hybridized composite primer is extended along the target polynucleotide strand by DNA polymerase. The nucleotide attached to the 3'-end of the hybridized primer is complementary to the corresponding nucleotide of the target polynucleotide to which the primer is hybridized. The incorporation of either labeled nucleotide or labeled terminator to the primer by primer extension is detectable. The extension of the hybridized primer is limited by the nature of available (provided) dNTPs and is stopped following incorporation of a terminator.

RNase H degrades the RNA strand of the primer-target polynucleotide complex. The cleavage of the hybridized composite primer results in reduction of the size of the non cleaved portion of the primer extension product which leads to the dissociation of the primer extension product from the target polynucleotide (as illustrated in Figure 1), and/or creation of a free site for a new primer to hybridize to the target polynucleotide by hybridization of the 5' end of the new primer to the corresponding target polynucleotide sequence (as illustrated in Figure 2). As illustrated in Figure 2, the 3' DNA portion of the newly hybridized primer could also displace the 5' DNA portion of the cleaved primer extension product to initiate extension of the newly hybridized primer along the template strand by the polymerase. The extended portion of the cleaved primer dissociates from the template. Another composite primer then hybridizes to the target polynucleotide following cleavage of the primer extension product. This sequence of steps is

repeated and results in the accumulation of multiple copies of the cleaved primer extension product in the reaction mixture.

The reactions are carried out at a single temperature (isothermally), which is a temperature that permits specific hybridization of the two primers but is too high for hybridization of the digested ligation product. Either the terminator or non-terminator dNTPs may be labeled with a distinct label. The cleaved primer extension product could be labeled by the incorporation of labeled nucleotide, which would be indicative of the nucleotide sequence of the target polynucleotide immediately adjacent to the sequence complementary to the 3'-end of the primer.

Various methods for the detection of the reaction product are envisioned. The multiple copies of the products of the reaction may be detected by hybridization to complementary oligonucleotide(s) immobilized on a solid support. When the reaction product is labeled, the immobilization of the primer extension product on the solid support and detection of the label indicates the presence of the product in the reaction mixture. The amount of the label immobilized on the support as a result of hybridization of the primer extension product to a complementary oligonucleotide immobilized on the support can be directly related to the amount of the product in the reaction mixture. When different labels are used for the different dNTPs or their analogs, the detection of a specific label can be indicative of the presence or absence of the nucleic acid sequence of interest in the sample.

Homogeneous detection of the primer extension product can also be used. The optical properties of a label may be altered subsequent to attachment to the primer. For example, fluorescence polarization of fluorescent dyes attached to free nucleotide triphosphates has been shown to change upon attachment to a primer by a polymerase.

It is also possible to detect altered spectral properties of a label by means of energy transfer. When the primer is labeled by a donor or acceptor dye, and the nucleotide triphosphates, or their analogs, are labeled with acceptor or donor dyes, respectively, the incorporation of the dyes into a primer extension product enables energy transfer between the donor and acceptor dyes, thus resulting in specific

spectral properties of the attached dyes. Fluorescence dyes useful for this detection mode are known in the art and described herein.

Detection of the primer extension product by hybridization to specific oligonucleotides immobilized on an array is useful for multiplexing. Simultaneous extension of a plurality of primers according to the methods of the invention results in formation of a plurality of primer extension products, each indicative of the presence/absence and/or quantity of multiple sequences of interest in a sample. The reaction products may be hybridized to a plurality of oligonucleotides which are immobilized, preferably at spacially defined sites.

Oligonucleotide attachment-based methods

This example is illustrated in Figure 3. A polynucleotide suspected of comprising a nucleic acid sequence of interest, a first composite primer comprising a 3' DNA portion and an RNA portion 5' to the DNA portion, a second composite primer comprising a 5' end DNA portion and an RNA portion 3' to the DNA portion, a ligase, RNaseH and buffer components are combined in a reaction mixture. The reaction mixture is subjected to conditions that support hybridization of the two composite primers to the respective sequences of interest on the target polynucleotide, to form a trimolecular complex. The two hybridized primers are ligated by the ligase. RNaseH digests the RNA portion of the primer in the RNA/DNA heteroduplex portions of the primer-target polynucleotide hybrid. The resultant ligation product comprising the DNA portions of the two primers dissociates from the target polynucleotide due to the reduced affinity of the smaller ligation product relative to that of the nondigested primers. The steps of hybridization of the primers to the target sequences, ligation of the DNA portions and digestion of the RNA portions by RNase H are repeated to product multiple copies of the ligation product.

The reactions are carried out at a single temperature (isothermally), which is a temperature that permits specific hybridization of the two primers but is too high for hybridization of the digested ligation product. The choice of temperature and length of the DNA portions of the two primers can be defined based on

known parameters for hybridization of nucleic acid sequences. When the reaction is carried out at an elevated temperature, for example between 37°C to 70°C, the ligase and RNaseH used are of the thermostable variety, which are known in the art.

The detection of the ligation product indicates the presence of the nucleic acid sequence of interest in a sample. The accumulation of the reaction product is linear, and can be used for quantification of the amount of the nucleic acid sequence of interest in the sample.

Insofar as ligation of the two oligonucleotides requires hybridization of the respective 3' and 5' most nucleotides of the first and second oligonucleotides, the method can also be used for the detection and quantification of specific genotypes. Any sequence alteration in the nucleic acid sequence of interest relative to a reference (control) sequence, which prevents hybridization of 3' and/or 5' end of the respective oligonucleotides, also prevents the production of the ligation product. Thus, genotype-specific primers can be used for detection of specific alleles.

The accumulated ligation product is detectable by various means. The ligation product is characterized by joining of the DNA sequence portions of the two primers. The formation of the ligation product may be detectable by association of two labels. Alternatively, only one label can used in which case the product is detected by hybridizing it to a labeled oligonucleotide under conditions which are not suitable for hybridization of the either the first or the second oligonucleotide individually (i.e., when not ligated to each other). The third oligonucleotide may be immobilized on a solid surface and the hybridization of the labeled ligation product to the oligonucleotide results in an altered detectable signal due to the interaction of the labels on the ligation product and the third (immobilized) oligonucleotide. The immobilized oligonucleotide may be a part of an array of multiple oligonucleotides, each oligonucleotide being specific for a ligation product comprising a distinct sequence. Use of an array of multiple oligonucleotides allows for the detection and quantification of multiple nucleic acid sequences of interest within a sample (i.e., multiplexing), through the

detection and/or quantification of multiple ligation products from a single reaction mixture.

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

5

EXAMPLES

Example 1: *Detection of sequence of interest by limited primer extension*

Overview

A DNA sample suspected of comprising a sequence of interest is combined in a reaction mixture with a composite primer. The reaction mixture comprises labeled primer extension terminators, with or without one, two or three dNTPs, buffer components, divalent ions such as Mg^{2+} or Mn^{2+} required for enzyme activity and specific primer hybridization, an RNase inhibitor and additives which enhance enzyme activity. The reaction mixture is incubated at a temperature that is sufficiently high to denature all secondary structures in the target. Following incubation for about 2 minutes, the mixture is incubated further at 50°C to 65°C. A mixture of a DNA polymerase and RNase H is added to the reaction mixture and the reaction mixture is further incubated for about 30 to 120 min. The reaction mixture is cooled to room temperature and an aliquot of the mixture is subjected to analysis by gel electrophoresis, using denaturing 15 to 20% polyacrylamide gel.

The incorporation of a label into the primer extension product is detected by means that are suitable for detection of the label. When the primer extension terminators are labeled by a flourophore, each specific to a specific terminator, fluorescence detectors are employed to determine the specific label attached to the primer extension product. The method allows the determination of the specific nucleotide in the target oligonucleotide sequence at the position adjacent to the site of hybridization of the 3'-end of the composite primer.

The reaction mixture may contain each of the labeled primer extension terminators, ddATP, ddCTP, ddGTP and ddTTP, either individually or combination thereof.

When the primer extension terminators are labeled with a fluorophore, the attachment of specific terminator to the truncated primer extension product can be measured by fluorescence polarization, a homogeneous detection method which

does not require separation of the reaction products. The measure of change in fluorescence polarization is directly related to the attachment of the specific nucleotide and the amount of the target nucleic acid sequence in the sample to be tested, in so far as the formation of truncated primer extension product is linear with respect to the target concentration.

Two synthetic target polynucleotides (GT01 & GT02) that differ in sequence from each other at two single nucleotide positions (hereinafter "variant sequence"; marked by bold, italic letter in the sequences listed below) are prepared.

Experimental details

Synthetic target 1: GT01

5'-
GGGAATTCGAATCTGCAGCT*T*TGTGGCTGCACCATCTGTCTTCAAGCTT
*A*ACACTGGAGACCGCATCCGTCAAAAAAAAA-3' (SEQ ID NO:1)

Synthetic target 2: GT02

5'GGGAATTCGAATCTGCAGCT*G*TGTGGCTGCACCATCTGTCTTCAAGC
TT*C*ACACTGGAGACCGCATCCGTCAAAAAAAAA-3' (SEQ ID NO:2)

Underlined sequence: complementary to primer 1.
Double underlined sequence: complementary to primer 2.

Two composite primers (primer 1 and primer 2) comprising a sequence complementary to the target polynucleotide immediately adjacent to either one of the variant sequences in the two target are employed.

Primer 1:

5'-*gaagacagatggtgca*GCCAGCA-3' (SEQ ID NO:3)

Italics: ribonucleotides

Primer 2:
5'-*gacggatgcggtct*CCAGTCT (SEQ ID NO:4)
Italics: ribonucleotides

Each of the composite primers comprises a 3'-end of 7 deoxynucleotides. The 5'-portion of primer 1 comprises 15 ribonucleotides and that of primer 2 comprises 14 ribonucleotides.

Reactions for the determination of the sequence identity at the sites marked using the two primers are carried out as follows. Target polynucleotides are combined with primers at a final concentration of 1 uM in a reaction mixture containing buffer (20 mM Tris, 5 mM $MgCl_2$, 0.1% NP40) and one dye-labeled terminator dNTP (final concentration of 100 mM). The reaction mixture is heated to 90°C for 1 min. and then incubated at 55°C for 5 min. A mixture of DNA polymerase (Bca, 2.5 units), RNase H (0.25 units) in the same buffer is added to a total volume of 20 ul. All mixtures contain 0.6 Units/ul recombinant ribonuclease inhibitor (rRNasin, Promega, Madison, WI). The mixture is further incubated for 30 min. Aliquots of each reaction (containing the individual dye-labeled terminator) are analyzed by gel electrophoresis on a commercial sequencing apparatus. The GT01 polynucleotide tested with primer 1 results in the formation of a labeled truncated primer extension product comprising the sequence: 5'-GCCAGCAA*-3, where A* denotes target-directed incorporation of a labeled ddATP. The GT02 polynucleotide results in the production of a truncated primer extension product comprising the sequence: 5'- GCCAGCAC*-3, where C* denotes target-directed incorporation of labeled ddCTP.

Similar experiments using primer 2 and target polynucleotide GT01 results in the production of a truncated extension product comprising the sequence 5-CCAGTCTT*-3', where T* denotes target-directed incorporation of labeled ddTTP. Similarly, when polynucleotide GT02 is tested with primer 2, the reaction results in the production of a truncated primer extension product

comprising the sequence 5'-CCAGTCTG*-3', where G* denotes target-dependent incorporation of labeled ddGTP. A sequence detection reaction that is carried out in the presence of the two target polynucleotides and one of the primers mimics the testing for a heterozygote genotype, where both alleles are present in a sample, and would result in the production of two truncated primer extension products with distinctive labels (detectable identifying characteristics).

General methods and material description

These general methods and materials are used in Examples 2-9 provided herein.

Use of primer-extension based methods of the invention to determine whether a single based pair polymorphism (SNP) is present in a sample

Target polynucleotide(s) suspected of comprising a sequence(s) of interest was combined in a reaction mixture with a composite primer. The reaction was performed in 20ul reactions containing isothermal buffer (20mM Tris-Cl, PH 8.5 and 5mM $MgCl_2$) or AcycloPrime-FP reaction buffer (PerkinElmer Life Sciences AcycloPrime-FP SNP Detection Kit), target DNA (various copy numbers of DNA target as noted in the descriptions of experiments), composite primer at final concentration of 1μM, ribonuclease inhibitor (6 units of Rnasin from Promega), fluorescence-tagged DNA terminator nucleotide (also called "ddNTP" herein) (5 pmol of Rox-ddUTP or R110-ddGTP obtained from Amersham; or 1μl of acyNTP terminators obtained from PerkinElmer Life Sciences AcycloPrime-FP SNP Detection Kit), DNA polymerases (4 units of Bst from New England Biolab or 0.005 μl of AcycloPol from PerkinElmer Life Sciences AcycloPrime-FP SNP Detection Kit), Hybridase (RNase H, 0.05 unit unless noted otherwise, Epicentre), BSA (2mg, New England Biolab), and DTT (0.5mM final, Life Technologies Inc.). Reaction buffer, DNase–RNase free H_2O, target DNA, Rnasin, and composite primer were combined in total of 12μl in 0.2 ml thin-wall

polypropylene PCR tubes. These pre-reaction mixtures were denatured at 95°C for 5 min, cooled to 0°C rapidly, and placed on ice. Eight µl of enzyme mixture containing reaction buffer, DNase–RNase free H_2O, DNA polymerase, Hybridase, BSA, and DTT was dispensed to each pre-reaction mixture tubes on ice. The reactions were carried out at 65°C for 40 min and stopped by heat inactivation at 95°C for 5 min.

Reaction products were treated with Shrimp Alkaline Phosphatase (at final concentration of 0.2 unit/µl, United States Biochemical) at 37°C for 60 min to hydrolyze unincorporated ddNTPs remaining in the reaction mixture, followed by phosphatase inactivation by incubation at 75°C for 15 min. Two µl of shrimp alkaline phosphatase-treated reaction product was diluted with 18 µl of Hi-Di formamide (Applied Biosystems). The samples were analyzed by capillary electrophoresis (ABI 310 Genetic Analyzer, Applied Biosystems) following the manufacturer's instructions. Briefly, oligonucleotides of different lengths were separated by capillary electrophoresis, then incorporation of fluorescently labeled nucleotides was detected using fluorescence detection. In some cases, fluorescence intensity was quantified, permitting quantification of fluorescently labeled cleaved primer extension product. In some cases, a set of length markers were included and the size of the products was analyzed. Further dilutions were made with Hi-Di formamide for re-analyzed when the signal intensity was too strong. Data analyses and electropherograms were processed and generated with GeneScan 3.1.2 software from Applied Biosystems.

Nucleotide sequence of composite primers and template polynucleotides used in the Examples 2-9

Composite primers

IA20 5′- *gac gga ugc ggu cu*C CAG TGT (SEQ ID NO:5)

Italics: ribonucleotides

IA30 5′-*aau acg acu cac uau* AGG CAG A (SEQ ID NO:6)

Italics: ribonucleotides

IA21.2 5′-*gaa gac aga ugg ugc* AGC CAC A (SEQ ID NO:7)

Italics: ribonucleotides

221s 5′-*ccc ucc aag gcu ccc* CAG TAT C (SEQ ID NO:8)

Italics: ribonucleotides

241r 5′-*aug gua ggu ggc agG* ATT CAG (SEQ ID NO:9)
Italics: ribonucleotides

Template DNA sequences

VL target sequence

5′-ATGGATAAAT AGCCTTGCTT GCTTCCTATT ATATCTTCCC
AAATTACCAA
TACATTACAC TAGCATCTGA ATTTCATAAC CAATCTCGAT ACACCAAATC
GACTCTAGAG GATCTAACCA TGGGATGGAG CTGGATCTTT CTCTTCCTCC
TGTCAGGAGC TGCAGGTGGT ACCTCAAGCG ACATTCAGCT GACCCAGTCT
CCAGCCTCCC TATCTGCATC TGTGGGAGAA ACTGTCACCA TCACATGTCG
AGCAAGTGAG AATATTTACA GTTATTTAGC ATGGTATCAA CAGAAACAGG
GAAAATCTCC TCAGTTCCTG GTCTATAGTG CAAAAACCTT AGCAGAAGGT
GTGCCATCAA GGTTCAGTGG CAGTGGATCA GGCACACAGT TTTCTCTGAA
GATCAACAGC CTGCAGCCTG AAGATTTTGG GAATTATTAC TGTCAACATT
ATTATGGTAG TCCGCGCACG TTCGGGTGCT GGGACCAAGC TTGAGATCAA
ACGAACTGTG GCTGCACCAT CTGTCTTCAT CTTCCCGCCA TCTGATGAGC
AGTTGAAATC TGGAACTGCC TCTGTTGTGT GCCTGCTGAA TAACTTCTAT
CCCAGAGAGG CCAAAGTACA GTGGAAGGTG GATAACGCCC TCCAATCGGG
TAACTCCCAG GAGAGTGTCA CAGAGCAGGA CAGCAAGGAC AGCACCTACA
GCCTCAGCAG CACCCTGACG CTGAGCAAAG CAGACTACGA GAAACACAAA
GTCTACGCCT GCGAAGTCAC CCA-3′
(SEQ ID NO:10)

Syn221

5′-TCTCAGGTTT CAGGGATTAG GGAGATATTA TTTGGCCAAA
CACACAAACG GAGATGAAAA GGGAAAGATG TGCCAGATAC TGGGGAGCCT
TGGAGGGTTG-3′
(SEQ ID NO:11)

41r21sGA
5′-
TGACTGCTGAATCCTGCCACCTACCATGACTGACAGATACTGGGGAGCCTTGGA
GGG TTG-3′
(SEQ ID NO:12)

41r21sTG
5'-
TGACTTCTGAATCCTGCCACCTACCATGACTGACGGATACTGGGGAGC
CTTGGAGG GTTG-3'
(SEQ ID NO:13)

GT01
5′-
GGGAATTCGAATCTGCAGCTTTGTGGCTGCACCATCTGTCTTCAAGCTTCACAC
TGGAGACCGCATCCGTCAAAAAAAAA-3′
(SEQ ID NO:1)

GT02
5′-
GGGAATTCGAATCTGCAGCTGTGTGGCTGCACCATCTGTCTTCAAGCTTAACAC
TGGAGACCGCATCCGTCAAAAAAAAA-3′
(SEQ ID NO:2)

GT03
5'-
GATGACGGATGCGGTCTCCAGTGTGGCCCAGGACCAGCTCGCTCCTACACTGGA
CCCAATTGGGAGCACCAAGCAAGTTGCG*TCTGCCTATAGTGAGTCGTATT*ACC-
3′
(SEQ ID NO:14)

Notes relating to template sequences:

1. Larger underlined letter: the single nucleotide sequence of interest (the nucleotide immediately adjacent to the site of hybridization of the composite primer).
2. Double underlined sequences: complementary to primer used in primer extension-based method.

3. For GT03 target polynucleotide: Single underlined sequence in Italics: complementary to primer used for single primer isothermal amplification of the template. Double underlined sequence: corresponds to primer extension primer sequence. Larger underlined letter is the nucleotide that will be incorporate during single base extension. During the reaction, the primer extension primer binds to the complement of the double underlined sequence (contained in the copies of the target sequence produced during single primer isothermal amplification).

Example 2: Detection of a single nucleotide sequence of interest using limited primer extension

This Example demonstrates that a single nucleotide sequence of interest in a single template can be detected using limited primer extension methods described herein.

An estimated 10^8 copies of VL were used as DNA template. Composite primer IA21.2 was used at final concentration of 1μM. Composite primer IA21.2 comprises a sequence complementary to the target sequence immediately adjacent to the nucleotide of interest, C. Limited primer extension in the presence of a terminator GTP (ddGTP) is expected to produce a limited primer extension product incorporating a single terminator dGTP nucleotide (ddGTP).

Incorporation of the labeled termination nucleotide analog R110-ddGTP (Amersham) was tested. Limited primer extension was conducted using either Bst- or AcycloPol-DNA polymerases, in either isothermal buffer or AcycloPrime-FP reaction buffer. Bst polymerase exhibits strong strand displacement activity, while AcycloPol-DNA polymerase lacks strand displacement activity. The reactions were carried out as described above, and cleavage primer extension products were analyzed by capillary electrophoresis, as described above.

The results are shown in Figure 4. Panel A is an electropherogram from the reaction performed with Bst polymerase in AcycloPrime-FP reaction buffer. Panel B is an electropherogram from the reaction performed with AcycloPol in AcycloPrime-FP reaction buffer. Panel C is an electropherogram from the reaction performed with Bst polymerase in isothermal buffer. Panel D is an

electropherogram from the reaction performed with AcycloPol in isothermal buffer. Signals of high intensity were observed in panels A and C, indicating that R110-ddGTP can be incorporated effectively by Bst polymerase with limited primer extension performed in either isothermal buffer or AcycloPrime-FP reaction buffer.

Example 3: Optimization of single base limited primer extension using Acycloterminators

An estimated 10^8 copies of VL were used as DNA template. Composite primer IA21.2 was used at a final concentration of 1μM. As noted above, composite primer IA21.2 comprises a sequence complementary to the target sequence immediately adjacent to the nucleotide of interest, C. Limited primer extension in the presence of a terminator GTP (ddGTP) is expected to produce a limited primer extension product incorporating a single terminator dGTP nucleotide (ddGTP).

Limited primer extension was performed as described above, except that it was performed in the presence of an acycloterminator mixture containing two labeled terminator nucleotide analogs, R110-acyGTP and Tamra-acyATP (PerkinElmer, acycloterminator mixture), either Bst- or AcycloPol-DNA polymerase, in either isothermal buffer or AcycloPrime-FP reaction buffer. In reaction A, the limited primer extension reaction was performed with Bst DNA polymerase in AcycloPrime-FP reaction buffer. In reaction B, the limited primer extension reaction was performed with AcycloPol DNA polymerase in AcycloPrime-FP reaction buffer. In reaction C, the limited primer extension reaction was performed with Bst DNA polymerase in isothermal buffer. In reaction D, the limited primer extension was performed with AcycloPol DNA polymerase in isothermal buffer. Cleaved primer extension product was analyzed as described above.

Analysis of electropherograms indicated that signals of high intensity were observed only in panel B and that the cleaved primer extension product incorporated R110-acyGTP (the expected terminator according to the target

sequence). As expected, no cleaved primer extension product incorporating Tamra-acyATP was observed. This example demonstrates that the correct acycloterminator can be incorporated effectively by AcycloPol DNA polymerase using AcycloPrime-FP reaction buffer.

Example 4: Limited primer extension performed in the presence and absence of single stranded DNA binding protein and RNase H activity

This example describes testing whether cleavage of the RNA portion of a composite primer in the primer extension product is required for the primer extension-based limited primer extension method of the invention and whether the use of single stranded DNA binding protein enhances limited primer extension.

An estimated 10^{10} copies of synthetic oligonucleotide synthetic target Syn221 were used as DNA template. Composite primer 221s was used at final concentration of 1μM. Composite primer 221s comprises a sequence complementary to the target sequence Syn221 immediately adjacent to the nucleotide of interest, A. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a limited primer extension product incorporating the complementary terminator nucleotide. Rox-ddUTP was used as a terminator, as it is expected to be incorporated during limited primer extension. Limited primer extension reactions were performed as described above, with or without single stranded DNA binding protein T4 gp32 (1.5 μg), and with or without Hybidase (RNase H, 0.05 unit). Reaction products were analyzed as described above.

Analysis of electropherograms indicated that the production of accumulated dissociated labeled cleaved primer extension product occurred only in samples containing RNase H, indicating that cleavage of the RNA portion of a composite primer in the primer extension product is required for the primer extension-based methods of the invention. The presence of the single stranded binding protein, T4gp32, resulted in a slight, insignificant, decrease in signal compared to the signal in reaction mixtures lacking T4gp32.

Example 5: Optimization of limited primer extension reaction for RNase activity

An estimated 10^{10} copies of synthetic oligonucleotide target Syn221 were used as DNA template. Composite primer 221s was used at the final concentration of 1μM. Composite primer 221s comprises a sequence complementary to the target sequence Syn221 immediately adjacent to the nucleotide of interest, A. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a limited primer extension product incorporating the complementary terminator nucleotide. Rox-ddUTP was used as a terminator, as it is expected to be incorporated during limited primer extension.

Limited primer extension reactions were performed as described above, except that different amounts of Hybridase (RNase H) were included in the reaction mixtures as follows: 0.2 Unit of Hybridase; 0.1 Unit of Hybridase; 0.05 Unit of Hybridase; 0.0125 Unit of Hybridase; 0.00315 Unit Hybridase; and a control reaction lacking Hybridase. Reaction products were analyzed as described above. Analysis of electropherograms revealed that no cleaved primer extension product (i.e., no accumulation of dissociated product) was produced if Hybridase was omitted. Cleaved reaction products were produced in all samples in which Hybridase was included; however, signal intensity dropped significantly between reactions carried out with 0.05 and 0.0125 Units of Hybridase.

Example 6: Multiplexed limited primer extension

This example demonstrates that limited primer extension using two composite primers can be used to detect multiple sequences of interest in a single template.

Limited primer extension reactions were performed with two composite primers (221s and 241r) using a defined target sequence 41r21sGA (estimated 10^9 copies) comprising binding sites for both composite primers. Composite primer 221s comprises a sequence complementary to the target sequence 41r21sGA immediately adjacent to the nucleotide of interest, A. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a

limited primer extension product incorporating the complementary nucleotide, T (ddTTP). Composite primer 241r comprises a sequence complementary to the target sequence 41r21sGA immediately adjacent to the nucleotide of interest, G. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a limited primer extension product incorporating the complementary nucleotide, C (ddCTP).

Limited primer extension reactions were carried out as described above using AcycloPro DNA polymerase (which lacks strand displacement activity), in the presence of either composite primer 221s or composite primer 241r, or in the presence of both composite primers. The reaction mixtures included a mixture of acycloterminators R110-acyCTP and Tamra-acyTTP, or a control mixture of R110-acyGTP and Tamra-acyATP. A control reaction was performed in the presence of both composite primers, but omitting fluorescent acycloterminators. Reaction mixtures containing template and reaction products were analyzed as described above.

The results are shown in Figure 5. Panels A, B, and C show electropherograms of cleaved primer extension products from reactions containing R110-acyCTP and Tamra-acyTTP acycloterminators. Panel D shows an electropherogram of product from a reaction containing R110-acyGTP and Tamra-acyATP acycloterminators. Reaction A was carried out in the presence of composite primer 221s, and cleaved primer extension product was produced that correctly incorporated acycloterminator Tamra-acyTTP (marked by arrow 1). Reaction C was carried out in the presence of composite primer 241r, and cleaved primer extension product was produced that correctly incorporated acycloterminator Tamra-acyCTP (marked by arrow 2). Reaction B was carried out in the presence of both composite primers 221s and 241r and cleaved primer extension product was produced that correctly incorporated either acycloterminator Tamra-acyTTP or acycloterminator Tamra-acyCTP (marked by arrows 1 and 2, respectively). Reaction D was carried out in the presence of both composite primers 221s and 241r and a mixture of R110-acyGTP and Tamra-

acyATP acycloterminators, and, as expected, no fluorescently labeled product was produced.

This example shows that single nucleotide limited primer extension of the invention may be used to detect a plurality of sequence of interest on a single template, thus permitting the determination of multiple genotypes of a single template in a single reaction.

Example 7: Limited primer extension reactions for the determination of a mixture of targets of different polymorphic genotypes (e.g., homozygous, heterozygous), as simulated by a mixture of defined genotype synthetic targets

This example demonstrates that that limited primer extension method as described herein can be used to detect two variant sequences of interest (polymorphic genotypes) in a mixture of two target polynucleotides.

Limited primer extension reactions were performed with composite primer 221s using a mixture of two target sequences which differ at a single defined polynucleotide. Defined target sequence 41r21sGA (Reaction A) possesses an A at the sequence of interest and thus simulates a A/A genotype at the sequence of interest. Defined target sequence 41r21sTG (Reaction C) possesses a G at the sequence of interest, and thus simulates a G/G genotype at the sequence of interest). A mixture of 50% 41r21sGA plus 50% 41r21sTG (Reaction B) contains template polynucleotide possessing an A at the sequence of interest (which is a single nucleotide), and template possessing a G at the sequence of interest (which is a single nucleotide), thus simulating a A/G genotype at the sequence of interest.

Composite primer 221s comprises a sequence complementary to the target sequence 41r21sGA immediately adjacent to the nucleotide of interest, A. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a limited primer extension product incorporating the complementary nucleotide, T (ddTTP). Composite primer 221s comprises a sequence complementary to the target sequence 41r21sTG immediately adjacent to the nucleotide of interest, G. Limited primer extension in the presence of a complementary terminator nucleotide is expected to produce a

limited primer extension product incorporating the complementary nucleotide, C (ddCTP).

Limited primer extension was performed as described above, in the presence of R110-acyCTP and Tamra-acyTTP acycloterminators. Cleaved primer extension products were analyzed as described above.

Reaction A contained template polynucleotide that simulated a homozygous A/A genotype at the sequence of interest), and, as expected, limited primer extension cleavage product was produced that incorporated Tamra-acyTTP. Reaction C contained template polynucleotides that simulated a homozygous G/G genotype at the sequence of interest and, as expected, limited primer extension cleavage product was produced that incorporated Tamra-acyCTP. Reaction B contained template polynucleotides that simulated heterozygous A/G genotype at the sequence of interest and, as expected, two limited primer extension cleavage products were produced that incorporated either Tamra-acyTTP or Tamra-acyCTP. No labeled limited primer extension cleavage product was produced in control reactions performed in the absence of a composite primer, or in the absence of template DNA. These results clearly demonstrate that the single nucleotide limited primer extension method may be used to analyze a mixture of target sequences which differ at defined single polynucleotides, and to detect multiple sequence of interest at a defined single polynucleotide. Put another way, the limited primer extension method may be used to correctly determine the presence of either a homozygous or heterozygous genotype.

Example 8: Quantification of genotypes in a single sample using single nucleotide limited primer extension

This example demonstrate that different genotypes (target polynucleotides) present in a single sample can be quantified using single nucleotide limited primer extension.

Limited primer extension reactions using composite primer IA20 were performed on mixtures of target sequences GT01 and GT02 at different ratios.

The two targets simulate two genotypes of a defined sequence. Composite primer IA20 is expected to incorporate T when reacted with template GT01 and C when reacted with template GT02. Single nucleotide limited primer extension was performed in the presence of various templates and either acycloterminators R110-acyCTP and Tamra-acyTTP or acycloterminators R110-acyGTP and Tamra-acyATP. The reactions were carried out as described above, except that two different concentrations of AcycloPol DNA polymerase were used: either 0.05 ul of polymerase solution (Perkin Elmer) or a 1:16 dilution of this amount. Reactions A1, B1, C1, D1, and E1 were performed with acycloterminators R110-acyCTP and Tamra-acyTTP. Reactions A2, B2, C2, D2, and E2 were performed with acycloterminators R110-acyGTP and Tamra-acyATP. Reactions mixtures for A1 and A2 contained only GT01 target (100% GT0). Reaction mixtures for B1 and B2 contained 75% GT01 and 25% GT02. Reaction mixtures for C1 and C2 contained 50% GT01 and 50% GT02. Reaction mixtures for D1 and D2 contained 25% GT01 and 75% GT02, and reaction mixtures for E1 and E2 contained 100% GT02. The reaction products were analyzed as described above, and further quantification of the two genotypes in the sample was assisted by incorporation of a reference dye, LIZ (ABI), to each injected reaction mixture, which served to normalized signal intensity for the two incorporated dyed terminators and variation of injection conditions.

Analysis of electropherograms indicated that single nucleotide primer extension using composite primer IA20 resulted in the correct incorporation of either Tamra-acyTTP or R110-acyGTP into the cleaved primer extension product when target GT01 and GT02 were used as templates, respectively. Correct incorporation of one or both labeled nucleotides was observed when each template was present alone in a sample or when the sample contained mixtures of the two templates. These results are shown in Table 1. Quantification of amount of cleaved primer extension product in each reaction was performed, and the ratios of fluorescence intensities and internal fluorescence references were tabulated in Table 1. These results demonstrate that the signal intensities of each cleaved primer extension product were proportional to the mixtures of GT01 and

GT02 target templates. Thus, quantification of the relative proportion of each cleaved primer extension product permits the quantification of the relative proportion of each template polynucleotide in the sample mixture. Further quantification of the two genotypes in the sample was assisted by incorporation of a reference dye to each injected reaction mixture, which served to normalized signal intensity for the two incorporated dye terminators and variation of analysis conditions. As shown in Table 1, the ratio of the specific dye terminator signal to the internal reference signal (normalized value) is proportional to the ratio of the two genotypes in the samples. RFU denotes "reference fluorescence units".

Table 1

Template	Genotype mixture	(GT01+GT02=100%)				
	GT01:	100%	75%	50%	25%	0%
	GT02:	0%	25%	50%	75%	100%
	(sample number)	(A1, A2)	(B1, B2)	(C1, C2)	(D1, D2)	(E1, E2)
Terminators						
1/16x Polymerase						
C, T	T Signal (RFU)	656.0	461.0	348.0	201.0	0.0
	Reference	128.0	126.0	121.0	126.0	123.0
	Ratio	5.1	3.7	2.9	1.6	0.0
G, A	G Signal (RFU)	0.0	1208.0	2500.0	3415.0	4465.0
	Reference	113.0	109.0	125.0	115.0	118.0
	Ratio	0.0	11.1	20.0	29.7	37.8
1x Polymerase						
C, T	T Signal (RFU)	1023.0	817.0	560.0	296.0	0.0
	Reference	122.0	124.0	114.0	118.0	122.0
	Ratio	8.4	6.6	4.9	2.5	0.0
G, A	G Signal	0.0	1956.0	3780.0	4840.0	5325.0

(RFU)					
Reference	117.0	115.0	112.0	114.0	116.0
Ratio	0.0	17.0	33.8	42.5	45.9

Example 9: Single primer isothermal linear amplification of target nucleic acid sequence (template pre-amplification) and genotyping of amplification product by the limited primer extension method of the invention

This Example demonstrates that a template polynucleotide can be preamplified using single primer isothermal linear amplification, then limited primer extension can be used to detect a sequence of interest on the target polynucleotide produced by amplification of the template.

Single primer isothermal amplification pre-amplification of target polynucleotide GT03

Single primer isothermal linear amplification (hereinafter, "SPIA") of target polynucleotide GT03 was carried out as follows: composite primer IA30, comprising a 3' DNA portion and a 5' RNA portion, was used for amplification of a defined target DNA sequence. The reaction was carried out in Tris buffer at pH 8.5, 0 to 50 mM KCl, 2 to 5 mM $MgCl_2$, 0.25 to 0.5 mM dNTPs, 3 ug T4gp32 (USB), Bst DNA polymerase, RNase H , and 1 to 5 mM DTT. Reactions A, C, D, E, and F contained 10^4 molecules of template polynucleotide GT03, and Reaction B contained 10^3 molecules of template polynucleotide GT03.

The reaction mixtures containing the composite primer and the template polynucleotide were first denatured by incubation at 95^0C for 2 to 5 min., and the primer was allowed to anneal to the respective target by incubation at 55^0C for 5 min. The enzyme mixture was then added to the reaction tubes and incubated at 55^0C for 30 min., permitting amplification of the template.

Single base-pair limited primer extension

Amplification reaction products were treated with shrimp alkaline phosphatase to hydrolyze residual dNTPs. For some of the samples, reaction products were diluted at a 1:10 ratio with water prior to shrimp alkaline phosphatase treatment. Amplification product was then diluted into limited

primer extension reaction mixture containing composite primer IA20 and a mixture of acycloterminators R110-acyGTP and Tamra-acyATP. Composite primer IA20 is expected to incorporate G (dGTP) when reacted with single stranded polynucleotide complementary to template polynucleotide GT03 (produced by SPIA amplification of template GT03 as described above). Negative controls of no Hybridase (D), wrong terminators (R110-acyCTP and Tamra-acyTTP) (E), and no composite primer (F) were also included.

The products of limited primer extension were analyzed by capillary electrophoresis (ABI 310) as described above. Cleaved primer extension product containing correctly incorporated R110-acyGTP was observed in Reactions A, B, and C. As expected, tamra-acyATP was not incorporated into the cleaved primer extension product. No labeled cleaved primer extension product was observed in Reactions D (no Hybridase control), E (wrong terminators) and F (no composite primer).

These results clearly demonstrate that very sensitive and specific genotyping (sequence identification) is possible using single base pair limited primer extension according the primer extension-based method of the invention, with single primer linear amplification pre-amplification of the target nucleic acid sequence.

CLAIMS

What is claimed is:

1. A method of generating multiple copies of a nucleic acid sequence of interest, said method comprising the steps of:

(a) hybridizing a composite primer to a target polynucleotide, wherein the composite primer comprises an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion hybridizes from about 1 to about 10 nucleotides from the sequence of interest;

(b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product is produced; and

(c) cleaving the RNA portion of the primer extension product of (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved primer extension product dissociates from the target polynucleotide,

wherein the primer extension product is of a size that when the RNA is cleaved the cleaved primer extension product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension, whereby multiple copies of the sequence of interest are produced.

2. The method of claim 1, wherein said conditions that permit primer extension comprise at least one terminator deoxyribonucleotide triphosphate or analog thereof.

3. The method of claim 1, wherein said conditions that permit primer extension comprise the absence of a nucleotide that must be provided for primer extension to continue.

4. The method of claim 1, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest.

5. The method of claim 4, wherein the conditions for primer extension comprise one terminator deoxyribonucleotide triphosphate or analog thereof.

6. The method of claim 1, wherein the sequence of interest is a single nucleotide base.

7. The method of claim 1, wherein the sequence of interest is more than a single nucleotide base.

8. The method of claim 1, wherein the RNA portion of the composite primer is 5' with respect to the 3' DNA portion.

9. The method of claim 8, wherein the 5' RNA portion is adjacent to the 3' DNA.

10. The method of claim 1, wherein the enzyme that cleaves RNA is RNase H.

11. The method of claim 1, wherein the DNA polymerase lacks strand displacement activity.

12. The method of claim 1, wherein the target polynucleotide is DNA or cDNA.

13. The method of claim 1, wherein the target polynucleotide is amplified.

14. The method of claim 13, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

15. The method of claim 13, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

16. A method of generating multiple copies of a nucleic acid sequence of interest, said method comprising the steps of:

(a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(b) optionally extending the first oligonucleotide;

(c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product; and

(d) cleaving the RNA portion of the attached oligonucleotide combination product of (c) with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved oligonucleotide combination product dissociates from the target polynucleotide,

wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved from the attached oligonucleotide combination product, the cleaved attached oligonucleotide product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, whereby multiple copies of the sequence of interest are produced.

17. The method of claim 16, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

18. The method of claim 16, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest.

19. The method of claim 16, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

20. The method of claim 16, wherein the 5' most nucleotide of the second oligonucleotide is hybridizable to at least one nucleotide of the sequence of interest.

21. The method of claim 16, wherein the sequence of interest is a single nucleotide base.

22. The method of claim 16, wherein the sequence of interest is more than a single nucleotide base.

23. The method of claim 16, wherein the first and second oligonucleotide each comprise a sequence hybridizable to at least one nucleotide of the sequence of interest.

24. The method of claim 16, wherein the enzyme that cleaves RNA is RNase H.

25. The method of claim 16, wherein the DNA polymerase lacks strand displacement activity.

26. The method of claim 16, wherein the target polynucleotide is DNA or cDNA.

27. The method of claim 16, wherein the target polynucleotide is amplified.

28. The method of claim 27, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

29. The method of claim 27, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

30. The method of claim 16, wherein the attaching is covalent.

31. The method of claim 30, wherein the covalent attachment is effected by ligase.

32. A method of determining whether a nucleic acid sequence of interest is present or absent in a sample, said method comprising the steps of:

(a) hybridizing a composite primer to a target polynucleotide, wherein the composite primer comprises an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the DNA portion of the composite primer hybridizes from about 1 to about 10 nucleotides from the sequence of interest;

(b) extending the composite primer with DNA polymerase under conditions that permit primer extension, whereby a primer extension product comprising a detectable identifying characteristic is produced if the sequence of interest is present; and

(c) cleaving the RNA portion of the primer extension product of (b), if any, with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved primer extension product dissociates from the target polynucleotide,

wherein the primer extension product is of a size that when the RNA is cleaved the cleaved primer extension product dissociates from the target polynucleotide under essentially the same conditions as those for primer extension, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates the presence of the sequence of interest.

33. The method of claim 32, wherein said conditions that permit primer extension comprise at least one terminator deoxyribonucleotide triphosphate or analog thereof.

34. The method of claim 32, wherein said conditions that permit primer extension comprise the absence of a nucleotide that must be provided for primer extension to continue.

35. The method of claim 32, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest.

36. The method of claim 32, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest, the sequence of interest is a single nucleotide, and the conditions for primer extension comprise one terminator deoxyribonucleotide triphosphate or analog thereof.

37. The method of claim 32, wherein the sequence of interest is a single nucleotide base.

38. The method of claim 32, wherein the sequence of interest is more than a single nucleotide base.

39. The method of claim 32, wherein the RNA portion of the composite primer is 5' with respect to the 3' DNA portion.

40. The method of claim 32, wherein the 5' RNA portion is adjacent to the 3' DNA portion.

41. The method of claim 32, wherein the enzyme that cleaves RNA is RNase H.

42. The method of claim 32, wherein the DNA polymerase lacks strand displacement activity.

43. The method of claim 32, wherein the target polynucleotide is DNA or cDNA.

44. The method of claim 32, wherein the target polynucleotide is amplified.

45. The method of claim 44, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

46. The method of claim 44, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

47. The method of claim 32, wherein the target polynucleotide is attached to an analyte.

48. The method of claim 47, wherein the analyte is selected from the group consisting of a polypeptide, an antibody, an organic molecule and an inorganic molecule.

49. The method of claim 32, wherein said detectable identifying characteristic is selected from the group consisting of size of the cleaved primer extension product, sequence of the cleaved primer extension product, and detectable signal associated with the cleaved primer extension product.

50. The method of claim 32, wherein the detectable identifying characteristic comprises the sequence of the cleaved primer extension product, wherein the sequence is detected by hybridizing the cleaved primer extension product with a nucleic acid probe that is hybridizable to the cleaved primer extension product.

51. The method of claim 50, wherein said nucleic acid probe comprises DNA.

52. The method of claim 50, wherein the nucleic acid probe is provided as an array.

53. The method of claim 52, wherein the array comprises the probe immobilized on a substrate fabricated from a material selected from the group consisting of paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, silicon, polystyrene, and optical fiber.

54. The method of claim 49, wherein said detectable signal is associated with a label on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

55. The method of claim 49, wherein said detectable signal is associated with interaction of two labels, wherein one label is on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension and another label is on a deoxyribonucleotide triphosphate or analog thereof located in the primer portion of the primer extension product.

56. A method of determining whether a nucleic acid sequence of interest is present or absent in a sample, said method comprising the steps of:

(a) hybridizing a first oligonucleotide and a second oligonucleotide to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(b) optionally extending the first oligonucleotide;

(c) attaching the first oligonucleotide and second oligonucleotide to each other when hybridized to said target polynucleotide to generate an attached oligonucleotide combination product comprising a detectable identifying characteristic, whereby the attached oligonucleotide combination product is produced if the sequence of interest is present; and

(d) cleaving the RNA portion of the attached oligonucleotide combination product of (c), if any, with an enzyme that cleaves RNA from an RNA/DNA hybrid such that the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide,

wherein the attached oligonucleotide combination product is of a size that when the RNA is cleaved the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide under essentially the same conditions as those for attachment of the oligonucleotides, whereby detection of the cleaved attached oligonucleotide combination product comprising the detectable identifying characteristic indicates the presence of the sequence of interest.

57. The method of claim 56, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

58. The method of claim 56, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest.

59. The method of claim 56, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

60. The method of claim 56, wherein the 5' most nucleotide of the second oligonucleotide is hybridizable to at least one nucleotide of the sequence of interest.

61. The method of claim 56, wherein the sequence of interest is a single nucleotide base.

62. The method of claim 56, wherein the sequence of interest is more than a single nucleotide base.

63. The method of claim 56, wherein the first and second oligonucleotide each comprise a sequence hybridizable to at least one nucleotide of the sequence of interest.

64. The method of claim 56, wherein the enzyme that cleaves RNA is RNase H.

65. The method of claim 56, wherein the DNA polymerase lacks strand displacement activity.

66. The method of claim 56, wherein the target polynucleotide is DNA or cDNA.

67. The method of claim 56, wherein the target polynucleotide is amplified.

68. The method of claim 67, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

69. The method of claim 67, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

70. The method of claim 56, wherein the target polynucleotide is attached to an analyte.

71. The method of claim 61, wherein the analyte is selected from the group consisting of a peptide, an antibody, an organic molecule and an inorganic molecule.

72. The method of claim 56, wherein said detectable identifying characteristic is selected from the group consisting of size of the cleaved attached

oligonucleotide combination product, sequence of the cleaved attached oligonucleotide combination product, and detectable signal associated with the cleaved attached oligonucleotide combination product.

73. The method of claim 56, wherein the detectable identifying characteristic comprises the sequence of the cleaved attached oligonucleotide product, and wherein the sequence is detected by hybridizing the cleaved attached oligonucleotide product with a nucleic acid probe that is hybridizable to the cleaved attached oligonucleotide product.

74. The method of claim 73, wherein said nucleic acid probe comprises DNA.

75. The method of claim 73, wherein the nucleic acid probe is provided as a array.

76. The method of claim 75, wherein said array comprises the probe immobilized on a substrate fabricated from a material selected from the group consisting of paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, silicon, polystyrene, and optical fiber.

77. The method of claim 72, wherein said detectable signal is associated with a label on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

78. The method of claim 72, wherein said detectable signal is associated with interaction of two labels, wherein one label is located in the first oligonucleotide and a second label is located in the second oligonucleotide.

79. The method of claim 72, wherein said detectable signal is associated with interaction of two labels, wherein one label is located either the first or

second oligonucleotide, and a second label is located on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

80. The method of claim 56, wherein the attaching is covalent attachment.

81. The method of claim 80, wherein the covalent attachment is effected by ligase.

82. A method of generating multiple copies of a nucleic acid sequence of interest comprising:

incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization, primer extension and RNA cleavage, such that a primer extension product is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product from the target polynucleotide.

83. The method of claim 82, wherein said conditions that permit primer extension comprise at least one terminator deoxyribonucleotide triphosphate or analog thereof.

84. The method of claim 82, wherein said conditions that permit primer extension comprise the absence of a nucleotide that must be provided for primer extension to continue.

85. The method of claim 82, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest.

86. The method of claim 85, wherein the conditions for primer extension comprise one terminator deoxyribonucleotide triphosphate or analog thereof.

87. The method of claim 82, wherein the sequence of interest is a single nucleotide base.

88. The method of claim 82, wherein the sequence of interest is more than a single nucleotide base.

89. The method of claim 82, wherein the RNA portion of the composite primer is 5' with respect to the 3' DNA portion.

90. The method of claim 82, wherein the 5' RNA portion is adjacent to the 3' DNA portion.

91. The method of claim 82, wherein the enzyme that cleaves RNA is RNase H.

92. The method of claim 82, wherein the DNA polymerase lacks strand displacement activity.

93. The method of claim 82, wherein the target polynucleotide is DNA or cDNA.

94. The method of claim 82, wherein the target polynucleotide is amplified.

95. The method of claim 94, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

96. The method of claim 94, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

97. A method of generating multiple copies of a nucleic acid sequence of interest comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(c) optionally a DNA polymerase;

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and

(e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide,

wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product is produced, and wherein the attached oligonucleotide combination product is of a size such that cleavage of RNA from the attached oligonucleotide combination product results in dissociation of the cleaved attached oligonucleotide product.

98. The method of claim 97, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

99. The method of claim 97, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest.

100. The method of claim 97, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

101. The method of claim 97, wherein the 5' most nucleotide of the second oligonucleotide is hybridizable to at least one nucleotide of the sequence of interest.

102. The method of claim 97, wherein the sequence of interest is a single nucleotide base.

103. The method of claim 97, wherein the sequence of interest is more than a single nucleotide base.

104. The method of claim 97, wherein the first and second oligonucleotide each comprise a sequence hybridizable to at least one nucleotide of the sequence of interest.

105. The method of claim 97, wherein the enzyme that cleaves RNA is RNase H.

106. The method of claim 97, wherein the DNA polymerase lacks strand displacement activity.

107. The method of claim 97, wherein the target polynucleotide is DNA or cDNA.

108. The method of claim 97, wherein the target polynucleotide is amplified.

109. The method of claim 108, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

110. The method of claim 108, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

111. The method of claim 97, wherein the attaching is covalent.

112. The method of claim 111, wherein the covalent attachment is effected by ligase.

113. A method of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization, primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product from the target polynucleotide, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

114. The method of claim 113, wherein said conditions that permit primer extension comprise at least one terminator deoxyribonucleotide triphosphate or analog thereof.

115. The method of claim 113, wherein said conditions that permit primer extension comprise the absence of a nucleotide that must be provided for primer extension to continue.

116. The method of claim 113, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest.

117. The method of claim 113, wherein the 3' most nucleotide of the 3' DNA portion of the composite primer hybridizes 1 nucleotide from the sequence of interest, the sequence of interest is a single nucleotide, and the conditions for primer extension comprise one terminator deoxyribonucleotide triphosphate or analog thereof.

118. The method of claim 113, wherein the sequence of interest is a single nucleotide base.

119. The method of claim 113, wherein the sequence of interest is more than a single nucleotide base.

120. The method of claim 113, wherein the RNA portion of the composite primer is 5' with respect to the 3' DNA portion.

121. The method of claim 113, wherein the 5' RNA portion is adjacent to the 3' DNA portion.

122. The method of claim 113, wherein the enzyme that cleaves RNA is RNase H.

123. The method of claim 113, wherein the DNA polymerase lacks strand displacement activity.

124. The method of claim 113, wherein the target polynucleotide is DNA or cDNA.

125. The method of claim 113, wherein the target polynucleotide is amplified.

126. The method of claim 125, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

127. The method of claim 125, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

128. The method of claim 113, wherein the target polynucleotide is attached to an analyte.

129. The method of claim 128, wherein the analyte is selected from the group consisting of a polypeptide, an antibody, an organic molecule and an inorganic molecule.

130. The method of claim 113, wherein said detectable identifying characteristic is selected from the group consisting of size of the cleaved primer extension product, sequence of the cleaved primer extension product, and detectable signal associated with the cleaved primer extension product.

131. The method of claim 113, wherein the detectable identifying characteristic comprises the sequence of the cleaved primer extension product, wherein the sequence is detected by hybridizing the cleaved primer extension product with a nucleic acid probe that is hybridizable to the cleaved primer extension product.

132. The method of claim 131, wherein said nucleic acid probe comprises DNA.

133. The method of claim 131, wherein the nucleic acid probe is immobilized on a surface.

134. The method of claim 133, wherein the surface comprises a substrate fabricated from a material selected from the group consisting of paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, silicon, polystyrene, and optical fiber

135. The method of claim 130, wherein said detectable signal is associated with a label on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

136. The method of claim 130, wherein said detectable signal is associated with interaction of two labels, wherein one label is on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension and another label is on a deoxyribonucleotide triphosphate or analog thereof located in the primer portion of the primer extension product.

137. A method of determining whether a nucleic acid sequence of interest is present or absent in a sample comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second

oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(c) optionally a DNA polymerase;

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and

(e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide,

wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product comprising a detectable identifying characteristic is produced, and wherein the attached oligonucleotide combination product is of a size such cleavage of the RNA from the attached oligonucleotide combination product results in dissociationg of the cleaved attached oligonucleotide combination product from the target polynucleotide, whereby detection of the cleaved attached oligonucleotide combination product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

138. The method of claim 137, wherein the first oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

139. The method of claim 137, wherein the 3' most nucleotide of the first oligonucleotide is hybridizable to the sequence of interest.

140. The method of claim 137, wherein the second oligonucleotide comprises a sequence hybridizable to at least one nucleotide of the sequence of interest.

141. The method of claim 137, wherein the 5' most nucleotide of the second oligonucleotide is hybridizable to at least one nucleotide of the sequence of interest.

142. The method of claim 137, wherein the sequence of interest is a single nucleotide base.

143. The method of claim 137, wherein the sequence of interest is more than a single nucleotide base.

144. The method of claim 137, wherein the first and second oligonucleotide each comprise a sequence hybridizable to at least one nucleotide of the sequence of interest.

145. The method of claim 137, wherein the enzyme that cleaves RNA is RNase H.

146. The method of claim 137, wherein the DNA polymerase lacks strand displacement activity.

147. The method of claim 137, wherein the target polynucleotide is DNA or cDNA.

148. The method of claim 137, wherein the target polynucleotide is amplified.

149. The method of claim 148, wherein amplification occurs simultaneously with generating multiple copies of a sequence of interest.

150. The method of claim 148, wherein amplification occurs prior to generating multiple copies of a sequence of interest.

151. The method of claim 137, wherein the target polynucleotide is attached to an analyte.

152. The method of claim 151, wherein the analyte is selected from the group consisting of a peptide, an antibody, an organic molecule and an inorganic molecule.

153. The method of claim 137, wherein said detectable identifying characteristic is selected from the group consisting of size of the cleaved attached oligonucleotide combination product, sequence of the cleaved attached oligonucleotide combination product, and detectable signal associated with the cleaved attached oligonucleotide combination product.

154. The method of claim 137, wherein the detectable identifying characteristic comprises the sequence of the cleaved attached oligonucleotide product, and wherein the sequence is detected by hybridizing the cleaved attached oligonucleotide product with a nucleic acid probe that is hybridizable to the cleaved attached oligonucleotide product.

155. The method of claim 154, wherein said nucleic acid probe comprises DNA.

156. The method of claim 154, wherein the nucleic acid probe is immobilized on a surface.

157. The method of claim 156, wherein said surface comprises a substrate fabricated from a material selected from the group consisting of paper, glass, plastic, polypropylene, nylon, polyacrylamide, nitrocellulose, silicon, polystyrene, and optical fiber.

158. The method of claim 153, wherein said detectable signal is associated with a label on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

159. The method of claim 153, wherein said detectable signal is associated with interaction of two labels, wherein one label is located in the first oligonucleotide and a second label is located in the second oligonucleotide.

160. The method of claim 153, wherein said detectable signal is associated with interaction of two labels, wherein one label is located either the first or second oligonucleotide, and a second label is located on a deoxyribonucleotide triphosphate or analog thereof that is incorporated during primer extension.

161. The method of claim 137, wherein the attaching is covalent attachment.

162. The method of claim 161, wherein the covalent attachment is effected by ligase.

163. The method of claims 32 or 113, wherein said method comprises determining whether two or more sequences of interest are present or absent in a sample, said method comprising using one or more different composite primers, wherein the detectable identifying characteristics of the cleaved primer extension products corresponding to two different sequences of interest are different from each other.

164. The method of claim 163, wherein said method comprises at least two different composite primers.

165. The method of claim 163, wherein at least one of the sequences of interest is a variant of the sequence of interest.

166. The method of 56 or 137, wherein said method comprises determining whether two or more different sequences of interest are present or absent in a sample, said method using a two or more sets of first and second oligonucleotides, wherein the detectable identifying characteristics of the cleaved oligonucleotide attachment products corresponding to two or more different sequences of interest are different from each other.

167. The method of claim 166, wherein at least one of the sequences of interest is a variant of the sequence of interest.

168. A method for identifying an altered sequence of interest in a sample comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the altered sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization and primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that a primer extension product comprising a detectable identifying characteristic is produced, and

wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension product dissociates from the target polynucleotide, whereby the cleaved primer extension product is characterized to identify the altered sequence of interest.

169. The method of claim 168, wherein the cleaved primer extension product produced from the target polynucleotide comprises a different detectable identifying characteristic as compared with a detectable identifying characteristic of a cleaved primer extension product produced from a reference template containing the sequence of interest, wherein the characterization comprises comparing the detectable identifying characteristics.

170. A method of identifying an altered sequence of interest in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the altered sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization, and primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and

wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension product dissociates from the target polynucleotide, and

wherein production of detectably fewer cleaved primer extension products from the target as compared to the amount of cleaved primer extension products produced from a reference template comprising the sequence of interest indicates that the target polynucleotide contains an altered sequence of interest.

171. The method of claim 170, further comprising characterizing the sequence of the template polynucleotide about 1 to about 10 nucleotides 5' from the sequence hybridizable to the composite primer.

172. A method of identifying an altered sequence of interest in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the

target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(c) optionally a DNA polymerase;

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and

(e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide,

wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product comprising a detectable identifying characteristic is produced, and wherein the attached oligonucleotide combination product is of a size that when RNA is cleaved from the attached oligonucleotide combination product, the cleaved attached oligonucleotide combination product dissociates from the target polynucleotide,

wherein production of detectably fewer cleaved oligonucleotide attachment products from the target as compared to the amount of cleaved oligonucleotide attachment products produced from a reference template comprising the sequence of interest indicates that the target polynucleotide contains an altered sequence of interest.

173. The method of claim 172, further comprising characterizing sequence of the target polynucleotide.

174. A kit for generation of multiple copies of a sequence of interest, comprising a composite primer comprising a 3' DNA portion and an RNA portion, and instructions for a method for generating multiple copies, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization, primer extension and RNA cleavage, such that a primer extension product is produced, and wherein the primer extension product is of a size such that cleavage of RNA from the primer extension product results in dissociation of the cleaved primer extension product, whereby multiple copies are generated.

175. The kit of claim 174, further comprising terminator deoxyribonucletide triphosphates.

176. The kit of claim 174 or 175, further comprising at least one but not all four types of deoxyribonucleotide triphosphates.

177. A kit for determining whether a sequence of interest is present or absent in a sample, comprising a composite primer comprising a 3' DNA portion and an RNA portion, and instructions for a method of determining whether a sequence of interest is present or absent in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a composite primer that hybridizes to the target polynucleotide, said composite primer comprising an RNA portion and a 3' DNA portion, the 3' DNA portion comprising a 3' most nucleotide, such that the 3' most nucleotide of the 3' DNA portion of the primer hybridizes from about 1 nucleotide to about 10 nucleotides from the sequence of interest;

(c) a DNA polymerase; and

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid,

wherein the incubation is under conditions that permit primer hybridization, primer extension to generate a primer extension product comprising a detectable identifying characteristic, and RNA cleavage, such that the primer extension product comprising a detectable identifying characteristic is produced, and wherein the primer extension product is of a size that when RNA is cleaved from the primer extension product, the cleaved primer extension dissociates from the target polynucleotide, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

178. A kit for generation of multiple copies of a sequence of interest, comprising a first oligonucleotide and a second oligonucleotide, wherein at least

one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, and wherein the oligonucleotides singly or in combination comprise at least one nucleotide of the sequence of interest; and an agent that effects attachment of said first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide.

179. The kit of claim 178, further comprising instructions for a method of generating multiple copies of a sequence of interest, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(c) optionally a DNA polymerase;

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and

(e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide,

wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product is produced, and wherein the attached oligonucleotide combination product is of a size such that cleavage of RNA results in dissociation of the cleaved attached oligonucleotide product.

180. A kit for determining whether a sequence of interest is present or absent in a sample, comprising a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, and wherein the oligonucleotides singly or in combination comprise at least one nucleotide of the sequence of interest; and an agent that effects attachment of said first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide.

181. The kit of claim 180, further comprising instructions for a method of determining whether a sequence of interest is present or absent in a sample, said method comprising incubating a reaction mixture, said reaction mixture comprising:

(a) a target polynucleotide;

(b) a first oligonucleotide and a second oligonucleotide that hybridize to non-overlapping portions of a target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second

oligonucleotide, wherein at least one of said oligonucleotides is a composite primer comprising an RNA portion and a DNA portion, and wherein at least one of said oligonucleotides comprises a sequence that is hybridizable to at least one nucleotide of the sequence of interest;

(c) optionally a DNA polymerase;

(d) an enzyme that cleaves RNA from an RNA/DNA hybrid; and

(e) an agent that effects attachment of the first oligonucleotide and second oligonucleotide to each other when said oligonucleotides are hybridized to the target polynucleotide,

wherein the incubation is under conditions that permit oligonucleotide hybridization, optionally oligonucleotide extension, RNA cleavage and attachment of the first oligonucleotide and the second oligonucleotide, such that an attached oligonucleotide combination product comprising a detectable identifying characteristic is produced, and wherein the attached oligonucleotide combination product is of a size such that cleavage of RNA from the attached oligonucleotide combination product results in dissociation of the cleaved attached oligonucleotide product from the target polynucleotide, whereby detection of the cleaved primer extension product comprising the detectable identifying characteristic indicates presence of the nucleotide sequence of interest.

182. The kit of claim 180 or 181, wherein the agent that effects attachment of said first oligonucleotide and second oligonucleotide is ligase.

183. The kit of claim 180, further comprising an enzyme that cleaves RNA from a RNA/DNA hybrid.

184. The kit of claim 183, wherein said enzyme is RNase H.

185. A reaction mixture comprising (a) a target polynucleotide; (b) a first oligonucleotide and a second oligonucleotide, wherein at least one oligonucleotide is a composite primer, wherein the two oligonucleotides hybridize to non-overlapping portions of the target polynucleotide, wherein the portion of the target polynucleotide that is hybridizable to the first oligonucleotide is 3' with respect to the portion of the target nucleotide that is hybridizable to the second oligonucleotide, and wherein the oligonucleotides singly comprise at least one nucleotide of the sequence of interest, or in combination comprise at least a portion of the sequence of interest; (c) optionally DNA polymerase; and (d) an agent that effects covalent attachment of the first oligonucleotide and second oligonucleotide to each other when the oligonucleotides are hybridized to the target polynucleotide.

Figure 1

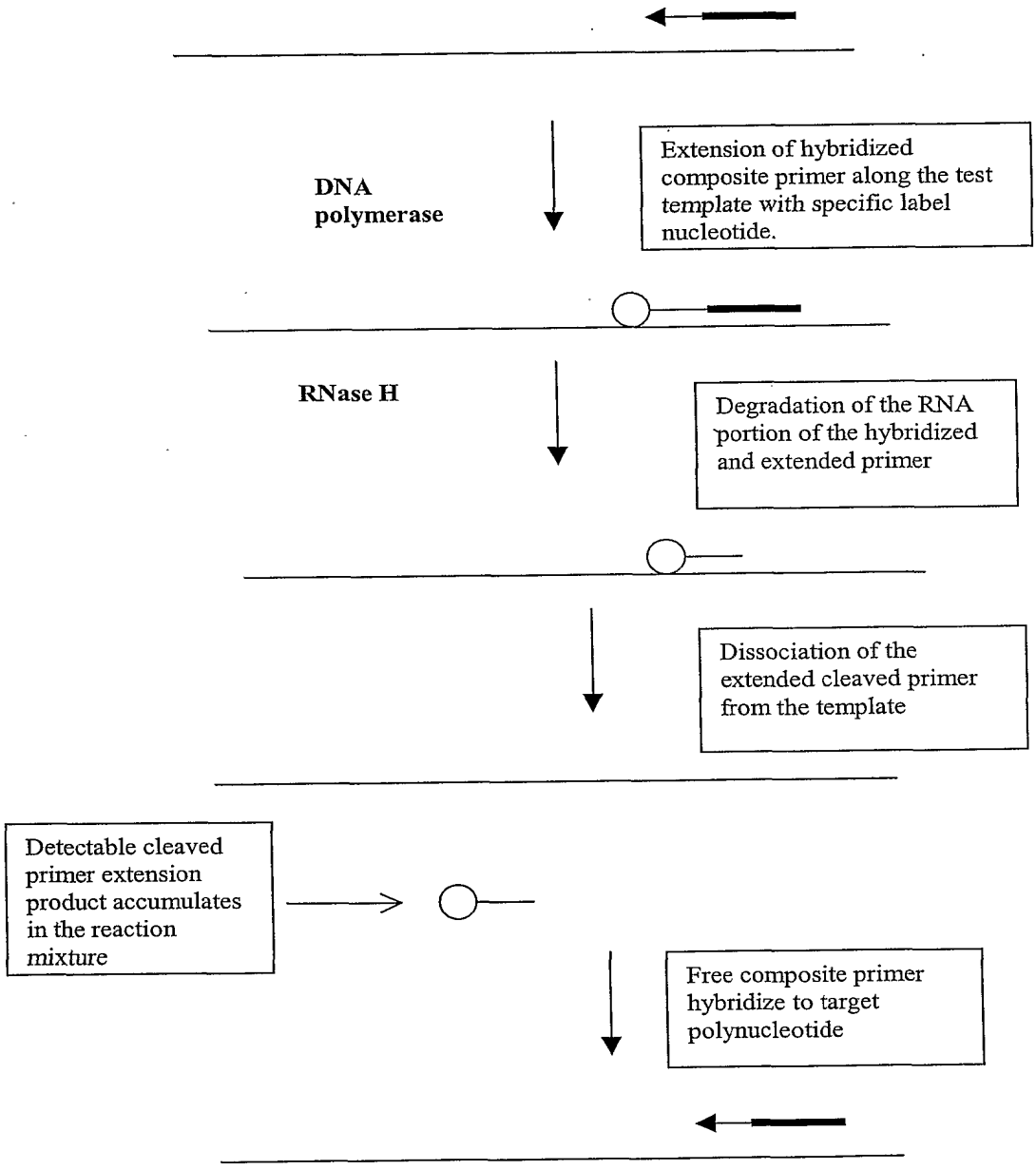

DNA polymerase
Extension of hybridized composite primer along the test template with specific label nucleotide.
RNase H
Degradation of the RNA portion of the hybridized and extended primer
Dissociation of the extended cleaved primer from the template
Detectable cleaved primer extension product accumulates in the reaction mixture
Free composite primer hybridize to target polynucleotide

Figure 2

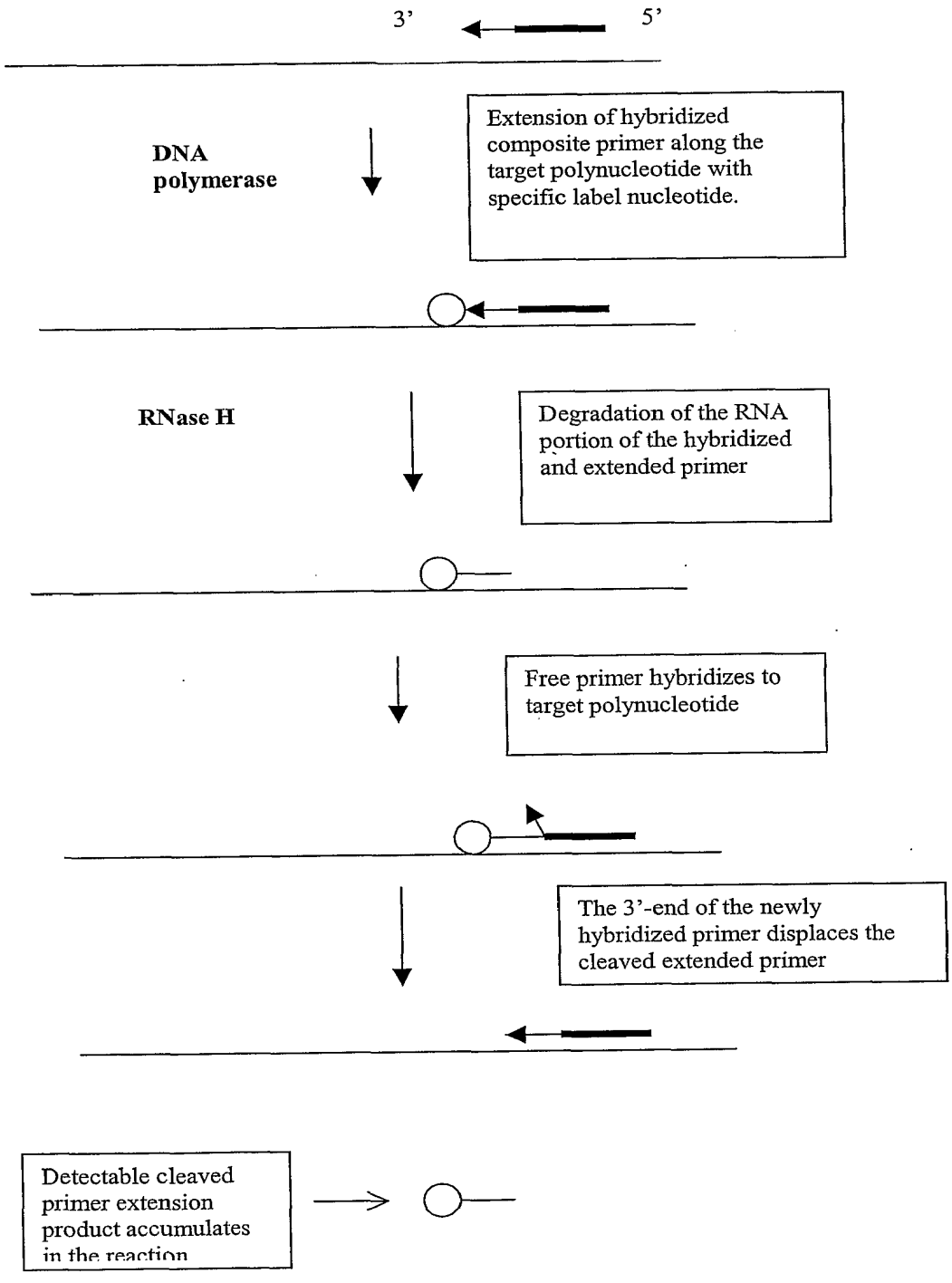
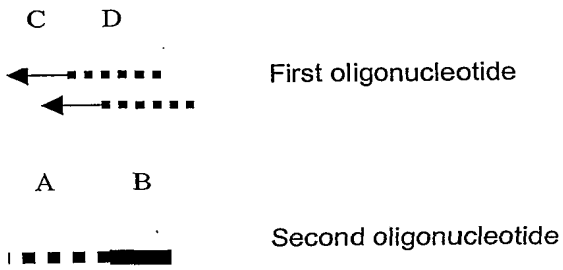

3'
5'
DNA polymerase
Extension of hybridized composite primer along the target polynucleotide with specific label nucleotide.
RNase H
Degradation of the RNA portion of the hybridized and extended primer
Free primer hybridizes to target polynucleotide
The 3'-end of the newly hybridized primer displaces the cleaved extended primer
Detectable cleaved primer extension product accumulates in the reaction

Figure 3



Formation of first target oligonucleotide complex

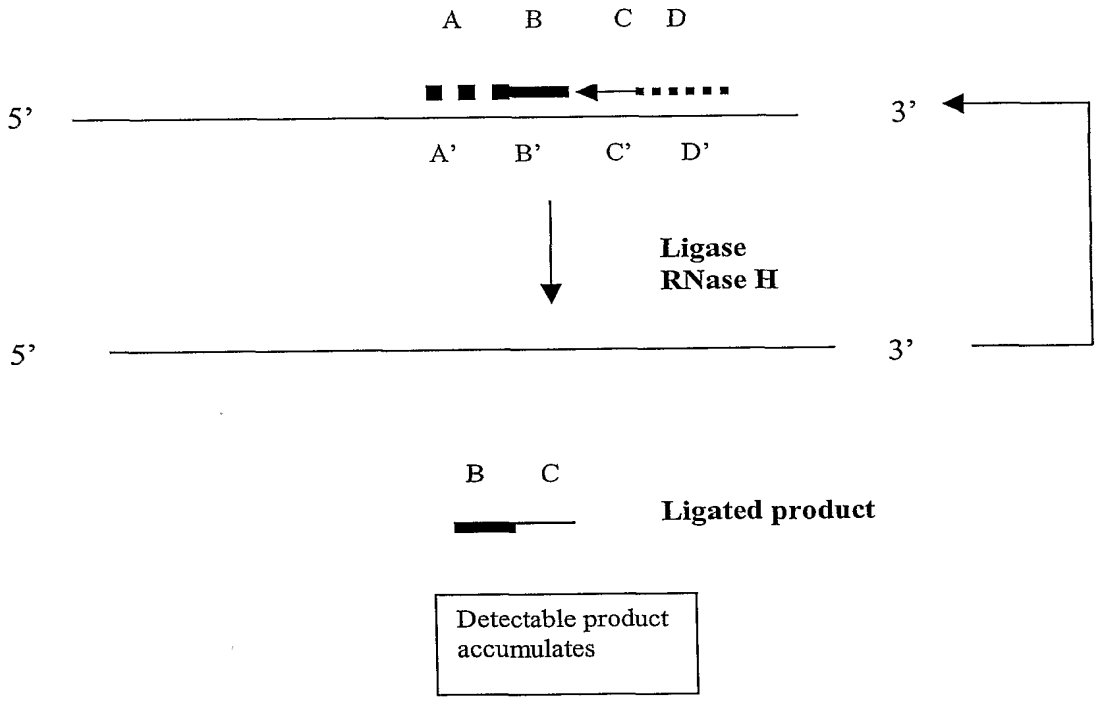


Figure 4

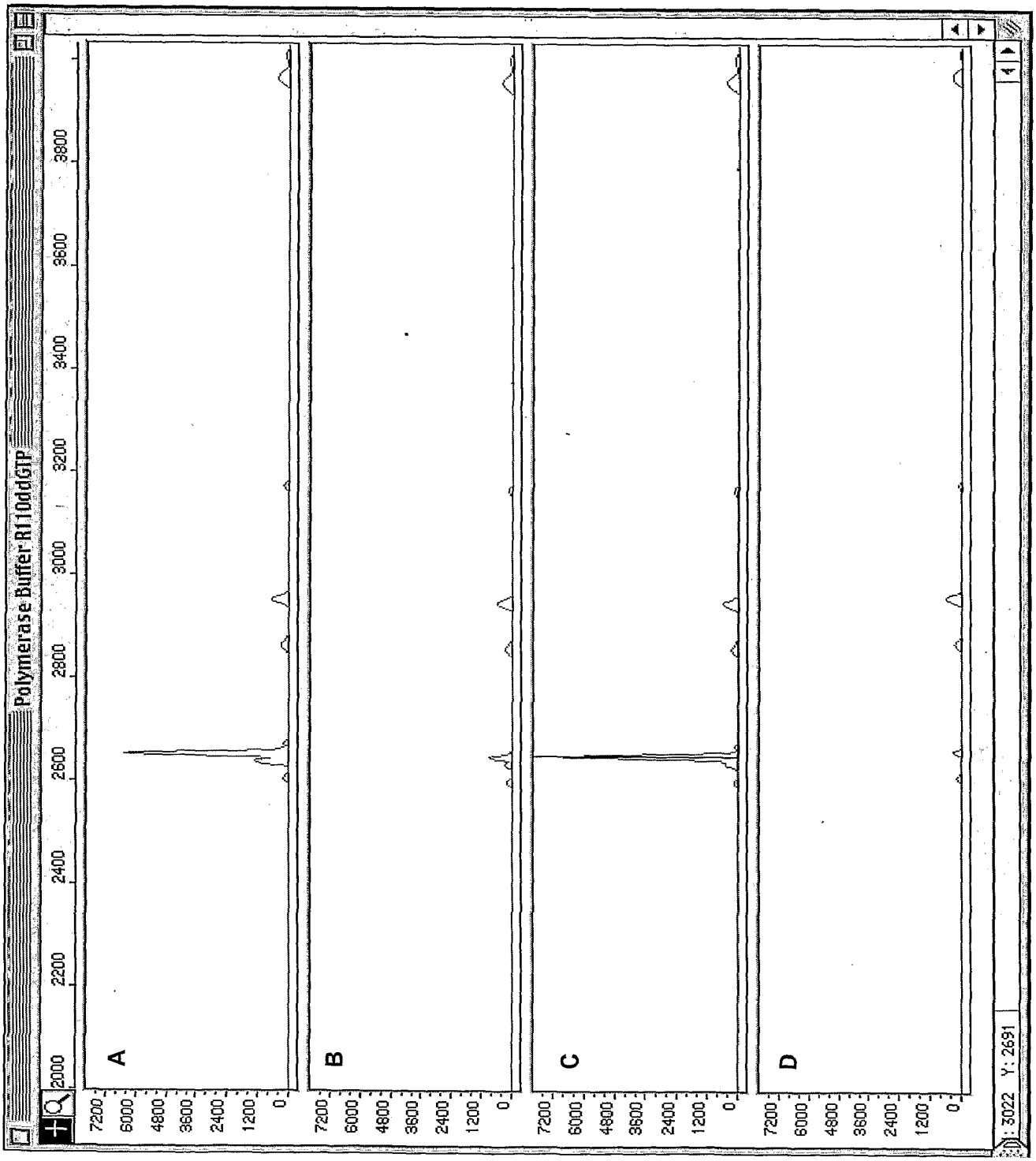
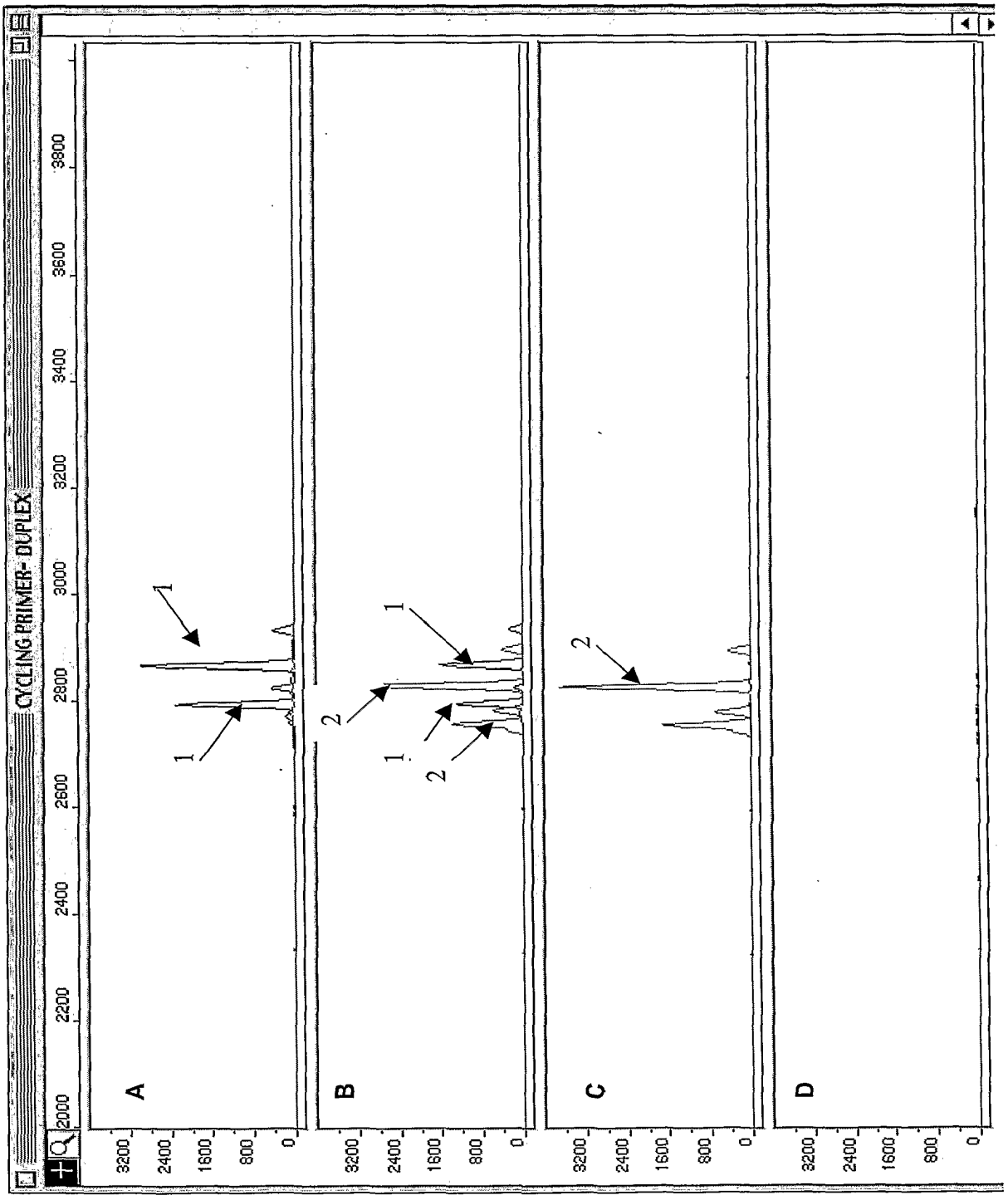

Polymerase Buffer R110ddGTP
A
B
C
D
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
7200
6000
4800
3600
2400
1200
0
X: 3022 Y: 2691

FIGURE 5


CYCLING PRIMER- DUPLEX
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
A
B
C
D
3200
2400
1600
800
0
1
2

Espacenet

Bibliographic data: CN1661102 (A) — 2005-08-31

Method for detecting sequence of double chain DNA based on procedure of DNA automaton

Inventor(s): ZHANG ZHIZHOU [CN]; HU JUN [CN]; HE LIN [CN] ± (ZHIZHOU ZHANG, ; JUN HU, ; LIN HE, ; ZHANG ZHIZHOU, ; HU JUN, ; HE LIN)

Applicant(s): UNIV SHANGHAI JIAOTONG [CN] ± (SHANGHAI JIAOTONG UNIV)

Classification: **- international:** ***C12Q1/68;*** (IPC1-7): C12Q1/68
- cooperative:

Application number: CN20041093095 20041216

Priority number (s): CN20041093095 20041216

Also published as: CN100413978 (C)

Abstract of CN1661102 (A)

A method based on DNA robot process for testing the double-strained DNA sequence includes such steps as using the double-strained DNA sequence as input molecular, severing it by restriction endonuclease to expose adhesive end, linking with state transition molecular via base, cyclic severing and linking to gradually shorten the sequence, and reading the variation of length to reflet the information about DNA sequence.

IDT Exhibit 1020.7058

Patent Translate
Powered by EPO and Google

Notice

This translation is machine-generated. It cannot be guaranteed that it is intelligible, accurate, complete, reliable or fit for specific purposes. Critical decisions, such as commercially relevant or financial decisions, should not be based on machine-translation output.

DESCRIPTION CN100413978C

Abstract A method for determining double-stranded DNA sequence based on DNA automata process. In the present invention, the double-stranded DNA sequence is used as an input molecule, and the DNA is cleaved by a restriction enzyme to expose a sticky end, which is combined with the state transfer molecule through base pairing and connected. Since the state transfer molecule still contains the recognition site of the restriction endonuclease, the ligated DNA molecule continues a new round of cutting and ligation cycle and gradually shortens the DNA sequence as the input molecule and gradually changes the DNA sequence as the input molecule Short, the change in the length of the input molecule reflects the information of the DNA sequence. The change in the length is readable. By designing the structure of the state transfer molecule, the speed of the length of the input molecule can be reduced by 1 base at a time. The invention effectively avoids the sequencing difficulties caused by the high-level structure of single-stranded DNA molecules, overcomes the shortcomings of the secondary structure of single-strand template in the prior art that affects sequencing, and reduces the cost on this basis, and has great application prospects.

Method for determining double-stranded DNA sequence based on DNA cutting process

Technical field

The invention relates to a new method for DNA sequence determination in the field of biotechnology, in particular to a method for double-stranded DNA sequence determination based on a DNA cleaving process.

Background technique

Since the publication of the human genome sequence, the world has entered the post-genomic era. Thousands of biological species require genome sequencing. Among the current DNA sequencing solutions, Sanger's dideoxy method is the most popular. At present, DNA sequencing has a number of international patented technologies {such as US patents, European patents and Japanese patents}, except for a few methods that are not very practical. Most of them are patents implemented on a series of technical issues around the dideoxy process. The human genome sequence is determined by the dideoxy method. However, all the large genomes currently published generally have "holes", that is, parts of the sequence that have not been completed. One of the reasons is that the existing sequencing technology is difficult to complete the final determination of certain DNA sequences such as highly repetitive sequences or sequences rich in secondary structure.

After searching the existing technical literature, it was found that Shapiro et al. of the Weizmann Institute in Israel published a DNA reaction process similar to the Turing machine in the journal Nature (2001, 414, 430-434). A continuous process of DNA cutting. We found that during the DNA cleavage process, the double-stranded DNA as the input molecule is continuously cleaved by the state-switching molecules under the circulation of restriction enzymes and DNA ligases. In this process, the sequence of the double-stranded DNA molecule The length changes from long to short. Because the sequence of double-stranded DNA molecules input during DNA cutting gradually becomes shorter, its length change can be tracked in real time, and the way of length change can be programmed completely, or even designed to be shortened by 1 base each time, so it constitutes A complete DNA double-strand sequencing solution. This scheme is completely different from the technology described by the dideoxy method.

Summary of the invention

IDT Exhibit 1020.7059

The purpose of the present invention is to overcome the shortcomings in the prior art and provide a method for determining the double-stranded DNA sequence based on the DNA cleaving process. It effectively avoids the sequencing difficulties caused by the high-level structure of single-stranded DNA molecules and reduces costs on this basis. The present invention can not only replace the existing technology in the technical route, but also overcome the defects of the existing technology and have potential application prospects. .

The present invention is achieved by the following technical solution, the specific method is as follows: double-stranded DNA sequence is used as an input molecule, the DNA is cleaved by a restriction endonuclease, the sticky end is exposed, and the state transition molecule is combined through base pairing and connected Ligation occurs under the action of an enzyme; since the state-transition molecule still contains the recognition site of the restriction enzyme, the ligated DNA molecule continues a new round of cutting and ligation cycles and gradually shortens the DNA sequence used as the input molecule as an input The DNA sequence of the molecule gradually becomes shorter, and the length change of the input molecule reflects the information of the DNA sequence. By designing the structure of the state transition molecule, the speed of the length of the input molecule can be reduced by 1 base at a time. Because it is based on double-stranded molecules, it can effectively avoid sequencing difficulties caused by the high-level structure of single-stranded DNA molecules.

The input molecule is double-stranded DNA, or double-stranded converted from single-stranded nucleic acid molecules.

A DNA ligase based on double-strand as a template for single-strand connection or a connection system capable of directly connecting single-strand DNA.

The restriction enzyme is a type II restriction enzyme, and its recognition site has a certain distance from the cleavage site, but it has a high degree of cleavage specificity; the end of the cut DNA molecule is a sticky end. The length is 2-5 bases, and the sticky end can be either 5'sequence or 3'sequence. Two or more different types of restriction enzymes can also be used simultaneously.

Four-color fluorescent labeling of state transition molecules using small molecule fluorescent labels. The state-switching molecule is covalently linked to the gradually shortened input molecule under the action of a ligase, and the shortened input molecule is identified by a fluorescent group.

The basic structure of the state-switching molecule used in the DNA cleavage reaction is: a double-stranded DNA molecule composed of two DNA single strands of different lengths, with an upper strand length of 19-38 bases and a lower strand length of 21-43 bases. The 3'sequence of the upper strand and the lower strand are completely paired, so that the 5'end of the lower strand has 2-5 bases as the sticky end for state transition; the length of the sticky end (n) determines the state to be synthesized The total number of converted molecules is 4<n>. The completely matched region contains the recognition sites of the above 3 enzymes. The recognition site has 6-25 bases from the 3'end of the lower strand. The corresponding double strand in this segment is the protective sequence; the recognition site and the sticky end The double-stranded sequence in between is a spacer sequence, and the length of the spacer sequence is designed so that only one base can be moved during a state transition during DNA cutting. The 3'end of the lower strand is a fluorescent labeling site; according to the 3'end of the lower strand being A, G, C or T, four different fluorescent labels or any one of the physical and chemical labels that can be used for detection are selected.

The components of the DNA cleavage reaction system include input molecules, state transition molecules, ligase, endonuclease, ATP, and buffer. The reaction system includes both endonuclease and ligase. The DNA cleaving reaction process also includes: 1) a reaction system composed of enzymes without protein properties, but ribozymes or deoxyribozymes are used, in which ribozymes or deoxyribozymes can replace restriction enzymes to drive DNA molecules to proceed Cleavage reaction, but the system also contains DNA ligase; 2) a reaction system composed of a combination of biological molecules through a physicochemical process, using double-stranded DNA molecules as input molecules to perform a similar cleavage process and reflect DNA sequence information. In the above reaction system, the DNA sequence information can be detected by directly observing the change in the length of the connection product between the double-stranded DNA input molecule and the state transition molecule.

Sequence length changes can be read by capillary electrophoresis.

Compared with the currently popular DNA sequence dideoxy determination method, the fluorescent label used in the method described in the present invention is exactly the same. But a set of state-switching molecules designed is universal for any DNA sequence, and the dideoxy method often needs to design corresponding primers for specific DNA sequences. In addition, the present invention uses double-stranded DNA as a template to overcome the secondary structure problem caused by the double-deoxygenation method using single-stranded DNA as a template, so the present invention obviously has technical advantages over the prior art.

IDT Exhibit 1020.7060

The following further describes the method of the present invention in detail:

(1)

The unknown DNA sequence is converted into a double strand and connected to a leader molecule, which contains a recognition site specified by an endonuclease. The leader molecule is about 10 to 20 bp (base pairs) in length, and can be a multiple cloning site on a general subcloning vector. The PCR (polymerase chain reaction) product can be linked to the leader molecule, or a primer-specific recognition site can be designed on the primer. The unknown DNA sequence may also be a double-stranded DNA converted from a single-stranded nucleic acid (DNA or RNA) molecule.

(2)

The endonuclease, FokI is one of the enzymes, BbsI, HgaI or BsMFI can also be used. The unknown sequence is digested with these enzymes before the determination to ensure that the part of the sequence to be tested does not contain a restriction site.

(3)

Design of state transition molecules. The sticky end exposed by the restriction enzyme of FokI is 4 bases, so the number of state transition molecules required is 4<4>, that is, 256 kinds. If the endonuclease used cuts double-stranded DNA and the sticky end is 3 bases or other numbers, the number of corresponding state transition molecules needs to be adjusted accordingly. The structure of the state transition molecule is shown in Figure 2. The spacer sequence allows only one base to move during a state transition during DNA cutting. The 8-base-pair sequences in all 256 state transition molecules can be all the same or different. The exposed sticky end requires 5'phosphorylation to facilitate ligation (in a ligase-free system 5'phosphorylation does not have much effect on the reaction rate), all state transition molecules can be synthesized at once. The fluorescent molecules of the state transition molecules are labeled at the 3'end of the lower strand. In order to prevent the self-ligation of state-switching molecules and ensure that the state-switching molecules can be efficiently recognized and cleaved by restriction enzymes, the left protection sequence can be specially designed similar to that shown in Figure 1, that is, the double-stranded end of the protection sequence is not necessarily It is a blunt end, or you can make the 3'end of the lower strand a few more bases than the upper strand. System description The state-switching molecule of the DNA cleavage reaction system for DNA sequencing is: a double-stranded DNA molecule composed of two single-stranded DNAs of different lengths but completely matched, with an upper strand length of 19-38 bases and a lower strand length of 21 --43 bases. The 3'end sequence of the upper strand and the lower strand are completely matched, so that the 5'end of the lower strand has 2-5 bases as sticky ends for state transition; the completely matched region contains the recognition sites of the above 3 enzymes , The recognition site is 6-25 bases from the 3'end of the lower strand, the corresponding double-strand is the protective sequence; the double-stranded sequence between the recognition site and the sticky end is the spacer sequence, the length of the spacer sequence Designed so that only one base is moved during a state transition during DNA cutting. The 3'end of the lower strand is a fluorescent labeling site; according to the 3'end of the lower strand being A, G, C or T, four different fluorescent labels or any one of the physical and chemical labels that can be used for detection are selected.

(4)

Selection of fluorescent markers. Small molecule fluorescent markers. You can refer to Applied BioSystems' fluorescent labels. According to the last base of the sticky end of the base in the state transition molecule, the different fluorescent label groups are specified, such as A label FAM, T label HEX, C label CY3, G label CY5, etc., select the labeling method according to the requirements of the sequence reading instrument . The sequence reading instrument will generate a four-color fluorescence map reflecting the changes in the length of the DNA double-stranded molecules.

(5)

Ligase selection. Generally, a DNA ligase that performs single-strand ligation based on double-strands as a template, or a DNA ligase or ligation system that can directly perform single-strand ligation are possible.

(6)

The general system of DNA cleavage reaction: the components include double-stranded DNA input molecule, state transition molecule, DNA ligase, type II restriction enzyme, ATP, buffer, and the reaction temperature of DNA cleavage reaction is 16-25°

IDT Exhibit 1020.7061

C, The respective concentrations are: 0.374 micromoles/liter for each state transition molecule, FokI: 0.2 units/microliter, ligase 18.7 units/microliter, ATP 0.8 millimoles/liter, input molecule DNA 0.15 micromoles/liter. Reaction mother liquor: FokI: 4 units/microliter, ligase 400 units/microliter, ATP 100 mmol/liter. Under the above conditions, the optimization of the reaction system can be continued.

(7)

How to read the sequence length. The core of the present invention lies in a completely new DNA sequencing reaction system, that is, a DNA cleavage reaction system; as for observing the change in sequence length, any existing technology can be used, such as capillary electrophoresis technology (such as ABI3100/MegaBase sequencer). In the general system of DNA cleavage reaction, DNA fragments with different lengths can be identified by DNA sequence using capillary electrophoresis technology and fluorescent reading system. Need to know the sequence of the first 3 bases of the tested sequence and the four-color map reflecting the change in the length of the DNA double-stranded molecule, then all subsequent sequences can be extrapolated. For example, if you already know that the first 3 bases are AGT, the subsequent The first base is completely determined by the fluorophore, such as A labeled FAM, T labeled HEX, C labeled CY3, G labeled CY5, and the base type can be obtained according to the fluorescent color. In other reaction systems for DNA cleavage, the DNA sequence information can be identified by directly detecting the change in the length of the double-stranded DNA input molecule.

BRIEF DESCRIPTION

Figure 1 Schematic diagram of the DNA cutting process.

In the figure: the numbers in the box (7, 21, 22, 15, 311) represent the length of the double-stranded DNA sequence in the box, T1, T4 and T6 are three examples of state-switching molecules, and the input molecule represents the double-stranded DNA The initial state in which the molecule begins to cut DNA. By designing the length of the spacer sequence in the state transition molecule, the number of bases that advance each time during the cutting process can be controlled.

Figure 2 Schematic diagram of state transition molecules used in the FokI reaction system.

In the picture: the black ellipsoid on the left indicates the fluorescent label, which is marked on the lower strand (3'), GGATC is the FokI recognition site, 6-25bp is the protection sequence, NNNNNNNN is the 8-base spacer sequence, and XXXX is all possible base combinations , P is the lower strand 5'phosphate label, which requires a total of 256 state transition molecules.

detailed description

Example 1

Determine the DNA sequence information of the unknown sequence that has been inserted into the multiple cloning site of the plasmid. Using DNA recombination technology to insert the sequence to be tested into the plasmid, the multi-cloning site should contain a recognition site for the restriction enzyme for DNA cleavage reaction such as FokI; the known sequence of the multi-cloning site is the start of the determination sequence. By designing state-switching molecules so that each DNA cut moves a distance of 1 base, the sequence of DNA input molecules can be directly read. If FokI is used, the exposed sticky ends are 4 bases, so the number of all state transition molecules to be consklered is 4<4>=256. Only need to know the first 3 bases of the measured sequence (the known sequence of the multi-cloning site) can infer all the subsequent sequences according to the fluorescence color of the state-switching molecule. Before the reaction, the plasmid containing the unknown sequence was cut with FokI to determine whether the unknown sequence contained another FokI recognition site. If there is no additional FokI recognition site, the sequencing reaction is performed directly, that is, the entire plasmid is added to the DNA cleavage reaction system as an input molecule for DNA cleavage; if there are additional FokI recognition sites, the scheme in Best Example 3 is followed. .

Examples of input molecules used in experimental research (where NNNN represents any sequence of unlimited length):

FokI

gtcactatggcgtgctgcta tgtgatc ctggctcgcagcctggcttgtcgc NNNN

cagtgataccgcacgacgat acactaggaccgagcgtcggaccgaacagcgNNNN

Examples of state transition molecules T1:

acgttgcatacggactgctgat tgtgatc c

FAM-tgcaacgtatgcctgacgacta acactaggccga-p

IDT Exhibit 1020.7062

Add input molecules, T1, FokI and T4DNA into the reaction system. The input molecule is digested by FokI:

ggctgcagcctggcttgtcgc NNNN

gcgtcggaccgaacagcgNNNN

T1 with ccga sticky end is connected under the action of T4 DNA and Ligase, the product is denoted as molecule M:

acgttgcatacggactgctgat tgtgatc cggctgcagcctggcttgtcgc NNNNFAM-tgcaacgtatgcctgacgacta acactaggccgag Observe directly. Due to the presence of FoKI, the molecule M will continue to be cut into two fragments:

gctgcagcctggcttgtcgc NNNN

cgtcggaccgaacagcgNNNN

和

acgttgcatacggactgctgat tgtgatc cg FAM-tgcaacgtatgcctgacgacta acactaggccgag

Examples of state transition molecules T2:

acgttgcatacggactgctgat tgtgatc c FAM-tgcaacgtatgcctgacgacta acactaggccgag-pT2 can continue to connect with the input molecule, and then Be cut. As can be seen from the above reaction, the input molecule moves to the right one base at a time. The above partial sequence of the input molecule containing the recognition site of FoKI can be designed as part of the sequence of the multiple cloning site. According to the existing experimental results, the degree of sequence that can be read using DNA cleavage reaction for DNA sequence determination will reach about 500 bases under unoptimized reaction conditions, which is close to the current good level of dideoxy method.

Example 2

The sequence of the double-stranded DNA of the PCR product was determined. It is necessary to design a FokI site in the PCR primer, so that one end of the double-stranded DNA molecule obtained by PCR amplification (the specific method of the PCR reaction is not limited) contains a FokI site, which is the most distant from the double-stranded DNA molecule. There is a distance of 6-10 bases at the end to ensure that this site can be effectively cleaved by FokI. In this way, the DNA cleavage reaction system corresponding to FokI can be used to sequence the PCR products. Before the reaction, the double-stranded DNA sequence amplified by PCR is cut with FokI to determine whether the sequence contains additional FokI recognition sites. If there is no additional FokI recognition site, the sequencing reaction is performed directly, that is, the entire PCR product double-stranded DNA is added to the DNA cleavage reaction system as an input molecule for DNA cleavage; if there are additional FokI recognition sites, then according to the best embodiment 3 in the program.

Determining the sequence of double-stranded DNA of PCR products accounts for a considerable proportion of all DNA sequence determination work, but it is often difficult to sequence PCR products using the existing dideoxy method. One of the reasons is that there are many intermediate products or By-products, they are not easy to clean everywhere, and not going clean will seriously interfere with the specificity of the sequencing reaction. Going clean will increase the cost of sequencing. Using the DNA cleaving method of the present invention will increase the specificity of the sequencing reaction because the endonuclease only recognizes and cleaves double-stranded DNA.

Example 3

For unknown sequences that already contain FokI recognition sites, the following strategy can be used. The sequence was cut with FokI, and the cut fragments were recovered and directly sequenced. There are many ways to recover the fragments. The products of the FokI cleavage reaction can be separated by agarose gel according to the size of the fragments, and each fragment can be purified by tapping. These recovery and purification processes use standard DNA recombination technology, and you can choose any kit on the market. High-pressure liquid phase systems can also be used for purification. The purified DNA fragment can be directly subjected to a DNA cleavage reaction to determine the sequence, that is, the purified DNA double-stranded fragment is added to the DNA cleavage reaction system as an input molecule for DNA cleavage. Finally, the sequences of these fragments are spliced together. If FokI is not suitable, you can use other restriction enzymes and design the corresponding reaction system.

IDT Exhibit 1020.7063

Patent Translate
Powered by EPO and Google

Notice

This translation is machine-generated. It cannot be guaranteed that it is intelligible, accurate, complete, reliable or fit for specific purposes. Critical decisions, such as commercially relevant or financial decisions, should not be based on machine-translation output.

CLAIMS CN100413978C

1.

A method for determining the double-stranded DNA sequence based on the DNA cleavage process, characterized in that the double-stranded DNA sequence is used as an input molecule, and the type II restriction enzyme cuts the DNA to expose the sticky end, and then passes through the state transition molecule Base pairing is combined, and ligation occurs under the action of ligase; the end of the DNA molecule cleaved by type II restriction enzyme is a sticky end, with a length of 2-5 bases, and the sticky end is a 5'sequence or 3 'Sequence, use two or more restriction enzymes to cut at the same time, because the state transition molecule still contains the recognition site of the restriction enzyme, the ligated DNA molecule continues a new round of cutting and ligation cycle The DNA sequence as the input molecule is gradually shortened, and the change in the length of the input molecule reflects the information of the DNA sequence. The way to reduce the length of the input molecule by designing the structure of the state transition molecule is 1 base at a time, thereby achieving DNA Unique determination scheme of sequence in double-stranded state.

2.

The method for determining a double-stranded DNA sequence based on a DNA cleaving process according to claim 1, characterized in that a small molecule fluorescent marker is used to perform four-color fluorescent labeling on the state-switching molecule, and the state-switching molecule is linked to the enzyme Under the action, it covalently connects with the gradually shortening input molecules, and uses fluorescent groups to identify the shortening input molecules.

3.

The method for determining a double-stranded DNA sequence based on a DNA cleaving process according to claim 1, wherein the basic structure of the state transition molecule is: a double-stranded DNA molecule composed of two DNA single strands of different lengths, on The chain length is 19-38 bases and the lower chain length is 21-43 bases. The upper strand and the lower strand are completely matched at the 3'end sequence, so that there are 2-5 bases at the 5'end of the lower strand. The sticky end is used for state transition; the length n of the sticky end determines the total number of state-switching molecules to be synthesized 4n. The completely matched region contains the recognition site of the above endonuclease, and the recognition site is 3'from the lower strand. There are 6-25 bases, the corresponding double-stranded sequence is the protection sequence; the double-stranded sequence between the recognition site and the sticky end is a spacer sequence, and the length of the spacer sequence is designed so that only one state transition moves during DNA cleavage One base, the 3'end of the lower strand is a fluorescent labeling site; according to the 3'end of the lower strand is A, G, C or T to select four different fluorescent labels or any one can be used for detection For physicochemical labeling, the 5'end of the lower strand should be phosphorylated.

IDT Exhibit 1020.7064

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 329 822 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **88113948.9**

(51) Int. Cl.[4]: **C12Q 1/68 , C07H 21/00**

(22) Date of filing: **26.08.88**

(30) Priority: **24.02.88 CA 559709**

(43) Date of publication of application:
30.08.89 Bulletin 89/35

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: **Cangene Corporation**
3403 American Drive
Mississauga Ontario L4V 1T4(CA)

(72) Inventor: **Davey, Cheryl**
175 Dinnick Crescent
Toronto Ontario M4N 1M4(CA)
Inventor: **Malek, Lawrence T.**
4 Sproule Drive
Brampton Ontario L6V 4A5(CA)

(74) Representative: **Kolb, Helga, Dr. Dipl.-Chem. et al**
Hoffmann, Eitle & Partner Patentanwälte
Arabellastrasse 4
D-8000 München 81(DE)

(54) Nucleic acid amplification process.

(57) This invention relates to a process for amplifying a specific nucleic acid sequence. The process involves synthesizing single-stranded RNA, single-stranded DNA and double-stranded DNA. The single-stranded RNA is a first template for a first primer, the single-stranded DNA is a second template for a second primer, and the double stranded DNA is a third template for synthesis of a plurality of copies of the first template. A sequence of the first primer or the second primer is complementary to a sequence of the specific nucleic acid and a sequence of the first primer or the second primer is homologous to a sequence of the specific nucleic acid. The amplification process may be used to increase the quantity of the specific nucleic acid sequence to allow detection, or to increase the purity of the specific nucleic acid sequence as a substitute for conventional cloning methodology.

EP 0 329 822 A2

IDT Exhibit 1020.7065

NUCLEIC ACID AMPLIFICATION PROCESS

FIELD OF THE INVENTION

This invention relates to a process for amplifying a specific nucleic acid sequence.

BACKGROUND OF THE INVENTION

The detection of a specific nucleic acid sequence present in a sample by probing the sample with a complementary sequence of nucleic acids is a known diagnostic technique. Nucleic acids are highly specific in binding to complementary nucleic acids and are thus useful to determine whether a specific nucleic acid is present in a sample. One must know the sequence of the specific nucleic acid to be detected and then construct a probe having a complementary nucleic acid sequence to the specific nucleic acid sequence.

In this application, the phrase "specific nucleic acid sequence" means a single stranded or double stranded nucleic acid which one wishes to amplify; "sample" means a mixture containing nucleic acids; "sufficiently complementary" means that two nucleic acids, a primer and a template, are capable of specific interaction which allows efficient, primer-dependent and template-directed synthesis of DNA, under given conditions of ionic strength and temperature.

Since nucleic acid probes are highly specific, it is preferable in some situations to probe the nucleic acid sequence itself rather than the protein produced by the nucleic acid sequence. As a particular example, a diagnostic method based solely on protein detection would be unreliable for determining the presence of infectious particles of hepatitis B virus, due to the presence of significant levels of non-infectious antigen particles which lack the DNA genome. In another example, the various subtypes of human papilloma virus found in either pre-cancerous or benign cervical tumors can be distinguished only by the use of nucleic acid probe hybridization. Also, the microbiology of AIDS makes it certain that an assay based on the presence of AIDS specific nucleic acid sequence would be superior as a diagnostic.

The greatest difficulty with applying existing nucleic acid probe technology, and the reason that the utility of the existing probe technology is limited, is the copy number problem. In a virus or cell, for example, there is usually a single copy of a particular gene. This one copy may give rise to many copies of gene product, either RNA or protein. For this reason, diagnostic techniques have often involved probing the protein, since the specific sequence of nucleic acid to be detected may give rise to many thousand copies of protein.

The naturally-occurring high number of ribosomal RNA, up to 100,000 copies per cell, has been used by GenProbe to facilitate diagnosis of certain bacterial pathogens, such as Legionella and Mycoplasma, using nucleic acid probes. However, this strategy cannot be used with non-cellular pathogens, such as viruses. Copy number is a particular problem with the development of a nucleic acid probe method for the detection of AIDS virus, where the integrated provirus may be present in less than one of ten thousand peripheral blood lymphocytes. Thus, if the particular nucleic acid sequence suspected to be present in a sample could be amplified, the copy number problem could be circumvented and probe assays could be more readily used.

In a normal biological sample, containing only a few cells, and consequently only a few copies of a particular gene, it is necessary to utilize an amplification process in order to overcome the copy number problem.

One method to amplify is to 'grow out' the sample, that is, to arrange conditions so that the living biological material present in the sample can replicate itself. Replication increases the quantity of nucleic acid sequences to detectable levels. In the food industry, for example, in order to test processed food for the food-poisoning bacteria Salmonella, food samples must be incubated for a number of days to increase the quantity of nucleic acids. In clinical samples, pathogens must also be allowed to increase their number by growing out over some considerable time.

United States Patent No. 4,683,195 issued on July 28, 1987 to Cetus Corporation and United States Patent No. 4,683,202 issued on July 28, 1987 to Cetus Corporation are each directed to a process for amplifying a target nucleic acid sequence contained in a sample. United States Patent No. 4,683,195 relates to a process in which a sample suspected of containing a target nucleic acid sequence is treated with oligonucleotide primers such that a primer extension product is synthesized which in turn serves as a template, resulting in amplification of the target nucleic acid sequence. The primer extension product is

IDT Exhibit 1020.7066

separated from the template in the preferred embodiment using heat denaturation. Similarly, United States Patent No. 4,683,202 relates to a process for amplifying a target nucleic acid sequence having two separate complementary strands. The process includes treating the strands with primers to synthesize extension products, separating the primer extension products from the templates, and in turn using the primer extension products as templates.

Both of the above United States patents require either manual or mechanical participation and multi-step operations by the user in the amplification process. The steps involved in these patents require the user to heat the sample, cool the sample, add appropriate enzymes and then repeat the steps. The temperature changes cause the enzymes to loose their activity. Hence, the user is required to repeatedly supplement the amplification mixture with aliquots of appropriate enzymes during the amplification process.

In addition, in United States Patents Nos. 4,683,195 and 4,683,202 each cycle of the amplification process takes place by the synthesis from a first template, of a second template, the second template in turn is used to synthesize the first template. This procedure is repeated, thus, each cycle of the amplification process is based on the synthesis of one product from one substrate.

Notwithstanding the amplification processes disclosed in the prior art, a need exists for improvements to the amplification process. It would be preferable if the amplification process required less participation and fewer manipulations by the user. Further, it would be advantageous if the amplification took place at a relatively constant ambient temperature so that the activity of the enzymes involved in the process would not be affected. It would be more expedient if a template could be used to generate more than one product from one substrate in each cycle of the amplification process.

SUMMARY OF THE INVENTION

This invention relates to an amplification process which is expedient and requires less participation and fewer manipulations by the user of the process than conventional amplification processes. The amplification takes place at a relatively constant ambient temperature. In addition, each cycle of the process generates a plurality of copies of product from one substrate. The amplification process of this invention may be used to increase the quantity of a specific nucleic acid thus circumventing the copy number problem. Hence, probe assays may be more readily used. The amplification process could also be used to increase the purity of a specific nucleic acid sequence as a substitute for conventional cloning methodology.

According to one aspect of the invention, a process for amplifying a specific nucleic acid sequence is used. The process involves the synthesis of single-stranded RNA, single-stranded DNA, and double stranded DNA. The single stranded RNA is a first template for a first primer. The single stranded DNA is a second template for a second primer. The double stranded DNA is a third template for the synthesis of a plurality of copies of the first template. A sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence and a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence. A 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

According to another aspect of the invention, a process for amplifying a specific nucleic acid sequence is used. The process involves:

(a) hybridizing a first primer to a first template. The first primer has a DNA sequence which is sufficiently complementary to a RNA sequence of the first template;

(b) synthesizing a first DNA sequence covalently attached to the first primer and complementary to the RNA sequence of the first template. The first DNA sequence and the first primer comprise a second template;

(c) separating the first template from the second template to allow hybridization of a second primer;

(d) hybridizing the second primer to the second template. The second primer has a DNA sequence which is sufficiently complementary to a DNA sequence of the second template. The second primer also has an antisense sequence of a promoter and an antisense sequence of a transcription initiation site for a RNA polymerase;

(e) synthesizing a second DNA sequence covalently attached to the second primer and complementary to the DNA sequence of the second template and synthesizing a third DNA sequence covalently attached to the second template and complementary to the DNA sequence of the second primer. The second and third DNA sequences, the second primer and the second template comprise a third template;

(f) synthesizing a plurality of copies of the RNA sequence of the first template from the third template.

IDT Exhibit 1020.7067

A sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence and a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence. A 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

In a further alternative of the invention, the second primer of DNA has a sequence at its 3′ end which is sufficiently complementary to the DNA sequence of the second template. The second primer has at its 5′ end an antisense sequence of a promoter and an antisense sequence of a transcription initiation site for a RNA polymerase.

In a further alternative of the invention, the third DNA sequence covalently attached to the second template is complementary to the DNA sequence at the 5′ end of the second primer.

In another alternative of the invention, a process for amplifying a specific nucleic acid sequence is used. The process involves combining a first primer, a second primer, ribonuclease H, a RNA-directed DNA polymerase, a DNA-directed DNA polymerase, a RNA polymerase, ribonucleoside triphosphates and deoxyribonucleoside triphosphates with a sample. The first primer of DNA has a sequence which is sufficiently complementary to a first template of RNA. The second primer of DNA has a sequence which is sufficiently complementary to a second template of DNA, and an antisense sequence of a promoter and an antisense sequence of a transcription initiation site which are recognized as substrate by the RNA polymerase. A sequence of the first primer or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence and a sequence of the first primer or the second primer is sufficiently homologous to a sequence of the specific nucleic acid. A 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

In a further alternative of the invention, a process for amplifying a specific nucleic acid sequence is used. The process involves adding a first primer, a second primer, avian myoblastosis viral polymerase, E. coli ribonuclease H, bacteriophage T7 RNA polymerase, ribonucleoside triphosphates and deoxyribonucleoside triphosphates to a sample. The first primer of DNA has a sequence which is sufficiently complementary to a first template of RNA. The second primer of DNA has a sequence which is sufficiently complementary to a second template of DNA, and an antisense sequence of a promoter and an antisense sequence of a transcription initiation site which are recognized as substrate by T7 RNA polymerase. A sequence of the first primer or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence and a sequence of the first primer or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence. A 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

BRIEF DESCRIPTION OF THE DRAWINGS

In drawings which illustrate embodiments of the invention,

Figure 1 is a general illustration the nucleic acid amplification process;

Figure 2 shows the synthetic oligonucleotides DNA sequences which are used for testing the amplification process:

Figure 2A, the gag test sequence;

Figure 2B, the gag2 test sequence;

Figure 3 is an autoradiogram of PAGE analysis of amplification reactions using different primer concentrations;

Figure 4 is an autoradiogram of PAGE analysis of amplification reactions using different template concentrations;

Figure 5 is an autoradiogram of Dot-blot hybridization on amplification reactions;

Figure 6 is an autoradiogram of PAGE analysis of amplification reaction using restriction fragments as template.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

This invention relates to a process for amplifying a specific nucleic acid sequence. The amplification involves the alternate synthesis of DNA and RNA and is generally illustrated in Figure 1. In this process, single-stranded RNA is converted to single-stranded DNA which in turn is converted to a functional template for the synthesis of a plurality of copies of the original single-stranded RNA. A first primer and a second

IDT Exhibit 1020.7068

primer are used in the amplification process. A sequence of the first primer or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence and a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence. In some instances, both the first primer and second primer are sufficiently complementary and sufficiently homologous to a sequence of the specific nucleic acid sequence, for example, if the specific nucleic acid sequence is double stranded DNA.

The RNA is converted to single-stranded DNA by hybridizing an oligonucleotide primer (the first primer) to the RNA (the first template) and synthesizing a complementary strand of DNA from the first primer, (the first DNA sequence) by using a RNA-directed DNA polymerase. The resulting single-stranded DNA (the second template) is separated from the first template by, for example, hydrolysis of the first template and by using a ribonuclease which is specific for RNA-DNA hybrids (for example, ribonuclease H). The second template is converted to a form which is capable of RNA synthesis by hybridizing a synthetic oligonucleotide (the second primer), which contains at its 3′ end a sequence which is sufficiently complementary to the 3′ end of the second template and toward its 5′ end a sequence containing the antisense strand of a promoter and antisense sequence of a transcription initiation site, and by synthesizing a second DNA sequence covalently attached to the 3′ end of the second primer using the second template as a template and synthesizing a third DNA sequence covalently attached to the 3′ end of the second template using the second primer as a template, using DNA-directed DNA polymerase. The resulting functional derivative of the second template, which is a third template, is used for the synthesis of a plurality of copies of RNA, the first template, by using a RNA polymerase which is specific for the promoter and transcription initiation site defined by the second primer. Each newly synthesized first template can be converted to further copies of the second template and the third template by repeating the cycle. In addition, repetition of the cycle does not require participation or manipulation by the user.

The amplification process commences with the addition of a suitable template nucleic acid to the appropriate enzymes, primers, and cofactors under the appropriate reaction conditions. This template nucleic acid is in a form which is capable of homogenous and continuous amplification and can function as an intermediate in the cycle set forth in Figure 1. The amplification process involves the net consumption of precursors (primers, ribonucleoside triphosphates and deoxyribonucleoside triphosphates) and the net accumulation of products (RNA and DNA). The processes of RNA and DNA synthesis will proceed asynchronously until sufficient levels of nucleic acids have been synthesized to allow detection. The amplification process may be monitored by, for example, the synthesis of a labeled product from a labeled precursor.

It is contemplated that amplification may involve another process either in addition to or in place of the one generally illustrated in Figure 1. Also possible are certain counter-productive enzymatic reactions which occur at permissibly low rates. Included among the possible non-productive side reactions is the synthesis of RNA and/or DNA in the absence of an added template nucleic acid. Such RNA and/or DNA products can be discriminated from desired products by determining whether a particular sequence which would be found only between the two priming sites of the specific nucleic acid sequence is present.

The first primer is an oligodeoxyribonucleotide which has at its 3′ end a sequence which is sufficiently complementary to the 3′ end of the first template. The sequence at the 3′ end of the first primer has a particular length and base composition to allow specific and efficient synthesis of the first DNA sequence, under the given conditions of ionic strength and temperature. The first primer may be sufficiently complementary to a region internal to the 3′ end of the first template in the first cycle. In subsequent cycles, the 5′ end of the first primer would be complementary to the 3′ end of the first template. It is contemplated that the first primer may be composed partially or completely of nucleotides or nucleotide analogs other than the natural deoxyribonucleotides. The 5′ end of the first primer may contain sequences which are not complementary to the first template in the first cycle. The non-complementary sequences may be complementary to a nucleic acid which can be immobilized, or to which can be bound a useful non-nucleic acid component, such as a reporter to facilitate detection. Alternatively, the non-complementary sequences may include an antisense sequence of a promoter and an antisense sequence of a transcription initiation site, which could be used for the synthesis of RNA. This RNA would be complementary to the first template and could be used as an intermediate in another amplification cycle.

The second primer is an oligodeoxyribonucleotide which contains at its 3′ end a sequence which is sufficiently complementary to the 3′ end of the second template. The second primer has a particular length and base composition to allow specific and efficient synthesis of the second and third DNA sequences, under the given conditions of ionic strength and temperature. In addition, the second primer contains the antisense sequence of a functional promoter and the antisense sequence of a transcription initiation site. This sequence, when used as a template for synthesis of the third DNA sequence, contains sufficient

IDT Exhibit 1020.7069

information to allow specific and efficient binding of a RNA polymerase and initiation of transcription at the desired site. The promoter sequence may be derived from the antisense strand of a functional promoter. The transcription initiation site may be derived from the 5′ terminal sequence of a natural RNA transcript. In the preferred embodiment, the 5′-terminal sequence of the second primer is AATTCTAATACGACTCAC-TATAGGGAG. This sequence contains the antisense sequence of the promoter and the antisense sequence of the transcription initiation site for T7 RNA polymerase. Alternatively, the transcription initiation site and promoter for another phage RNA polymerase may be used. In addition, sequences which are unrelated to the promoter function may be included at the 5′ end of the second primer or between the transcription initiation site and the sequence at the 3′ end which hybridizes to the second template. It is contemplated that the second primer may be composed partially or completely of nucleotides or nucleotide analogs other than natural deoxyribonucleotides.

All of the enzymes used in this invention should meet certain practical specifications. Each enzyme or enzyme preparation should be free of deleterious deoxyribonuclease ("DNase") activities, such as the 5′ or 3′ exonuclease activities which are often associated with certain DNA polymerases and single-strand or double-strand specific exonuclease or endonucleases. Each enzyme or enzyme preparation should be free of deleterious ribonuclease ("RNase") activities, with the exception of the preferred addition of a ribonuclease activity which is specific for hybrids of RNA and DNA (for example, ribonuclease H). In addition, each enzyme should be reasonably active under the common reaction conditions which are used for the other enzymatic processes, and non-enzymatic processes, such as hybridizing oligonucleotide primers to the RNA or DNA templates.

The DNA-directed RNA polymerase which is used in this invention may be any enzyme capable of binding to a particular DNA sequence called a promoter and specifically initiating in vitro RNA synthesis at a defined initiation site within close proximity to the promoter. The promoter and the initiation site form part of the second primer. In addition the RNA polymerase should be capable of synthesizing several copies of RNA per functional copy of template in a reasonable amount of time. In the preferred embodiment, the bacteriophage T7 RNA polymerase is used. In addition other bacteriophage RNA polymerases, such as phage T3, phage ØII, Salmonella phage sp6, or Pseudomonas phage gh-1 may be used. In another embodiment, other prokaryotic or eukaryotic DNA-directed RNA polymerase may be used. It should be understood that if alternative RNA polymerases are used, then the necessary changes to the promoter and initiation sequences of the second primer should be made according to the template specificity of the particular RNA polymerase.

The RNA-directed DNA polymerase which is used in this invention may be any enzyme capable of synthesizing DNA from an oligodeoxyribonucleotide primer and a RNA template. In addition this enzyme may contain activities for DNA-directed DNA polymerase and RNase H. In the preferred embodiment, the avian myoblastosis viral polymerase ("AMV reverse transcriptase") is used. In addition, the RNA-directed DNA polymerase could be from another retrovirus, such as Maloney murine leukemia virus.

Alternatively, other eukaryotic RNA-directed DNA polymerases could be used.

The DNA-directed DNA polymerase which is used in this invention may be any enzyme capable of synthesizing DNA from an oligodeoxyribonucleotide primer and a DNA template. This enzyme should not contain either 5′- or 3′- exonuclease activities, which are associated with many types of DNA polymerase. In the preferred embodiment, the AMV reverse transcriptase is used. However, other DNA-directed DNA polymerases which naturally lack the 5′- or 3′- exonuclease activities could be used. These could include certain eukaryotic DNA polymerases, such as, DNA polymerase γ or β those DNA polymerases which could be isolated from a mammalian tissue, such as calf thymus. An otherwise unsuitable DNA polymerase could be made useful by removing the undesirable exonuclease activities either by alteration of the DNA polymerase gene followed by expression of the altered polymerase in a suitable host cell, or by chemical modification of the DNA polymerase protein. Altered versions of DNA polymerase could be made from the Klenow fragment of E. coli DNA polymerase I or the bacteriophage T7 DNA polymerase. It should be understood that such alternative DNA-directed DNA polymerase activities are added to supplement the activity contributed by the RNA-directed DNA polymerase, since in the preferred embodiment, both RNA-directed and DNA-directed DNA polymerase activities are supplied by the same enzyme.

The RNase H which could be used in this invention may be any enzyme capable of hydrolyzing a RNA which is annealed to a complementary DNA. This enzyme should not be capable of hydrolyzing single or double-stranded RNA or any DNA. In the preferred embodiment, the E. coli RNase H is used. In addition, other RNase H enzymes could be used, such as calf thymus RNase H. Since RNase H is an intrinsic activity of AMV reverse transcriptase, the E. coli RNase H will be supplemented in the preferred embodiment by the RNase H of AMV reverse transcriptase. Alternatively, any other enzyme capable of separating the second template from the first template could be used.

IDT Exhibit 1020.7070

The abovementioned enzymes and primers are mixed together in a reaction vessel which contains the necessary buffers and cofactors for both DNA and RNA synthesis. In addition, the ionic conditions and reaction temperature should be compatible with specific hybridization of the primers to the DNA and RNA templates as is known to those skilled in the art. The reaction mixture should be free of such agents which would interfere with the amplification process, specifically substances which could greatly inhibit the activity of the enzymes, interfere with the hybridizing of primers and templates, or degrade non-productively the nucleic acid intermediates and products.

The description of possible detection schemes may be useful to the application of the amplification process. It should be understood that schemes which may be used for detecting the nucleic acids which are synthesized in the amplification process are not limited to those described herein, and it is contemplated that other methods may be used.

In one embodiment, a labeled precursor may be added to the reaction mixture. Amplification is determined by quantitive or qualitative analysis of labeled products, which can be separated from the labeled precursor by using methods known in the art.

A labeled precursor may be a ribonucleoside triphosphate for detecting RNA synthesis, or a deoxynucleoside triphosphate or an oligonucleotide primer for detecting DNA synthesis. The type of label may be a radioisotope or a useful chemical group, such as biotin, a chromophore, a fluorophore, or a hapten which could bind to an antibody, or possibly a protein or an enzyme. The labeled products may be separated from the labeled precursors on the basis of solubility, charge, or size. In addition, the labeled DNA or RNA may be hybridized to a nucleic acid which contains a complementary sequence and which can be immobilized.

In another embodiment, the products of the amplification process may be bound to an immobilized support, hybridized to a nucleic acid probe containing a complementary sequence, and separated from the unhybridized nucleic acid probe which remains in solution. The products, DNA or RNA, may be bound directly to a solid support by any stable interaction, such as hydrophobic, electrostatic, or covalent interaction. In addition, the products may contain certain chemical groups, for example, biotin, which may be incorporated into the products during the amplification process to allow binding to an immobilized protein, for example, avidin or streptavidin. In addition, the products may be hybridized to a nucleic acid which contains a complementary sequence and which can be immobilized. The nucleic acid probe would contain a complementary sequence which forms a sufficiently stable interaction with a product of the amplification process to allow binding under the conditions of hybridization and sustained binding under the conditions used for removal of the unhybridized nucleic acid probe. In the preferred embodiment the complementary sequence would be derived from that part of the specific nucleic acid sequence which is between the sequences of the first primer and the second primer. The nucleic acid probe may be a single-stranded DNA or RNA, or a double-stranded DNA or RNA which can be made single-stranded, or an oligonucleotide which can be composed of deoxyribonucleotides and/or ribonucleotides. In addition, the nucleic acid probe may contain a chemical group which could covalently bind to a product DNA or RNA under the appropriate conditions. The nucleic acid probe may be labeled with a radioisotope or a useful chemical group, such as biotin, a chromophore, a fluorophore, or a hapten which could bind to an antibody. In addition, the nucleic acid probe could be conjugated to a protein or enzyme, for example, a phosphatase or a peroxidase. In addition, the nucleic acid probe may contain sequences which would allow in vitro replication of the probe.

It is contemplated that the products of the amplification process may be analyzed by methods which are typically used for nucleic acids that have been enriched by molecular cloning techniques. In one alternative, the synthesis of a specific DNA sequence may be detected by digestion of the synthesized DNA with a restriction endonuclease, followed by electrophoretic separation and detection using methods known in the art. In another alternative, the sequence of amplified RNA may be determined by DNA synthesis using a RNA-directed DNA polymerase, the first primer, and dideoxynucleoside triphosphates (Stoflet et al., 1988). In another alternative, the sequence of the amplified third template may be determined by RNA synthesis using the DNA-directed RNA polymerase used in the amplification process, and 3′-deoxyribonucleoside triphosphates (Axelrod & Kramer, 1985). In another alternative, the amplified RNA may encode a polypeptide which could be translated, in vitro. The polypeptide product of the in vitro translation could be analyzed by using an antibody.

A sample suspected of containing or known to contain the specific nucleic acid sequence is added to the reaction mixture in the form of a template nucleic acid which is capable of homogeneous and continuous amplification and may be any intermediate in the cycle set forth in Figure 1. In particular, the template nucleic acid may be a single-stranded RNA which contains at its 5′ end a sequence which is sufficiently homologous to that which is at the 3′ end of the second primer, and contains a sequence which

IDT Exhibit 1020.7071

is sufficiently complementary to the first primer. A template nucleic acid of this form would function as a first template in the amplification process. Alternatively, the template nucleic acid may be a single-stranded DNA which contains at its 3′ end a sequence which is sufficiently complementary to at least the 3′ end of the second primer, and contains a sequence which is sufficiently homologous to that which is at the 3′ end of the first primer. A template nucleic acid of this form would function as a second template in the amplification process. Alternatively, the template nucleic acid may be a double-stranded DNA, one strand of which contains at its 5′ end the entire sequence of the second primer and contains a sequence which is sufficiently complementary to the first primer. The double-stranded DNA functions as a third template in the amplification process.

Although the preparation of a template nucleic acid is not part of the amplification process, the description of possible schemes for generating template nucleic acids may be useful to the application of the amplification process. It should be understood that the schemes which may be used for obtaining the template nucleic acid are not limited to the alternatives which are described herein, and it is contemplated that other methods may be used.

In one alternative, a template nucleic acid which could function as a first template could be a naturally occurring RNA or a RNA fragment which could be generated from a larger RNA molecule by using site specific hydrolysis methods known in the art (Shibahara et al., 1987).

In another alternative, a template nucleic acid which could function as a second template could be generated from a double-stranded DNA by digestion with a restriction endonuclease which has a site immediately flanking the sequence which is sufficiently complementary to the 3′ end of the second primer. The resulting double-stranded DNA fragments could then be made single-stranded by using chemical or thermal denaturation methods.

In another alternative, a template nucleic acid which could function as a second template could be generated from a single-stranded DNA or RNA to which has been hybridized an oligonucleotide which is capable of blocking DNA synthesis. This blocking oligonucleotide may contain a chemical group, which could covalently bind to the template, under the appropriate conditions. DNA synthesis from this blocked template using the first primer could result in a synthesized DNA with the same 3′ end as the second template. If the original template is RNA, then the resulting DNA-RNA hybrid may be used directly as a template nucleic acid. If the original template is DNA, then the resulting copy of the second template could then be separated from the original template by using chemical or thermal denaturation methods.

In another alternative, a template nucleic acid which could function as a third template could be generated from a single-stranded DNA or RNA by DNA synthesis from the DNA or RNA template using the second primer. The resulting synthesized DNA could then be separated from the original template by using chemical or thermal denaturation methods. In addition, a RNA template could be hydrolyzed by using chemical or enzymatic methods. The resulting single-stranded DNA has the sequence of the second primer covalently attached to its 5′ end and contains a sequence which is sufficiently complementary to the first primer. This single-stranded DNA could be converted to a transcriptionally functional double-stranded DNA by hybridizing the first primer to the single-stranded DNA, and by synthesizing a DNA sequence which is covalently attached to the first primer and complementary to the single-stranded DNA.

In a further alternative, a single-stranded DNA or RNA template could be obtained from a double-stranded DNA, double-stranded RNA or a DNA-RNA hybrid by using chemical, thermal, or possibly enzymatic methods. Then, by using one of the alternative schemes proposed above, the resulting single-stranded DNA or RNA could then be used to generate a template nucleic acid which could function as a first, second or third template. In addition, an alternative scheme involving the first primer and one strand of nucleic acid, and another alternative scheme involving the second primer and the other (complementary) strand of the nucleic acid may be used concurrently to generate template nucleic acids.

MATERIALS AND METHODS

Materials

Oligonucleotides were synthesized using an Applied Biosystems 380A DNA synthesizer. Columns, phosphoramidites, and reagents used for oligonucleotide synthesis were obtained from Applied Biosystems, Inc. through Technical Marketing Associates. Oligonucleotides were purified by polyacrylamide gel electrophoresis followed by DEAE cellulose chromatography. The radioisotope [α -32p] UTP (800 Ci/mmol) was from Amersham. Enzymes for digesting and ligating DNA were purchased from New England Biolabs, and

IDT Exhibit 1020.7072

used according to the supplier's recommendations. Preparations containing the large fragment of DNA polymerase 1 (Klenow) were also purchased from New England Biolabs. RNasin and T7 RNA polymerase from Promega Biotec were purchased through Bio/Can Scientific Inc. Reverse transcriptase and RNase H were obtained from Pharmacia. The supplier for proteinase K was Boehringer Mannheim Canada. E. coli strain HB101 (ATCC 33694) was used for all transformations. The plasmid pUC19 (Norrander et al., 1983) was purchased from Bethesda Research Laboratories.

Isolation of DNA and sequencing

E. coli transformants were grown on YT medium (Miller, 1972) containing 50 ug/ml ampicillin. Plasmid DNA was purified by a rapid boiling method (Holmes and Quigley, 1981). DNA fragments and vectors used for all constructions were separated by electrophoresis on low melting point agarose, and purified from the molten agarose by phenol extraction and ethanol precipitation (Maniatis et al., 1982). Plasmid DNA was sequenced using a modification (Hattori et al., 1985) of the dideoxy method (Sanger et al., 1977). Reactions were run using the -20 universal primer (New England Biolabs).

TCA precipitation

Aliquots (5ul) of amplification reactions were quenched in 20ul 10 mM EDTA and placed on ice until all time point samples had been collected. The quenched samples were then applied to glass filter discs, and immediately dropped into ice-cold 5% trichloroacetic acid ("TCA") - 1% sodium pyrophosphate for 10 min with occasional mixing. Two 5 min washes with ice-cold 5% TCA were followed by two additional washes with 95% ethanol and lyophilization to dryness. Radioactivity was determined in a liquid scintillation counter.

Polyacrylamide gel electrophoresis

Samples (1 to 6 ul) were mixed with 4-5 ul formamide dye (90% deionized formamide, 10 mM TrisHCl (pH 8.0), 1 mM EDTA, xylene cyanol and bromophenol blue), and applied to a pre-run 12-cm- long 7% denaturing polyacrylamide gel. Gels were run at 350 volts until the bromophenol blue dye had reached the bottom. In some cases the gels were fixed and dried prior to autoradiography. Fixing involved a 15 min wash in 10% methanol- 7% acetic acid. The profiles of the RNA products separated by this procedure were visualized by autoradiography at room temperature.

Example 1: Design and synthesis of oligonucleotides for a gag test system

A synthetic DNA sequence (Figure 2A) was designed to include an EcoRI site, a T7 phage promoter, a sequence required for initiation of transcription by T7 RNA polymerase and a 19 bp hybridization region (hybridization region 1). The 47 b antisense strand oligonucleotide (T7H1.GAG) involved in the cloning of these elements also serves as the first primer. Hybridization region 2 lies 53 bp away from hybridization region 1 and is 20 bp in length. The primer made to this region (H2.GAG) is a 20 b oligonucleotide duplicate of the sense strand and is not used for cloning. The sequence spanning and including the hybridization regions is a 92 bp segment of the gag portion of the HTLV-III genome, the causative agent of AIDS. This particular gene segment was chosen because the primers were predicted to hybridize efficiently and because the distance between the two hybridization regions was relatively short. In addition, a XbaI site was placed at the end of the sequence for cloning ease. The gag test sequence also contains SphI and PstI sites which can aid in the screening of recombinants.

A total of four oligonucleotides were used in the cloning of this fragment. N1.GAG, used in the construction of both the gag test and the gag2 test sequence, completes the antisense strand and is used only in the cloning process. Similarly, T74.PRO is the sense strand component of the T7 promoter. N2.GAG however, was used in the construction of both test fragments, and has also been used as an intermediate (second template) in two steps of the amplification cycle. The entire cloned gag test fragment can also represent an intermediate of the amplification cycle (third template). Once cloned into an appropriate vector the gag test DNA could be transcribed by T7 RNA polymerase to produce a RNA fragment (first template) useful as an amplification intermediate involved in three of the steps. In addition, T7H1.GAG and H2.GAG

IDT Exhibit 1020.7073

serve as primers in the test system.

The gag2 test synthetic DNA fragment (Figure 2B) does not contain the T7 promoter, but the remainder of the sequence is identical to the gag test sequence and therefore, both N1.GAG and N2.GAG were involved in its construction. The oligonucleotide required to complete the antisense strand is called H1.GAG. Once cloned, the gag2 test fragment can be used as a template for testing amplification, using a DNA restriction fragment as template nucleic acid.

Example 2: Construction of the gag test plasmids

The oligonucleotides T74.PRO and N1.GAG (2 ug each), were phosphorylated separately in 20 ul reactions containing 70 mM TrisHCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM ATP and 5 units T4 polynucleotide kinase, at 37° C for 30 min. Phosphorylated T74.PRO and N1.GAG (10 ul of each) were mixed with 1 ug each of unphosphorylated T7H1.GAG and N2.GAG, and 3 ul 100 mM TrisHCl (pH 7.8) - 500 mM NaCl, in a final volume of 29 ul for the gag test assembly. The gag2 test mixture contained 10 ul phosphorylated N1.GAG, 1 ug each of unphosphorylated H1.GAG and N2.GAG, and 1.8 ul 100 mM TrisHCl (pH 7.8) - 500 mM NaCl, in a final volume of 18 ul. The oligonucleotide mixtures were hybridized separately by placing them at 90° C for 10 min followed by slow cooling to room temperature for 10 to 16 h 60 ul reactions containing 50 mM TrisHCl (pH 7.8), 10 mM $MgCl_2$, 20 mM DTT, 1mM ATP and 50 ug/ml BSA were used to ligate the hybridized oligonucleotides together. 400 units T4 DNA ligase was added to the gag test reaction and it was incubated at 15° C for 2h while the gag2 test reaction was incubated for 14 to 16 h with 200 units T4 DNA ligase.

The isolated and purified synthetic DNA segments were mixed with plasmid pUC19 which had been linearized by digestion at restriction enzyme sites within the polylinker region. T4 DNA ligase was used to ligate the gag test sequence into the EcoRI- XbaI fragment of pUC19, while the gag2 test sequence was ligated to the SmaI-XbaI fragment. Plasmid DNA from transformants obtained after these reactions were used to transform E. coli were screened by restriction analysis , and the final plasmids (pGAG.TEST and pGAG2.TEST) were determined to be correct by sequence analysis.

Example 3: Effect of primer concentration on RNA amplification

The reaction mixtures (25 ul) which were used to amplify RNA transcribed from the gag test oligonucleotides contained 50 mM TrisHCl (pH 8.45), 6 mM $MgCl_2$, 40 mM KCl, 10 mM dithiothreitol, 0.5 mM NTP (ATP, CTP, GTP, UTP), 1 mM dNTP (dATP, dCTP, dGTP dTTP), 20 units RNasin, 10 units T7 RNA polymerase, 10 units reverse transcriptase, 0.4 units RNase H, and 10 uCI [α -32p] UTP. Two of the reactions contained 0.5 ng (0.015 pmoles) N2.GAG while the other two reactions contained no template. The primers T7H1.GAG and H2.GAG were each added at final concentrations of 3.4 uM or 0.34 uM to reactions containing either N2.GAG or no template. The reactions were incubated at 42° C for 2 h. Total synthesis of RNA was monitored by determining the incorporation of TCA insoluble cpm at 30 min. intervals. The effect of the primer concentration on template- dependent RNA synthesis is shown in Table 1. Aliquots of each reaction, containing equal amounts of synthesized RNA, were analyzed by PAGE and autoradiography (Figure 3, lanes 1-4 numbered the same as the reactions).

Table 1

RNA amplification from N2.GAG after 2 h.			
Reaction	Concentration of each primer (uM)	Template (ng)	RNA Synthesized (ug)
1	3.4	0.5	2.8
2	3.4	-	2.1
3	0.34	0.5	1.8
4	0.34	-	0.7

It was found that while reaction 1 resulted in the greatest incorporation of isotope, the no template

IDT Exhibit 1020.7074

control, reaction 2, was also high (73% of reaction 1) and produced a very similar electrophoretic profile. It would therefore appear that in the presence of high primer concentrations, a RNA transcript of identical size to that expected in amplification is produced in the absence of any template. Results using samples with a 10- fold decrease in primer concentration were dramatically different. The amount of RNA produced in reaction 3 was 2.6 times that of reaction 4, but while virtually all of the transcript was found in a single band of the expected size in reaction 3, no fragments greater than 60 to 70 b were found in reaction 4. Primer concentration therefore plays a significant role in the accuracy and efficiency of RNA amplification.

A control RNA transcript used to represent the size of fragment expected to be generated by the amplification system (lane 0 of Figure 3) was prepared by transcription from the test plasmid. pGAG.TEST was linearized by digestion with XbaI, proteinase K treated (Maniatis et al., 1982), phenol extracted and ethanol precipitated. T7 RNA polymerase was then used according to the supplier's recommendations to transcribe 0.5 ug of the resulting fragment in a 25 ul reaction mixture containing 10 uCi [α -32p] UTP.

Example 4: Effect of template concentration on RNA amplification

The standard 50 ul reaction mixture used to amplify RNA transcribed from the gag test oligonucleotides contained 0.34 uM T7H1.GAG, 0.34 uM H2.GAG, 50 mM TrisHCl (pH 8.45), 6 mM $MgCl_2$, 40 mM KCl, 10 mM DTT, 0.5 mM NTP, 1 mM dNTP, 40 units RNasin, 20 units T7 RNA polymerase, 20 units reverse transcriptase, 0.8 units RNase H and 10 - 20 uCi [α-32p] UTP. The reactions contained amounts of template (N2.GAG) varying from 1 ng to 1 fg. One reaction contained no template. The reactions were incubated at 42° C for 3h, during which total synthesis of RNA was monitored by determining the incorporation of TCA insoluble cpm at 30 min intervals. As indicated in Table 2, total RNA synthesis was higher than the no template control for all template concentrations tested. Although the total synthesis of RNA generally decreased with decreasing template concentration, this decrease in synthesis was not quantitative. Thus, the degree of amplification of RNA per starting template generally increased with decreasing template concentration. Amplification of 8 x 10^8 fold was achieved by synthesis of 0.8 ug RNA from 1 fg of N2.GAG template. One fg of the 102-b N2.GAG oligonucleotide represents approximately 2 x 10^4 molecules.

Table 2

RNA amplification from N2.GAG after 3h.			
Reaction	Template	RNA Synthesized (ug)	Fold amplification
1	1 ng	3.5	3.5 x 10^3
2	100 pg	4.4	4.4 x 10^4
3	10 pg	4.1	4.1 x 10^5
4	1 pg	3.0	3.0 x 10^6
5	100 fg	2.7	2.7 x 10^7
6	10 fg	1.9	1.9 x 10^8
7	1 fg	0.78	7.8 x 10^8
8	-	0.046	-

The RNA which was synthesized after a reaction time of 3h was analyzed by PAGE, for each template concentration (Figure 4 lanes 1-8, numbered the same as the reactions). A major band representing a RNA of about 100 b was present in all reactions except the ones containing 1 fg template and no template. The reaction containing 1 fg template did not have much of this 100 b product at 3 h but the overall RNA synthesis was higher than and qualitatively different from the no template reaction.

Example 5: Hybridization analysis of RNA products.

Amplification reactions containing amounts of N2.GAG template varying from 1 pg to 0.1 fg were performed following the teaching of example 4, except the radiolabeled UTP was omitted. The reactions were incubated at 42° C for 3 h. Aliquots were removed from each reaction at 30 min intervals and applied

IDT Exhibit 1020.7075

to a nylon membrane (Amersham). The nucleic acids that were contained in these reaction aliquots were fixed by exposure to ultraviolet light. The membrane was prehybridized at 50 ° C for 1 h in prehybridization buffer consisting of a final concentration of 50% v/v formamide, 5 X SSC and 5 X Denhardt's solution (Maniatis et al, 1982; Southern et al, 1975) at a volume equivalent to 5 mls of solution per 100 cm^2 and hybridized with a radiolabeled probe with a specific activity of 10^6 cpm/ml of hybridization solution. Hybridization was performed at 50 ° C for 16 h in 50% formamide, 5 X SSC and 5 X Denhardt's solution (Maniatis et al, 1982; Southern et al, 1975). The radiolabeled probe was the synthetic oligonucleotide 5′ GATCTGGGATAGAGTACATCCA 3′ which had been labeled at the 5′ end using T4 polynucleotide kinase and (α-32p) ATP. After the membrane was washed at 50 ° C in a series of 2, 3 min. washes consisting of 2 × SSC, 0.1% v/v SDS and 0.2 X SSC, 0.1% v/v SDS (Southern et al, 1975; Maniatis et al, 1982; Szostak et al, 1979).

Figure 5 shows the results of the hybridization analysis performed on the amplification reactions, containing various amounts of N2.GAG template, which were sampled at different times of incubation.

Each column of Figure 5 represents a different time point (1, 30 min; 2, 60 min; 3, 90 min; 4, 120 min; 5, 150 min; 6, 180 min) and each row represents a different amount of added N2.GAG template (1, 1pg; 2, 100 fg; 3, 10 fg; 4, 1 fg; 5, 0.1 fg; 6, no template). Amplification of nucleic acids which hybridized to the labeled probe were observed for rows 1-3 (1 pg - 10 fg), however the hybridization to specific nucleic acids in rows 4 - 5 (1 fg, 0.1 fg) was not higher than row 6 (no template). The apparent non-specific binding of labeled probe in row 6 seems to be related to DNA or RNA synthesis since the hybridization signal increases with increasing time.

Example 6: Use of DNA restriction fragment as template.

The plasmid pGAG2.TEST was digested with MspI, treated with proteinase K, purified by phenol extraction and ethanol precipitation, and denatured by boiling for 5 min. Amplification reactions were performed and analyzed following the teaching of Example 4, except the MspI digested pGAG2.TEST was used as a template instead of the N2.GAG oligonucleotide. The amounts of plasmid added to each reaction varied from 55 ng to 5.5 pg, and no template. To simulate additional DNA which would be present in an actual sample, alternate reactions contained 1 ng of calf thymus DNA which had been similarly digested, purified and denatured. After a 3 h incubation at 42 ° C, the synthesis of RNA was determined by TCA precipitation and PAGE analysis. As indicated in Table 3, total RNA synthesis was higher than the no template controls for all template concentrations tested. The degree of amplification was calculated based on RNA synthesis from the actual template which was 1.8% of the total plasmid DNA.

The total RNA synthesis (degree of amplification) from a particular initial level template concentration was consistently lower for the restriction fragment (Table 3) as compared to that for the synthetic oligonucleotide template (Table 2). This could be due to competition with the complementary strand of the restriction fragment template under the conditions used.

IDT Exhibit 1020.7076

Table 3

RNA amplification from MspI-digested pGAG2.TEST			
Reaction	Template*	RNA Synthesized**	Fold amplification**
1 2	55.0 ng [1 ng]	3.65 (4.05)	3.7 x 103 (4.1 x 103)
3 4	5.5 ng [100 pg]	3.54 (3.16)	3.5 x 104 (3.2 x 104)
5 6	550.0 pg [10 pg]	2.29 (2.79)	2.3 x 105 (2.8 x 105)
7 8	55.0 pg [1 pg]	2.62 (0.67)	2.6 x 106 (0.7 x 106)
9 10	5.5 pg [100 fg]	1.37 (2.26)	1.4 x 107 (2.3 x 107)
11 12	-	1.25 (0.08)	-

* Numbers in brackets indicate equivalent amounts of N2.GAG.

** Numbers in parentheses indicate RNA synthesis in presences of 1 ug MspI-digested calf thymus DNA.

The RNA which was synthesized after a reaction time of 3h was analyzed by PAGE (Figure 6, lanes 1-6, 11 and 12, numbered the same as the reactions). The major band representing a RNA of about 100 b was present in reactions (lanes) 1-6 but absent in the no template reactions (lanes 11 and 12). The RNA in lane 0 was a standard, which was prepared following the teaching of Example 3. There was no apparent qualitative difference in the synthesized RNA either with (lanes 2, 4 and 6) or without (lanes 1, 3, and 5) the additional of 1 ug of MspI-degested calf thymus DNA.

Example 7: Use of RNA fragment as template

The plasmid pGAG.TEST is digested with XbaI, treated with proteinase K, and purified by phenol extraction and ethanol precipitation. RNA of a sequence complementary to N2.GAG is transcribed from the linearized pGAG.TEST plasmid using T7 RNA polymerase. The resulting RNA is purified by digestion with DNase (Pro Mega BioTec), followed by phenol extraction and ethanol precipitation. The purified RNA is used as a template for amplification reactions following the teaching of Example 5. Amounts of RNA are added to each reaction and vary from 55 ng to 5.5 pg, and no template. After a 3 h incubation at 42 ° C, the synthesis of specific RNA is determined by hybridization to labeled oligonucleotide probe, following the teaching of Example 5.

Example 8: Use of ribosomal RNA as a template; amplification of internal sequences

Two primers are used for amplifying RNA sequences which are complementary to a part of E. coli 16S ribosomal RNA (rRNA). One of these primers T7HIRIB3.PR2 (AATTCTAATACGACTCACTATAGGGAGTATTACCGCGGCTGCTG) contains the antisense strand of the T7 promoter and initiation site, and a sequence which is complementary to 16S rRNA. The other RIB8.PR (AATACCTTTGCTCATTGACG) is complementary to the DNA synthesized by using T7H1RIB3.PR2 as a primer and 16S rRNA as a template. A third synthetic oligonucleotide RIB5.PR (AGAAGCACCGGCTAAC) which allows detection of amplification is complementary to the RNA products of the amplification reaction, which are in turn complementary to the original rRNA template.

Reaction mixtures (25 ul) contain 50 mM TrisHCl (pH 8.45), 6 mM $MgCl_2$, 40 mM KCl, 10 mM DTT, 0.5

IDT Exhibit 1020.7077

mM NTP, 1 mM dNTP, 20 units RNasin, 10 units T7 RNA polymerase, 10 units AMV reverse transcriptase, 0.4 units RNase H, 0.34 um T7H1RIB3.PR2, and 0.34 um RIB8.PR.

Amounts of E. coli rRNA varying from 50 ng to 50 fg are added to the reactions. One reaction contains no added rRNA. The reactions are incubated at 42° for 3 h, during which aliquots are removed at 30, 60, 120, and 180 minutes. The reaction aliquots are quenched, fixed to a nylon membrane, and hybridized to the ^{32}P 5′-end labeled RIB5.PR probe, following the teaching of Example 5.

Example 9: Use of ribosomal RNA as a template; amplification of 5′-terminal sequences.

Two primers are used for amplifying RNA sequences which are homologous to a part of E.Coli 16S rRNA. One of these primers RIB12.PR (TTACTCACCCGTCCGCC) is complementary to 16S rRNA. The other T7H1RIB5.PR (AATTCTAATACGACTCACTATAGGGAGAAATTGAAGAGTTTGATCAT) is complementary to the 3′end of the DNA synthesized by using RIB12.PR as a primer and 16S rRNA as a template. A third synthetic oligonucleotide RIB11.PR (GTTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCC) which allows detection of amplification is complementary to both the RNA products of the amplification and the original rRNA template. The amplification reactions for rRNA and detection of the synthesized RNA are performed following the teaching of Example 8, except that T7H1RIB5.PR and RIB12.PR are used as primers (in place of T7H1RIB3.PR2 and RIB8.PR) and RIB11.PR is used as an oligonucleotide probe (in place of RIB5.PR).

Although preferred embodiments of the invention have been described in detail, it will be understood by those skilled in the art that variations may be made thereto without departing from either the spirit of the invention or the scope of the appended claims.

Claims

1. A process for amplifying a specific nucleic acid sequence comprising:
synthesizing single-stranded RNA, single stranded DNA and double stranded DNA in which the single stranded RNA is a first template for a first primer, the single stranded DNA is a second template for a second primer, the double stranded DNA is a third template for synthesis of a plurality of copies of the first template, where a sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

2. A process as set forth in claim 1 in which the first primer selectively hybridizes to the first template and promotes the synthesis of DNA covalently attached to the first primer and complementary to the first template; and the second primer, having an antisense sequence for a promoter and an antisense sequence for a transcription initiation site for a RNA polymerase, selectively hybridizes to the second template thereby permitting the synthesis of DNA covalently attached to the second primer and complementary to the second template.

3. A process for amplifying a specific nucleic acid sequence comprising:
(a) hybridizing a first primer of DNA to a first template, the first primer having a sequence sufficiently complementary to a RNA sequence of the first template;
(b) synthesizing a first DNA sequence covalently attached to the first primer and complementary to the RNA sequence of the first template, wherein the first DNA sequence and the first primer comprise a second template;
(c) separating the first template from the second template to allow hybridization of a second primer;
(d) hybridizing the second primer of DNA to the second template, the second primer having a sequence sufficiently complementary to a DNA sequence of the second template, the second primer having an antisense sequence of a promoter and an antisense sequence of a transcription initiation site for a RNA polymerase;
(e) synthesizing a second DNA sequence covalently attached to the second primer and complementary to the DNA sequence of the second template and synthesizing a third DNA sequence covalently attached to the second template and complementary to the DNA sequence of the second primer wherein the second and third DNA sequences, the second primer and the second template comprise a third template;
(f) synthesizing a plurality of copies of the RNA sequence of the first template from the third template;
where a sequence of the first or the second primer is sufficiently complementary to a sequence of the

IDT Exhibit 1020.7078

specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

4. A process as set forth in claim 3 in which the second primer has a sequence at its 3′ end sufficiently complementary to the DNA sequence of the second template and has at its 5′ end an antisense DNA sequence of a promoter and an antisense sequence of a transcription initiation site for a RNA polymerase.

5. A process as set forth in claim 4 in which the third DNA sequence covalently attached to the second template is complementary to the DNA sequence at the 5′ end of the second primer.

6. A process as set forth in claim 3 in which the first DNA sequence is synthesized by a RNA directed DNA polymerase.

7. A process as set forth in claim 3 in which the first template is separated from the second template by a ribonuclease specific for hybrids of RNA-DNA.

8. A process as set forth in claim 7 in which the ribonuclease is E. coli ribonuclease H.

9. A process as set forth in claim 7 in which the ribonuclease is calf thymus ribonuclease H.

10. A process as set forth in claim 3 in which the second and third DNA sequences are synthesized by a DNA directed DNA polymerase.

11. A process as set forth in claim 3 which the first template is synthesized from the third template by a DNA directed RNA polymerase.

12. A process as set forth in claim 3 in which the first primer has an antisense sequence of a promoter and an antisense sequence of a transcription initiation site.

13. A process as set forth in claim 3 in which the first primer has a sequence which hybridizes to a support means.

14. A process as set forth in claim 3 in which the second primer has from its 3′ to its 5′ end, a sufficiently complementary sequence, an antisense sequence of a transcription initiation site, and an antisense sequence of a promoter.

15. A process as set forth in claim 14 in which the antisense sequence of the transcription initiation site and the antisense promoter sequence bind a bacteriophage RNA polymerase.

16. A process as set forth in claim 14 in which the antisense sequence of the transcription initiation site and the antisense promoter sequence bind T7 RNA polymerase.

17. A process as set forth in claim 14 in which the antisense sequence of the transcription initiation site and the antisense sequence of the promoter is AATTCTAATACGACTCACTATAGGGAG.

18. A process as set forth in claim 11 in which the DNA directed RNA polymerase is T7 RNA polymerase.

19. A process as set forth in claim 11 in which the DNA directed RNA polymerase is a bacteriophage RNA polymerase.

20. A process as set forth in claim 11 in which the DNA directed RNA polymerase is a phage T3 polymerase.

21. A process as set forth in claim 11 in which the DNA directed RNA polymerase is phage ØII polymerase.

22. A process as set forth in claim 11 in which the DNA directed RNA polymerase is salmonella phage sp6 polymerase.

23. A process as set forth in claim 11 in which the DNA directed RNA polymerase is Pseudomonas phage gh-1 polymerase.

24. A process as set forth in claim 6 in which the RNA directed DNA polymerase is a retroviral polymerase.

25. A process as set forth in claim 24 in which the retroviral polymerase is an avian myoblastosis viral polymerase.

26. A process as set forth in claim 24 in which the retroviral polymerase is a Maloney murine leukemia viral polymerase.

27. A process as set forth in claim 6, 10 or 11 in which the polymerase lacks exonuclease or endonuclease activity.

28. A process as set forth in claim 10 in which the DNA directed DNA polymerase is an avian myoblastosis viral polymerase.

29. A process for amplifying a specific nucleic acid sequence comprising:

combining a first primer, a second primer, ribonuclease H, RNA-directed DNA polymerase, DNA directed DNA polymerase, RNA polymerase, ribonucleoside triphosphates and deoxyribonucleoside triphosphates with a sample, where the first primer of DNA has a sequence sufficiently complementary to a first template of RNA; the second primer of DNA has a sequence sufficiently complementary to a second template of

IDT Exhibit 1020.7079

DNA, an antisense sequence of a promoter, an antisense sequence of a transcription initiation site which are recognized as substrate by the RNA polymerase;
where a sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

30. A process for amplifying a specific nucleic acid sequence comprising:
combining a first primer, a second primer, avian myoblastosis viral polymerase, E. coli ribonuclease H, bacteriophage T7 RNA polymerase, ribonucleoside triphosphates and deoxyribonucleoside triphosphates with a sample, where the first primer has a sequence sufficiently complementary to a first template of RNA; the second primer has a sequence sufficiently complementary to a second template of DNA, an antisense sequence of a promoter and an antisense sequence of a transcription initiation site which are recognized as substrate by T7 RNA polymerase;
where a sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

31. A process as set forth in claims 1 or 3 further comprising detecting the presence of the specific nucleic acid sequence by a comparison between an amount of amplification of a first sample suspected to contain the nucleic acid sequence and an amount of amplification of a second sample in which the nucleic acid sequence is not present.

32. A process as set forth in claims 29 or 30 further comprising detecting the presence of the specific nucleic acid sequence by a comparison between an amount of amplification of a first sample suspected to contain the nucleic acid sequence and an amount of amplification of a second sample in which the nucleic acid sequence is not present.

33. A process as set forth in claims 1 or 3 further comprising detecting the presence of the specific nucleic acid sequence with a probe which hybridizes to the specific nucleic acid sequence or a product of the specific nucleic acid sequence.

34. A process as set forth in claims 1 or 3 further comprising detecting the presence of the specific nucleic acid sequence using restriction endonucleases and electrophoretic separation.

35. A process as set forth in claims 31 or 32 further comprising detecting the presence of the specific nucleic acid sequence using restriction endonucleases and electrophoretic separation.

36. A process as set forth in claims 31 or 32 further comprising detecting the presence of the specific nucleic acid sequence with a probe which hybridizes to the specific nucleic acid sequence or a product of the specific nucleic acid sequence.

37. The specific nucleic acid sequence amplified by the process set forth in claim 1.

38. The specific nucleic acid sequence amplified by the process set forth in claim 3.

39. The specific nucleic acid sequence amplified by the process set forth in claim 29.

40. The specific nucleic acid sequence amplified by the process set forth in claim 30.

41. A kit for amplifying a specific nucleic acid sequence comprising:
(a) first primer having a DNA sequence sufficiently complementary to a RNA sequence of a first template;
(b) second primer having a DNA sequence complementary to a DNA sequence of a second template;
(c) ribonuclease H;
(d) RNA-directed DNA polymerase;
(e) DNA-directed RNA polymerase;
(f) DNA directed DNA polymerase;
(g) ribonucleoside triphosphates; and
(h) deoxyribonucleoside triphosphates;
where a sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

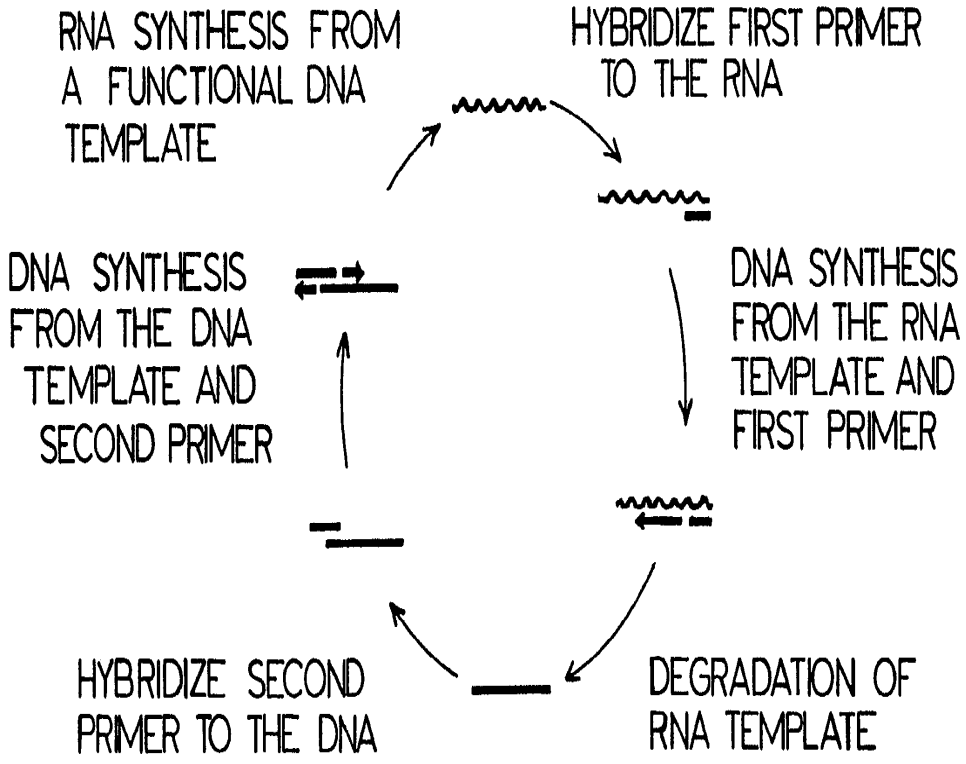
42. A kit for amplifying a specific nucleic acid sequence comprising:
(a) first primer having a DNA sequence sufficiently complementary to a RNA sequence of a first template;
(b) second primer having a DNA sequence sufficiently complementary to a DNA sequence of a second template;
(c) bacteriophage T7 RNA polymerase;
(d) E. coli ribonuclease H;

IDT Exhibit 1020.7080

(e) avian myoblastosis viral polymerase;
(f) ribonucleoside triphosphates; and
(g) deoxyribonucleoside triphosphates;
where a sequence of the first or the second primer is sufficiently complementary to a sequence of the specific nucleic acid sequence, a sequence of the first or the second primer is sufficiently homologous to a sequence of the specific nucleic acid sequence, and a 3′ end of the first primer is oriented towards a 3′ end of the second primer on complementary strands.

IDT Exhibit 1020.7081

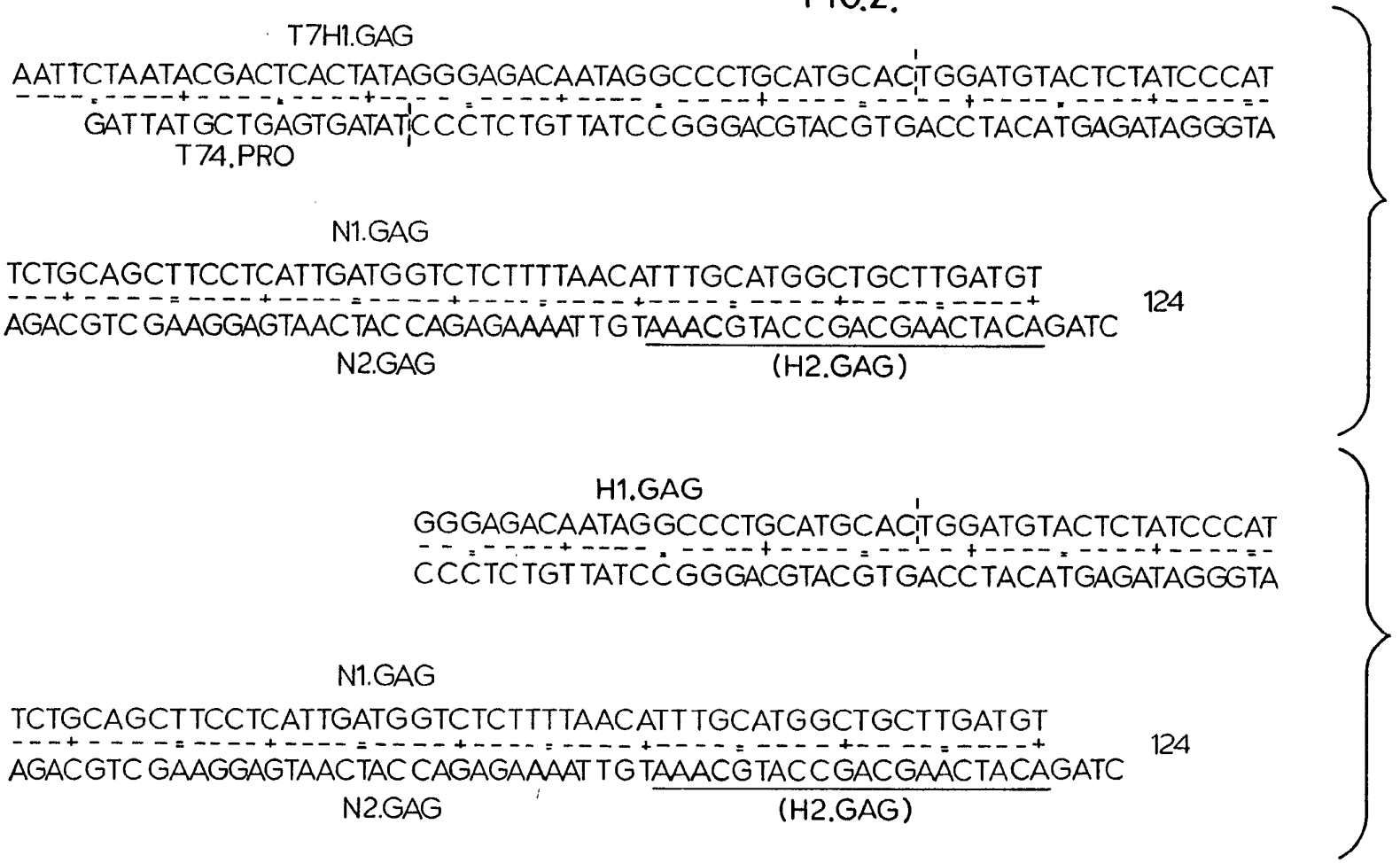
FIG. 1.



IDT Exhibit 1020.7082

A

FIG.2.

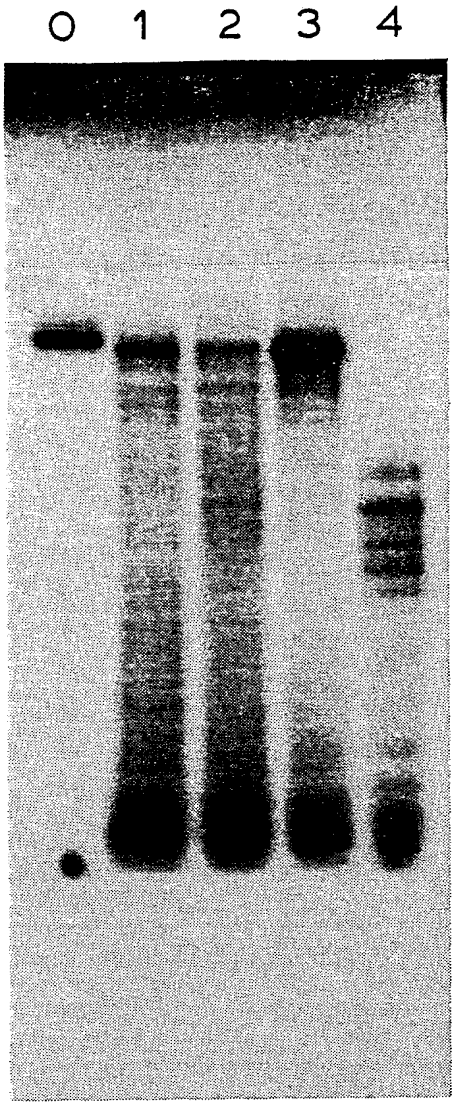


B

EP 0 329 822 A2

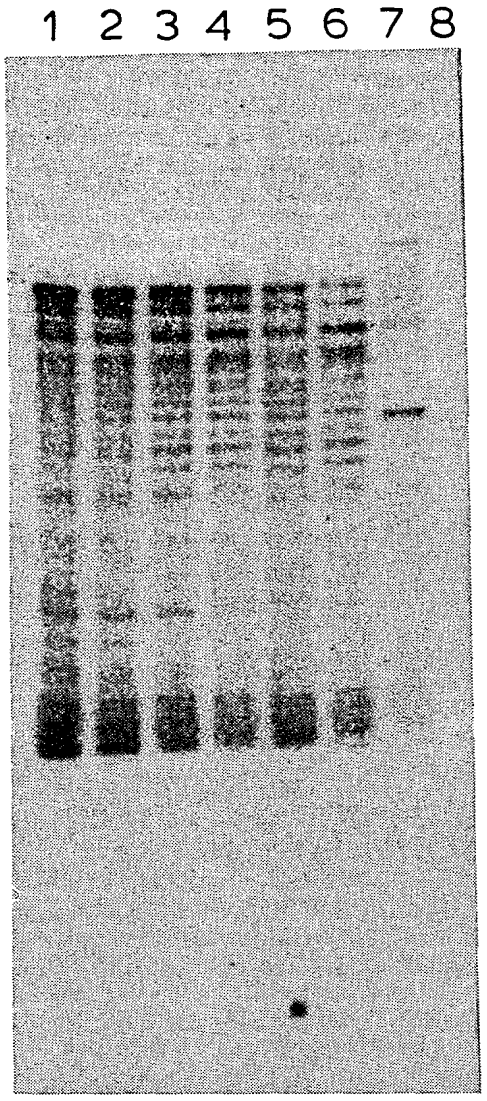
IDT Exhibit 1020.7083

FIG.3.



IDT Exhibit 1020.7084

FIG.4.



IDT Exhibit 1020.7085

FIG.5.

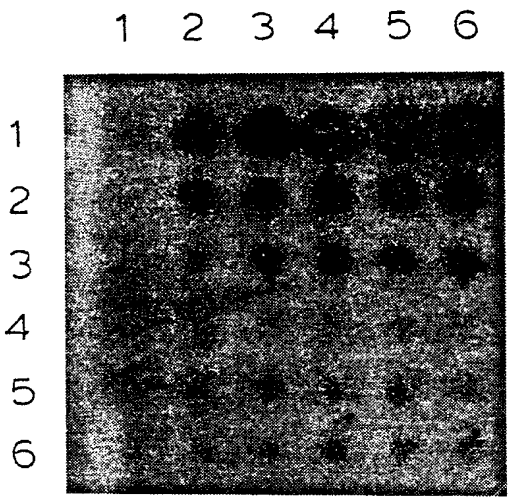
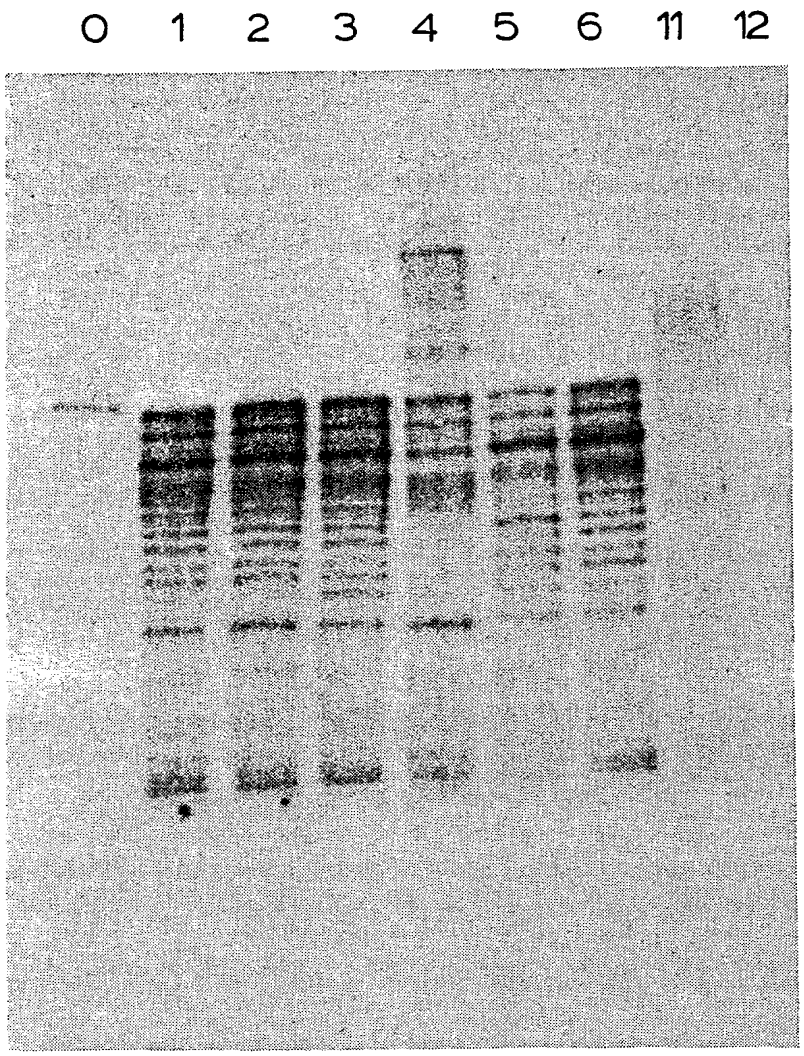


FIG.6.



IDT Exhibit 1020.7087

Espacenet

Bibliographic data: EP2451973 (A1) — 2012-05-16

METHOD FOR DIFFERENTIATION OF POLYNUCLEOTIDE STRANDS

Inventor(s): SOLDATOV ALEKSEY [DE]; BORODINA TATIANA [DE]; LEHRACH HANS [DE] ± (SOLDATOV, ALEKSEY, ; BORODINA, TATIANA, ; LEHRACH, HANS)

Applicant(s): MAX PLANCK GESELLSCHAFT [DE] ± (MAX-PLANCK-GESELLSCHAFT ZUR FOERDERUNG DER WISSENSCHAFTEN E.V, ; MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V)

Classification: **- international:** ***C12Q1/68***
- cooperative: **C12Q1/6848**; **C12Q1/6853**; **C12Q1/6855**; **C12Q1/6869**; **C12Q1/6876** more

Application number: EP20100737282 20100706 Global Dossier

Priority number(s): WO2010EP04399 20100706 ; EP20090008808 20090706 ; US20090213730P 20090708 ; EP20100737282 20100706

Also published as: EP2451973 (B1) EP2272976 (A1) US2012237943 (A1) US2015105299 (A1) US2016053315 (A1) more

Abstract not available for EP2451973 (A1)
Abstract of corresponding document: EP2272976 (A1)

Objective of the present invention is to provide a method for keeping of directional information in double-stranded DNA. We suggest to convert polynucleotide into a hybrid double-stranded DNA. One particular strand of this hybrid double-stranded DNA should be synthesised using at least one modified nucleotide. Thus, this particular strand would contain modified nucleotides along the whole length. Density of directional markers would not depend on the length of polynucleotides. Any internal fragments of the hybrid double-stranded DNA would have directional information. When it is necessary the modified strand may be easily degraded or separated from the other strand. It was found that such hybrid double-stranded DNA may be easily generated in a number of molecular biology tasks and may be used for molecular cloning, library preparation and strand separation.

IDT Exhibit 1020.7088

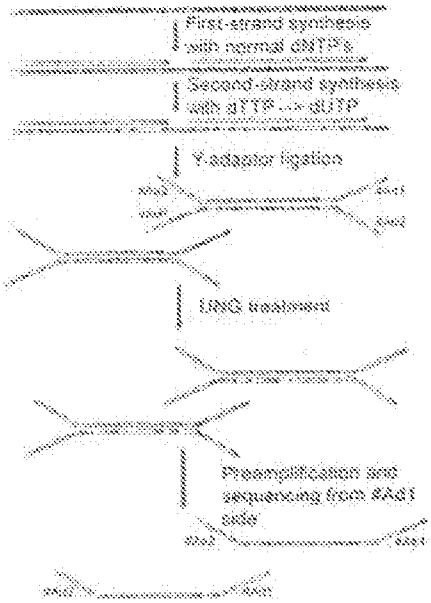


Figure 4

IDT Exhibit 1020.7089

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 January 2011 (13.01.2011)

PCT

(10) International Publication Number
WO 2011/003630 A1

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

(21) International Application Number:
PCT/EP2010/004399

(22) International Filing Date:
6 July 2010 (06.07.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09008808.9	6 July 2009 (06.07.2009)	EP
61/213,730	8 July 2009 (08.07.2009)	US

(71) Applicant *(for all designated States except US)*: **MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E. V.** [DE/DE]; Hofgartenstraße 8, 80539 München (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SOLDATOV, Aleksey** [RU/DE]; Taylorstrasse 12b, 14195 Berlin (DE). **BORODINA, Tatiana** [RU/DE]; Taylorstrasse 12b, 14195 Berlin (DE). **LEHRACH, Hans** [AT/DE]; Terrassenstr. 31, 14129 Berlin (DE).

(74) Agent: **ARTH, Hans-Lothar**; ABK Patent Attorneys, Jasminweg 9, 14052 Berlin (DE).

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

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

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: METHOD FOR DIFFERENTIATION OF POLYNUCLEOTIDE STRANDS

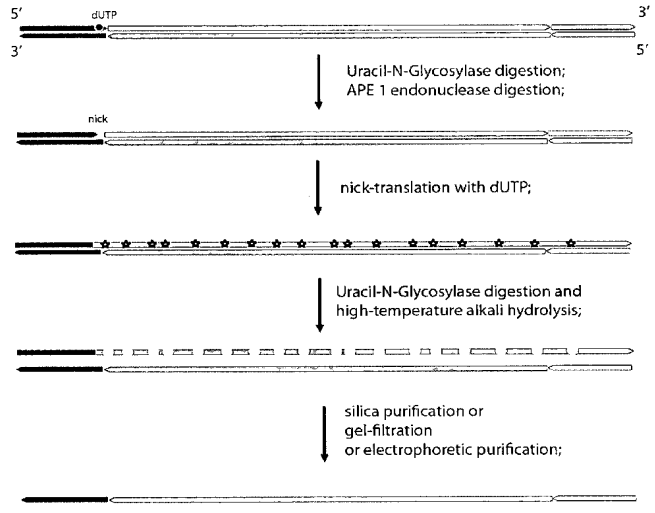


Figure 3A

(57) **Abstract**: Objective of the present invention is to provide a method for keeping of directional information in double-stranded DNA. We suggest to convert polynucleotide into a hybrid double-stranded DNA. One particular strand of this hybrid double-stranded DNA should be synthesised using at least one modified nucleotide. Thus, this particular strand would contain modified nucleotides along the whole length. Density of directional markers would not depend on the length of polynucleotides. Any internal fragments of the hybrid double-stranded DNA would have directional information. When it is necessary the modified strand may be easily degraded or separated from the other strand. It was found that such hybrid double-stranded DNA may be easily generated in a number of molecular biology tasks and may be used for molecular cloning, library preparation and strand separation.

WO 2011/003630 A1

Method for differentiation of polynucleotide strands

Description

Cloning of single-stranded RNA and DNA molecules is much easier if they are transformed into double-stranded DNA. The invention relates to methods of preserving information about the direction of the original single-stranded molecules after being transformed into a double-stranded form.

Single-stranded RNA and DNA molecules play an important role in molecular biology. A number of viruses have single-stranded genomes. Most RNAs are present in single-stranded form. Recent studies have demonstrated an unexpected complexity of eukaryotic transcriptomes. A transcriptome is the set of all RNA transcripts produced in one or a population of cells. In addition to classical mRNAs (messenger ribonucleic acids) which cover approximately 1.5 percent of the genome in most higher eukaryotes, a large number of non-coding RNAs (ribonucleic acids) with widely varying expression levels have been identified. The biological functions of these novel transcripts are largely unknown and represent a new research area requiring high throughput transcriptome studies.

The polarity of the transcript is important for correct annotation of novel genes because it provides essential information about the possible function of a gene both on the RNA (structure and hybridization to other nucleic acid molecules) and the protein level. In addition, many genomic regions give rise to transcripts from both strands simultaneously. Antisense transcription is characteristic for eukaryotic genes and is thought to play an important regulatory role. Overlapping genes are common for compact genomes of prokaryotes and lower eukaryotes. Knowledge of transcript orientation helps to resolve the problem of colliding transcripts and to determine correctly gene expression levels in the presence of antisense transcription.

Background of the invention

From the art several methods are known that preserve directional information using distinguishable adapters for different ends of double-stranded DNA.

Adapters may be ligated directly to single-stranded RNA molecules.

Special composite primers may be used for first-strand cDNA synthesis: Herein the 5'- part of the primer provides a directional information, the 3'-part (oligo(dT)) or random nucleotides) is able to form a duplex with the RNA. Recently, first-strand cDNA synthesis was performed from a tagged random hexamer primer and second strand synthesis from a DNA-RNA template-switching-primer.

Special homopolymeric adapters may be synthesised at the 3'-end of DNA molecules by Terminal Nucleotide Transferase.

DNA-RNA template-switching is used to attach a specific adapter to 3'-end of the DNA copy of the original single-stranded RNA molecule.

The main disadvantage of the adaptor-based approach is that only the ends of the clones preserve directional information. The problems are:

(i) strand-specific operations (directional immobilization on a solid phase; production of a single-stranded DNA for hybridization, etc.) should act on the adaptor sequence. For long clones the efficiency of such operations is normally low. For example, strand separation where one of the strands contains a biotin molecule at the 5'-end is getting complicated if the length of the duplex is higher than 1kb.

(ii) after fragmentation internal parts of the clones lose the directional information.

A special approach which is able to keep directional information in any part of the sequence was suggested recently: All cytidine residues in the RNA were converted to undines by bisulfite treatment prior to cDNA synthesis. The problem is that this approach is laborious and also leads to the loss of about 30% of uniquely matched sequencing reads because a part of the genome complexity is lost during the transformation from the four-bases into the three-bases code.

A lot of biologically important nucleic acids function in a single-stranded form where the polarity is extremely important. But for basic molecular cloning procedures single-stranded polynucleotides should be converted into double-stranded DNA molecules, because

* most of restriction enzymes work only on double-stranded DNA;
* ligation of double-stranded molecules is much simpler in comparison with the ligation of single-stranded molecules;
* the result of PCR is a double-stranded DNA.

At the moment most of the protocols keep directional information in the form of distinguishable adaptors on the ends of double-stranded molecules (Fig. 1B).

US 2007/0281863 discloses a method for characterizing the amounts of nucleic acids, including plus/minus determinations. To this end the mRNA is transcribed into double stranded cDNA and two different RNA promoters are integrated in both cDNA strands respectively. By using the different RNA promoters in the presence of allylamine-UTP nucleotides plus or minus RNA with modified allylamine-UTP nucleotides is generated. The modified allylamine-UTP nucleotides can be labeled with NHS-Cy3/5 and are then hybridized in different nucleic acid arrays.

The main disadvantage of this approach is that the RNA cannot be used for high-throughput complementary DNA sequencing (RNA-Seq). Direct cDNA sequencing (RNA-Seq) is a new tool for whole-transcriptome analysis. Second generation sequencing machines have increased sequencing throughput by about two orders of magnitude, making global transcriptome sequencing feasible. Since sequencing costs are constantly decreasing (contrary to those of microarrays) cDNA sequencing will capture a considerable portion of transcriptome analyses in the future.

The RNA-Seq procedure is simple, has a large dynamic range and high sensitivity, and can unequivocally identify splicing and RNA editing products as well as allele-specific transcripts. RNA-Seq provides a number of further advantages over other high throughput approaches like microarray hybridization, gene-specific and tiling arrays or SAGE-analyses. The depth of RNA-Seq analyses is flexible, providing a dynamic range typically an order of magnitude greater than one can achieve with hybridization arrays. The digital character of the RNA-Seq data permits to compare and pool results from different laboratories. No prior information about transcripts sequences is required, allowing detection of novel transcripts. It is possible to estimate the absolute level of gene expression and to study structure of transcripts.

However, the weakness of RNA-Seq was the inability to determine the polarity of RNA transcripts without laborious modification of the protocol. The method according to the invention solves this problem allowing the strand specific differentiation with an easy, robust and cheap procedure. The method is less susceptible and more robust compared to the method described in US 2007/0281863 because there is no need to transcribe the mRNA first in cDNA then inserting two different promoters in the cDNA and then using two different polymerases to transcribe the cDNA again in RNA including an additional step to label the marked RNA with NHS-Cy3/5.

Thus, the modified allylamine-UTP RNA according to US 2007/0281863 can only be used in hybridization arrays that are very expensive, less sensitive and cannot detect

novel transcripts. Additionally, the method is laborious including the integration of promoters and different polymerases and works with RNA as an end product, which is very susceptible to degradation because of abundant RNases and can only be used in expensive hybridization arrays that are less sensitive then e.g. RNA-Seq.

Further disadvantages include that many basic molecular cloning procedures are performed on double stranded cDNA, which are necessary for many subsequent analyzing methods. The method according to US 2007/0281863 generates a double stranded cDNA with two promoters as an intermediate product. However, these intermediate cDNAs are not preferred for cloning procedures,

As we already explained above the main disadvantage of adaptor-based approaches is that only the ends of the clones preserve directional information. The integration of promoters in the cDNA is basically the same procedure. During cloning procedures the promoters can be changed, damaged etc., especially because most cloning procedures modify the ends of the cDNA. The subsequent strand specific transcription of the cDNA with promoter specific polymerases is thus not very robust and the strand specific information can be lost during molecular cloning procedures. Taking into account the susceptibility of RNA to degradation, the method according to the instant application is more functional, i.e. the RNA is transcribed directly into cDNA and the strand specific information is already integrated in this step. Furthermore, the inventive method is much more robust, e.g. the cDNA already contains the strand specific information along the whole length of the transcript and no molecular cloning procedures can compromise this information. Consequently, the inventive method is convenient, reliable and highly reproducible.

WO 97/12061 discloses a method for characterizing nucleic acid molecules, comprising synthesizing DNA with a nucleic acid matrix, primers, polymerase, four canonical deoxynucleotides and at least one non-canonical deoxynucleotide. The DNA is then contacted with N-glycosylase and the abasic sites are treated in a way to leading to breakage of the phosphodiester backbone. The resulting DNA fragments are separated according to their size.

This method characterizes DNA according to their size by using non-canonical deoxynucleotide to label one synthesized DNA strand, but the method is not able to preserve the information about the direction of the original single-stranded molecules after being transformed into a double-stranded form.

Description of the invention

The objective of the present invention is to provide a method for keeping the directional information in double-stranded DNA. Surprisingly, it was found that this objective can be reached by the inventive method as disclosed by the description, the figures and the examples. A polynucleotide strand is converted into a hybrid double-stranded DNA. One particular strand of this hybrid double-stranded DNA should be synthesised using at least one modified nucleotide. Thus, this particular strand would contain modified nucleotides along the whole length of the strand. The density of directional markers would not depend on the length of hybrid double-stranded DNA. Any internal fragment of the hybrid would have directional information. It was found that such hybrid double-stranded DNA may be easily generated in a number of molecular biology tasks and may be used for molecular cloning, library preparation and strand separation. If necessary, one of the strands (modified or non-modified) may be easily degraded or separated from the other strand. Also some modifications might prevent the strand containing modified nucleotides from participation in downstream applications.

The directional library construction method and the method of synthesis of single-stranded DNA according to the invention are easy to perform, inexpensive and highly reliable and thus overcome the shortcomings of the state of the art approaches.

The inventive method for preserving information about the direction of single-stranded nucleic acid molecules during molecular cloning operations with double-stranded DNA is provided. One of the strands of the double-stranded DNA is synthesized by using the other strand as a matrix and contains at least one modified dNTP. Thus, one strand of double-stranded DNA would contain unmodified nucleotides and the other strand at least one modified nucleotide along the whole length (Fig. 1C).

The new method of keeping of direction of polynucleotides is fully compatible with methods known from the art. Double-stranded DNA molecules may have simultaneously distinguishable primers and distinguishable strands (Fig. 1D). Fig. 2A shows how molecules with distinguishable strands may be converted into molecules with distinguishable primers. In an opposite operation the conversion of molecules with distinguishable primers into molecules with distinguishable strands is shown in Fig. 2B.

The main advantages compared to the adaptor-based method are:

(i) the density of modified nucleotides (directional markers) does not depend on the length of the double-stranded molecule. This is important for strand separation or strand-specific digestion of long DNA molecules.

(ii) any internal fragment of the clone carries directional information. After fragmentation internal parts of the clones do not lose the directional information.

Two examples demonstrate the convenience of the inventive method for differentiation of polynucleotide strands.

The method of synthesis of a single-stranded DNA is described in Fig. 3.

There is a number of methods of synthesis of single-stranded DNA relying on the difference between flanked adaptors:

* asymmetric PCR may be performed with the different amounts of PCR primers;

* the 5' end of one of the strands may contain thionucleotides which protect this strand from 5' to 3' exonucleases;

* protective thionucleotides may be introduced into the 3'-end part of only one strand to protect it from 3' to 5' exonucleases;

* the 5'-part of one of the strands may bear biotin for attachment of this strand to streptavidine-coated solid particles.

The problem is that all these methods proved to be unstable and show low efficiency for molecules larger than 1 kb.

Fig. 3 shows two possible ways for generating single-stranded DNA using the inventive method for differentiation of polynucleotide strands.

In a first embodiment of the present invention (Fig. 3A) one strand of PCR product is substituted by a strand where dTTPs are substituted by dUTPs. After UNG treatment and APE 1 digestion (or alkalic hydrolisis) this strand is converted into a number of short single-stranded fragments which may be easily separated from the long unmodified strand by electrophoresis, anion-exchange chromatography, silica-based purification or gel-filtration.

In a second embodiment of the present invention (Fig. 3B) one strand of the PCR product is substituted by a strand where a significant percentage of the dTTPs is substituted by biotin-associated dUTPs, respectively biotin-associated dTTPs. The

concentration of biotin in a modified strand does not depend on the length of the double-stranded DNA, so the efficiency of the purification on streptavidine-coated particles should not depend on the length of the fragment.

The key steps of the method according to the invention include the generation of a cDNA strand including at least one modified nucleotide that in the course of the method is recognized by a selection factor which allows for selection of the strand containing the at least one modified nucleotide from the strand containing the corresponding unmodified nucleotide or nucleotides.

The key steps of the inventive method are to be read in the following order:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on said polynucleotides; and

III) Performing strand-specific selection by using a selection factor for selecting strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide.

The key steps of the inventive method can also be read in the following order:

I) Providing polynucleotide templates, 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 polynucleotides; and

III) 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.

The term *strand specific differentiation* refers to the ability to differentiate in the double stranded form between the template and the synthesized matching strand.

The term *specific strand synthesis* refers to synthesis of the matching strand containing the at least one modified nucleotide from the template.

Another embodiment of the present invention relates to single-stranded RNA. Thus, if the polynucleotide template is a single-stranded RNA a second cDNA strand has to be synthesized. For this purpose the at least one modified dNTP has to be removed from the solution and has to be replaced by the corresponding unmodified dNTP or dNTPs or analogues thereof. By the use of all four unmodified dNTPs a second cDNA strand is generated.

In this case the key steps according to the inventive method read as follows:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on said polynucleotides;

III) Performing second-strand synthesis with all four unmodified dNTPs or analogues thereof after removing the at least one modified dNTP; and

IV) Performing strand-specific selection by using a selection factor for selecting strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide.

As said before, if the polynucleotide templates are single-stranded RNA a second-strand cDNA has to be synthesized after the exchange of the at least one modified dNTP by the corresponding unmodified dNTP or dNTPs. If this double-stranded cDNA shall undergo cloning it will be advantageous to perform one or more of the modifications for the molecular cloning operation as described before. By combining these conditions the key steps of the invention read as follows:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on said polynucleotides;

III) Performing second-strand synthesis with all four unmodified dNTPs or analogues thereof after removing the at least one modified dNTP;

III) 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

V) Performing strand-specific selection by using a selection factor for selecting strands containing the at least one modified nucleotide from strands containing the corresponding unmodified nucleotide.

The initially provided polynucleotide templates may be a RNA such as mRNA, tRNA, rRNA, hnRNA, snRNA, snoRNA, scaRNA, asRNA, micro RNA, siRNA, short transcripts, single-strand oligonucleotides, single-strand polynucleotides, synthetic polynucleotides and semisynthetic polynucleotides as well as physiological or externally induced modifications thereof, wherein mRNA is preferred.

mRNAs are protein-coding RNA. tRNAs are transfer RNAs responsible for providing correct amino acids in peptide (protein) synthesis in ribosomes. rRNAs are ribosomal RNAs, being a constitutive element of ribosomes. hnRNAs (heterogeneous nuclear RNA) is a precursor mRNA in the nucleus of eukaryotic cells. It is often named also pre-mRNA or precursor mRNA. snRNA (small nuclear RNA) is responsible in the eukaryotic nucleus for the processing of hnRNAs (spliceosomes). snoRNA (small nucleolar RNA) occurs in the nucleolus of eukaryotic cells. scaRNA (small Cajal body-specific RNA) is strongly related to snoRNA and occurs in Cajal bodies. asRNA (antisense RNA) plays an important role in regulating gene expression. microRNAs are involved in the regulation of cellular processes such as apoptosis and cell death. siRNA (small interfering RNA) is processed from cellular RNA into fragments with a length of ca. 22 base pairs by the enzyme Dicer and integrated into the enzyme complex RISC (RNA-induced silencing complex). RISC can thus regulate the expression of DNA segments or RNAs by switching off the respective coding region through analogous base pairing (silencing). This method by also used in gene therapeutic approaches. Retroviruses use RNA as the carrier of their genetic information.

Short transcripts refer to small RNA fragments with up to 20-40 base pairs.

Alternatively, the polynucleotide templates used within the inventive methods may be a DNA such as siDNA, DNA copies from retroviral or hepadnaviral RNA, single-stranded DNA, synthetic polynucleotides and semisynthetic polynucleotides as well as physiological or externally induced modifications thereof, wherein single-stranded DNA is preferred.

It shall be pointed out that in this invention the term *single-stranded DNA* can also refer to the first-strand cDNA derived from a RNA as templates.

siDNA (small interfering DNA) is the corresponding DNA of siRNAs and mainly fulfils the same tasks in the cell. Single-stranded DNA is a single-stranded type of satellite DNA, occurring among others in certain soil-dwelling bacteria species.

According to the invention also double-stranded DNA can be the polynucleotide templates. Specific primers and optionally adjuvants are needed to allow the generation of a cDNA strand against a particular strand of the double-stranded DNA (see below and Fig. 2B). Herein, the double-stranded DNA is destabilized in the region of hybridization of said specific primer for allowing annealing of the specific primer. Such a destabilization can be achieved for example by Uracil-N-Glycosylase digestion of deoxyuridines.

In a specific embodiment the double-stranded DNA templates are PCR products.

Sources for RNAs and DNAs can be eukaryotes (animals including humans and other mammals, plants, fungi), prokaryotes (bacteria, archaea), viruses, synthetic or semisynthetic constructs wherein these constructs or semisynthetic constructs may mirror at least partially a naturally occurring nucleic acids or may have a widely artificial sequence.

Said RNAs and DNAs can be isolated by well-known methods (Molecular Cloning: A Laboratory Manual, Third Edition. J. Sambrook, D. Russell).

The term *classic dNTPs*, respectively *unmodified dNTPS* refers to dATP (deoxyadenosine triphosphate), dCTP (deoxycytidine triphosphate), dGTP (deoxyguanosine triphosphate) and dTTP (deoxythymidine triphosphate). These four dNTPs (deoxynucleoside triphosphates) are usually used in nature for the generation of DNA strands.

It shall be pointed out that the plural form *classic dNTPs* refers to the total of dATP, dCTP, dGTP and dTTP. A referral in a quantifying sense to one or more of these classic dNTPs is not intended.

The term *analogues of dNTPs*, respectively *dNTPs and analogues thereof* refers to any molecule suitable for replacing a dNTP in presence and behavioural characteristics in a respective polynucleotide-containing solution and particularly in

its matching characteristics to be able to undergo a base pair matching identical or similar to the classic dNTP it replaces. In particular, such analogues of dNTPs should be able to build the same or similar selective hydrogen bonds that allow for a selective pairing of base pairs (2 hydrogen bonds between A/T, T/A, A/U or U/A, and 3 hydrogen bonds between G/C and C/G), thus allowing for a correct translation of a polynucleotide sequence into its antisense sequence. Furthermore, it should be sterically similar to its corresponding classic dNTP so that it allows integration into the base pair sequence. Not only the nucleotide can be varied in such an analogue, but also the saccharide moiety (usually a ribose or deoxyribose) or the phosphate backbone. For example, the phosphate backbone may be partially or entirely substituted by a sulphonated backbone and/or by a phosphonate backbone. Also combinations of such variations are subsumed under this term.
Moreover the analogues of the dNTPs cannot be used for stand specific selection or strand differentiation. Thus the analogues of the dNTPs do not react under the conditions which are used for the modified dNTPs such as degradation (e.g. dUTP, heat), differentiation by e.g. electrophoretic methods, adsorption or binding to suitable counterparts such as biotin – avidin. Thus the analogues of the dNTPs are used as substitutes for the classic or canonical dNTPs and not as substitutes for the modified dNTPs.

The quantitative term *sufficient amount* herein refers to any amount, respectively concentration of an agent which allows for a qualitative and quantitative implementation of the method according to the invention in respect of the expected amounts which shall be needed for this purpose. Usually, the respective agents, in particular dNTPs, are added in exceeding amounts in order to ensure that all possible copies can be effected qualitatively and quantitatively.

The term *template* refers in molecular biology to a DNA or RNA strand which sets the base pair sequence to which a corresponding strand with an antisense sequence shall by polymerized. In a broader sense it may also refer to a RNA encoding a sequence for translation of this sequence into an amino acid sequence such as a peptide or a protein in ribosomes. The term *templates* in plural refers to the total sum of polynucleotides provided, each single template representing a single polynucleotide. Templates also refers to the originally provided sum of polynucloetid strands. to A referral in a quantifying sense to one or more of these templates is not intended.
The term *template* refers also to the original single stranded polynucleotide strand, where the preservation of information about the direction is desired after being

transformed into a double-stranded form. If the inventive method is applied to double stranded polynucleotides (e.g. DNA/RNA, DNA/DNA with distinguishable adapters), in this case the template can be freely chosen from both strands, depending for which strand the preservation of information about the direction is desired.

The term *matching strand* refers to the complementary strand to the template.

The term *first-strand synthesis* refers to the synthesis of the first strands using the templates as matrix for synthesizing the complementary strands. The nucleotide sequence of the first strand corresponds to the sequence of the matching strand.

The term *second-strand synthesis* refers to the synthesis of the second strands that uses the first strands as matrix for synthesizing the complementary strands, i.e. the second strands. The nucleotide sequence of the second strand corresponds to the sequence of the template.

As used herein the term "matrix" or "matrices" refers to any polynucleotides which can serve as starting material for polymerase-reactions.

As used herein the term "at least one nucleotide" or "at least one modified nucleotide" is not meant as a referral in a quantifying sense to one or more nucleotides but refers to a species of modified dNTPs. Thus, this term refers to a plurality of dNTP molecules of the same kind. Thus using "one" modified nucleotide does not mean that a single modified nucleotide is used. Using "one" modified nucleotide means that one kind or one sort of nucleotides are used which are modified such as modified dATP or modified dCTP or modified dGTP or modified dTTP.

The term "replaced by a corresponding modified dNTP" refers to a substitution of one of the classic or canonical dNTP species in the reaction mixture used for strand synthesis by a modified dNTP species able to fulfil the same biological function. This means that the substitution is not performed on the completed strand after strand synthesis. The term "same biological function" refers to modified dNTPs able to preferably build the same or similar selective hydrogen bonds that allow for a selective pairing of base pairs (2 hydrogen bonds between A/T, T/A, A/U or U/A, and 3 hydrogen bonds between G/C and C/G). Furthermore, it should preferably be sterically similar to its corresponding classic dNTP so that it allows integration into the base pair sequence. Not only the nucleotide can be varied in such a modified dNTP, but also the saccharide moiety (usually a ribose or deoxyribose) or the phosphate backbone.

The term *cDNA* stands for complementary deoxyribonucleic acid. It refers to a DNA which is generated as a copy of a RNA (usually by means of Reverse Transcriptase) or from a single DNA strand (usually by means of a DNA polymerase). Subsequently, a second cDNA strand may be generated by using the first cDNA strand as a matrix (usually by means of a DNA polymerase).

Enzymes with reverse transcriptase (RNA-dependent DNA polymerase) activity are able to transcribe a RNA sequence into a corresponding cDNA sequence. This group of enzymes includes among others HIV-1 reverse transcriptase from human immunodeficiency virus type 1 (PDB 1HMV), M-MLV reverse transcriptase from the Moloney murine leukemia virus, AMV reverse transcriptase from avian myeloblastosis virus and telomerase reverse transcriptase fulfilling a role in the maintenance of telomeres of eukaryotic chromosomes as well as commercially available forms such as "M-MLV reverse transcriptase RNase H Minus", "SuperScript® II", "SuperScript® III" or Tth polymerase. Among this group M-MLV reverse transcriptase RNase H Minus, SuperScript® II and SuperScript® III polymerases are preferred.

The terms *primers* and *random primers* refer to oligonucleotides which shall serve as starting point for DNA-replicating enzymes such as DNA polymerases. They can consist of a few base pairs, but in selected cases can reach up to 2.000 base pairs. Preferred are 5 to 50 base pairs. DNA polymerases need a hydroxy group for starting transcription. This hydroxy group is provided by the primer with its 3' hydroxy end. Often, the DNA sequence to be investigated is known. Hence suitable primers can be synthesized precisely according to the antisense sequence of aforesaid polynucleotide sequence. If this polynucleotide sequence is not or only incompletely known so-called random primers can be used. In such a method a large number of primers with a random base pair sequence are provided. One or more of these random primers statistically anneal to the polynucleotide sequence that shall be transcribed. Thus corresponding transcripts of different length can be obtained, depending where the random primer anneals to the polynucleotide sequence of interest. Alternatively, also composite primers with a random 3' part can be used. A further variation corresponding to the inventive method is the use of primers which are complementary to a homo-oligomeric tail synthesized by terminal transferase on the polynucleotide which serves as a matrix.

If the polynucleotide templates are RNA with a polyA tail it is preferred that said primers are oligo(dT), oligo(dU) or oligo(U) primers or composite primers with oligo(dT), oligo(dU) or oligo(U) region on 3' part. Alternatively, specific primers complementary to said single-stranded RNA templates can be used.

If double-stranded DNA templates are used it is likewise possible to mark one particular strand of the double-stranded DNA with at least one modified nucleotide. Herein specific primers complementary to a particular strand of said double-stranded DNA can be used. In this case RecA protein may facilitate the hybridization of said primers to the double-stranded DNA templates. Specific primers may appear as a result of introducing a nick in a particular strand of said double-stranded DNA templates. Such nicks can for example result by the use of nicking endonuclease. Alternatively, such nicks may result from the digestion of the at least one modified nucleotide in a particular strand of said double-stranded DNA templates.

In a specific embodiment the double-stranded DNA templates are PCR products. In this case, the recognition site of a nicking exonuclease or at least one modified nucleotide is located in one of the PCR primers. Herein it is preferred that said modified nucleotide is deoxyuridine.

If double-stranded DNA templates are used it is preferred that said polymerase is a DNA-dependent DNA polymerase with 5' to 3' exonuclease activity and said strand-synthesis on such double-stranded DNA templates are effected by means of a nick-translation.

In analogy, the polynucleotide templates may be also a DNA::RNA hybrid. Hybrid herein means that a DNA strand is matching with a RNA strand having a corresponding antisense sequence. For example, but not limiting, it is possible that such a DNA::RNA hybrid is the product of a first-strand synthesis with unmodified nucleotides generated on single-strand RNA as templates. For generating a cDNA of the DNA strand specific primers complementary to a particular DNA region of said double-stranded DNA::RNA hybrid have to be used. It is preferred that said primers are complementary to a homo-oligomeric tail synthesized by terminal transferase. In another preferred embodiment this specific primer is a template-switch-primer. Such template-switch primers bind to the extra random nucleotides attached to the 3' end of the newly synthesized cDNA strand by using SuperScript® II enzyme (Invitrogen). This allows for full-length reverse transcription of the RNA strand. If the cDNA strand shall be generated against the RNA strand of the DNA::RNA hybrid said polymerase is DNA-dependent DNA polymerase with 5' to 3' exonuclease activity.

If the cDNA strand shall be generated against the DNA strand of the DNA::RNA hybrid the specific primers may appear as a result of introducing a nick into the RNA strand of said DNA::RNA hybrid. Such a nick may be set by using the RNA-specific enzyme RNase H. In a preferred embodiment RNase H, DNA polymerase I and

E.coli ligase are used simultaneously. DNA ligases are enzymes able to ligate DNA strands by building an ester bond between the phosphate residue and the saccharide deoxyribose in the backbone of the DNA strands.

For second-strand cDNA synthesis using the first strand cDNA as a matrix an enzyme with DNA polymerase activity is needed. Also if the original polynucleotide is a single-stranded DNA and only one cDNA strand is generated such an enzyme with DNA polymerase activity is needed. DNA-dependent DNA polymerases are usually enzymes which generate a DNA copy strand from a DNA matrix. This group includes DNA polymerase I, II, III holoenzyme and IV, Taq polymerase, Tth polymerase, Pfu polymerase, wherein DNA polymerase I and Taq polymerase are preferred. This enzyme with polymerase activity does not have to be the same as used optionally for amplifying the remaining intact cDNA.

If the cDNA strand shall be generated against the RNA strand of the DNA::RNA hybrid the procedure is similar to that for generating a cDNA strand against one strand of a single-stranded RNA.

The term *modified dNTP* refers to any molecule suitable for substituting exactly one corresponding classic dNTP in presence and behavioural characteristics in a respective polynucleotide-containing solution and particularly in its matching characteristics to be able to undergo a base pair matching identical or similar to the classic dNTP it replaces. Additionally, this modified dNTP must be suitable either for a specific degradation in which selectively this modified dNTP is degraded by a suitable degrading agent, thus rendering the DNA strand containing at least one of this degraded dNTP essentially unfit for amplification and/or hybridization, or it must render the DNA strand containing this modified dNTP eligible for specific removal of this DNA strand from the solution, or at least allowing a local and thus functional separation of polynucleotide strands. Such a removal respectively separation can be achieved by molecules or particles binding selectively to such exchange dNTPs, thus selectively only to one polynucleotide strand. Removal can occur by purification, adsorption to suitable structures or surfaces, binding to suitable selective structures such as antibodies, by selective degradation through certain chemicals, heat or radiation, or by centrifugation, magnetic or electrophoretic methods, isoelectric focusing. Also any kind of immobilization can be perceived.

The term *canonical* is used to refer to the four common nucleic acid bases adenine, cytosine, guanine and thymine that are commonly found in DNA or to the respective deoxyribonucleosides, deoxyribonucleotides or 2′-deoxyribonucleoside-5′-

triphosphates that contain a canonical base, The tern "non-canonical" is used to refer to nucleic acid bases in DNA other than the four canonical bases, or to the respective deoxyrinonucleosides, deoxyribonucleotides, or 2´ -deoxyribonucleoside- 5´-triphosphates that contain a non-canonical base. Although uracil is a common nucleic acid base in RNA, uracil is a non-canonical base in DNA. "Non canonical bases" are found in nucleic acids as a result of incorporation of non-canonical nucleotides or as a result of modification of existing bases (canonical or non-canonical).

Different purification methods may be used for removal of modified and unmodified dNTPs from the solution: electrophoretic separation, salt-alcohol precipitation, *anion*-exchange chromatography, silica purification, gel filtration. Purification on silica and gel-filtration are preferred.

Degradation respectively degradable herein means that a deoxynucleotide is specifically modified by addition of at least one enzymatic, chemical, heat, radiation (X-ray, UV, microwave) or optical methods such as lasers or combinations thereof in such a way that the resulting deoxynucleotide does not serve anymore for hybridization, cloning or as a suitable matrix in an amplification system such as PCR.

Analogously, the plural form *modified dNTPs* refers to the total of the respective modified dNTPs. A referral in a quantifying sense to one or more of these modified dNTPs is not intended.

The term *at least one modified dNTP* refers to the option that not only one dNTP is exchanged. According to the invention either one dNTP such as dATP, dCTP, dGTP or dTTP, or two dNTPs can be exchanged in the combinations dATP + dCTP, dATP + dGTP, dATP + dTTP, dCTP + dGTP, dCTP + dTTP or dGTP + dTTP, or three dNTPs can be exchanged in the combinations dATP + dCTP + dGTP, dATP + dCTP + dTTP, dATP + dGTP + dTTP or dCTP + dGTP + dTTP, or all four dNTPs are exchanged in the combination dATP + dCTP + dGTP + dTTP. It is understood that also analogues of these dNTPs as defined before shall be subsumed under the respective dNTPs. Also possible quantitative combinations thereof shall be subsumed under this combinational scheme.

The term *selection*, *selecting* or *strand specific selection* refers to selecting in ds (double stranded) polynucleotides the strand with at least one modified nucleotide to distinguish the template from the matching strand. The selected strand with the at least one modified nucleotide in ds (double stranded) polynucleotides can be

selectively removed or specifically destroyed, while the strand without modified nucleotide is not impaired preferably throughout its whole length. The removed or destroyed or destroyable strand with the modified nucleotide can be used separately. If the strand with the modified nucleotide is not removed and/or destroyed, it can remain in the same solution with the matching strand. The modified nucleotide in the strand prevents full-length amplification of the strand.

There are several principles by which a selection factor can be used to select between a strand marked with at least one modified nucleotide and a strand with the corresponding unmodified nucleotide or nucleotides. For example, the at least one modified nucleotide has a chemically active group that can be modified after strand-specific synthesis. Another possibility is that the at least one nucleotide carries an affinity group and the selection factor binds selectively to this affinity group. Such am affinity group may for example be recognized by an antibody. A further possibility is that the at least one modified nucleotide carries a group which allows for selective recognition and binding or attachment by another agent which then enables selective removal of the such recognized strand. A further selective factor may be a different sensitivity to specific nucleases. Also the incorporation of ribonucleotides (NTPs) instead of deoxyribonucleotides (dNTPs) is suitable for enabling selection between two strands.

The at least one modified nucleotide may display a different sensitivity to a nuclease, in comparison with the corresponding unmodified nucleotide or nucleotides. Thus such nuclease degrades specifically the at least one modified nucleotide while the unmodified nucleotides are not or only insignificantly affected by such nuclease action. Alternatively, such nuclease degrades specifically the unmodified nucleotides. In both cases the strand containing the such degraded nucleotide or nucleotides has become unfit for amplification or hybridization. Thus both strands can be differentiated, and only one particular strand may be further used.

In a preferred embodiment the selection of strands of polynucleotides is carried out with the modified dNTP being dUTP (deoxyuridine triphosphate). Herein the modified dNTP is dUTP (deoxyuridine triphosphate) as a substitute for dTTP. dUTP displays the same matching characteristics as dTTP, i.e. it is able to match with an opposed adenosine, but not with opposed guanosines, cytosines, thymidines or uridines. dUTP can be specifically degraded by the nuclease Uracil-N-Glycosylase (UNG). UNG is neutral towards dTTP and the other dNTPs or analogues thereof. This enzyme degrades uracil specifically by cleaving the N-glycosylic bond resulting in an abasic site. The such degraded DNA strand is unfit for undergoing appropriate

amplification. At temperatures of over 50°C in slightly alkali conditions nicks occur in the DNA strand at the abasic positions. Additionally, Phusion DNA polymerase preferentially used for amplification (see below) is strongly suppressed by deoxyuridine-containing matrices (Hogrefe, H.H., Hansen, C.J., Scott, B.R. and Nielsen K.B. 2002. Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation. Proc Natl Acad Sci USA 99: 596-601). Consequently, only the DNA strand that does not contain dUTP can be amplified properly.

The combination of dUTP as exchange dNTP and UNG as the degrading agent of the exchange dNTP is a preferred embodiment. It is also preferred to replace only dTTP.

In another preferred embodiment the at least one modified nucleotide is thio modified and said nuclease is an exonuclease that degrades thio bonds with low efficiency. Such exonucleases sensitive to thio bonds can be lambda-exonuclease, Exo III nuclease, T4 polymerase and T7 exonuclease.

Other preferred dNTP-specific degradation systems include the use of specific biotinylated dNTPs. Herein NHS esters of biotin are reacted with free amino groups of the respective nucleotide. These biotinylated nucleotides can be recognized by polymerase enzymes as eligible for integration into the actually formed strand. The matching properties to their counter nucleotide on the other strand are comparable. The biotinylated strand might be removed using streptavidin- or avidin-coated magnetic particles. Such specifically attached magnetic particles can be removed by applying a suitable magnetic field.

In another preferred embodiment the at least one modified polynucleotide is an aminoallyl nucleotide (5-(3-aminoallyl)-nucleotide, aa-dNTP). Herein the nucleotide has an aminoallyl group attached to carbon 5 of the pyrimidine ring of uracil and cytosine. This aminoallyl group reacts with dyes from the cyanine series such as HiLyte Fluor® or Alexa Fluor® dyes that also been modified to be amine-reactive. As these dyes use to react also with standard buffer mediums a carbonate buffer is typically used. Herein it is preferred that the modified nucleotide is aminoallyl-dUTP.

In a further preferred embodiment the modified nucleotide can be selectively recognized by an antibody. The antibody binds specifically to this modified nucleotide and the strands are selected by binding or immobilizing said antibody. It is further preferred that such antibody recognizes specifically the modified nucleotide 5-

bromodeoxyuridine. It is also further preferred that such antibody or antibodies recognize specifically at least one modified nucleotide that has been modified by biotinylation.

If the strand carrying at least one modified nucleotide which carries a group which interacts which an affinity factor it is preferred that such strand is immobilized on a solid phase.

In a further preferred embodiment the at least one modified nucleotide is a ribonucleotide. Then it is preferred that said polymerase is a DNA polymerase displaying enhanced ability to incorporate modified substrates such as ribonucleotides (for example, Therminator™ II DNA Polymerase). The such incorporated at least one ribonucleotide is then specifically degraded by a specific nuclease, in this case a RNase.

If at least one classic dNTP is replaced by the corresponding modified dNTP(s) to generate a modified DNA strand the ratio of the at least one modified dNTP shall amount to at least 5% of the corresponding classic dNTP in the solution. Preferred are 20% ratio, more preferred 30%, more preferred 40%, more preferred 50%, more preferred 60%, more preferred 70%, more preferred 80% and particularly preferred more then 90%. Theoretically, only one degraded nucleotide is enough to render the degraded strand unfit for appropriate amplification. This requires that the DNA is long enough that at least one modified dNTP is integrated into the DNA strand. This is a matter of probability (when an even distribution of all four nucleotides is ideally assumed for the original polynucleotides) according to the distribution of the corresponding classic dNTPs and exchange dNTPs. With a length of at least 60 base pairs the probability is close to nil that not at least one modified dNTP has to be integrated into the second-strand cDNA. In order to ensure that also cDNAs with a length less than 60 base pairs always include at least one degradable nucleotide it is preferred that a sufficiently large percentage of the respective nucleoside triphosphate is modified.

If an amplification step is carried out in any of the embodiments of this invention an enzyme with polymerase activity is needed which is able to use the selected cDNA strand as a matrix. In theory, any such enzyme with polymerase activity is suitable for this step. However, those enzymes that are used in conventional PCR protocols are preferred. They should grant a high reliability and a high throughput under conventional PCR conditions.

To preserve information about the direction of the strand selected from a polynucleotide duplex for the further amplification., it is essential to mark differently the 5' and 3' ends of the said selected strand. This way the polarity of the original strand is preserved even after amplification. This may be achieved by for example the following ways:

1) by ligation of Y-shaped adapters to ds polynucleotide with differentiating strands, where Y-shaped adapters comprise a complementary double stranded sequence (e.g. ~13nt long) which ligates to the ds DNA, and further comprises at least one single stranded tail (e.g. 20nt long) of a different sequence. After ligation of such adapters 5'ends of each strand in ds cDNA fragment are specifically marked with the same sequence, different from that attached to the 3' ends. The adapter sequences correspond to primer sequences used later in the procedure. After selection of the strand for amplification, where the other one is either removed, or destroyed or incapable to be amplified due to sensitivity of the polymerase used to the at least one modified nucleotide, the amplification is carried out with the selected strand, wherein 5` and 3' ends bear different sequences for PCR primers (Fig. 2A-1).
2) By ligation of directional adapters (Fig.2A-2)
3) by ligation of two types of at least partly ds adapters to the ds polynucleotide; wherein one of the strands of one of adapters bears at least one biotinylated nucleotide. Through the biotin group, ds DNA with at least one biotinylated adapter is attached to streptavidine coated beads, and the non-biotinylated strand is washed away, leaving on the beads only biotinylated strand which may be with and without modified nucleotides. Then the strand for amplification is selected, while the other one is either removed, or destroyed or incapable to be amplified due to sensitivity of the polymerase used to the at least one modified nucleotide, and amplified with primers corresponding to the adapters sequences. Since the biotinylated adapter is known (whether it is attached to 5' or 3' end of the selected strand), the polarity of the selected strand is preserved after amplification (Figure 2A-3). Instead of biotin, it is possible to use any chemical group capable to bind to some surface and not interfering with amplification.

The term *PCR* refers to polymerase chain reaction. This is a general term comprising all protocols, conditions, automatized and half-automatized devices in which at least one DNA template is amplified by default by means of a enzyme with polymerase activity. In general, such molecules with polymerase activity are enzymes such as

Taq DNA polymerase, Pfu DNA polymerase, Phusion DNA polymerase, wherein Taq and Phusion DNA Polymerases are preferred.

It is particularly preferred that the at least one modified nucleotide is dUTP and said amplification is a Pfu DNA polymerase-dependent amplification.

If an amplification step is carried out in any of the embodiments of this invention a useful PCR (polymerase chain reaction) protocol shall be used by means of which the remaining intact DNA strand shall be amplified. Suitable PCR protocols of the include the known protocols for Standard PCR, Allele-Specific PCR, Alu PCR, Assembly PCR, Asymmetric PCR, Helicase Dependent Amplification HDA, Hot Start PCR, Inverse PCR, Intersequence-Specific PCR (ISSR PCR), Late-PCR, Ligation-mediated PCR, Long PCR, Methylation-Specific PCR, Miniprimer PCR, Multiplex PCR, Nested PCR, Overlap-Extension PCR, Quantitative PCR, Real Time PCR, Solid Phase PCR, Thermal Assymetric Interlaced PCR (TAIL-PCR), Touchdown PCR or combinations thereof. Preferred is the use of Standard PCR.

Herein the term "at least one intact DNA strand" means that the number of the intact DNA strands to be amplified can vary from one to many billions of different polynucleotides.

In a further optional step the resulting sequences can be aligned with already known sequence patterns of the corresponding DNA, if applicable. Such results can be plotted in a way such as shown in Fig. 6.

Optionally, the method according to the invention comprises the step of contig (contiguous) construction from the obtained sequences. Herein the sequences of different overlapping (contiguous) DNA fragments derived from the same source can be used for (re-)constructing the original respectively complete DNA sequence of this source.

Alternatively to the amplification step the at least one intact DNA strand can be hybridized to a nucleic acid structure displaying an antisense sequence of the at least one intact DNA strand. Suitable hybridization methods include *in situ* hybridization (hybridization of a DNA to a complementary sequence in isolated cells and/or tissues), Southern Blot (a method for verifying the presence of a DNA sequence in a DNA sample by separating the DNAs in the sample by size using gel electrophoresis and bringing said sample to a filter membrane for probe hybridization), Northern Blot (the equivalent of Southern Blot with RNA samples) and Southwestern Blot

(identifying the presence of a DNA sequence in a DNA sample prepared according to the Southern blot method by hybridizing the sample on a nitrocellulose gel with characterizing DNA-binding proteins) microarray hybridization and correspondent "reverse" hybridizations were intact DNA's are loaded on a membrane and hybridised with particular probes.

When an amplification is intended the key steps of the inventive method are to be read in the following order:

I) Providing polynucleotide templates, primers, buffer solution(s), 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 selection;

II) Performing first-strand synthesis on the polynucleotide templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) If said polynucleotide templates on step I) are single-stranded RNA, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof is performed;

IV) Mark differently the 5' and 3' ends of the ds cDNA to preserve the polarity of the original strand after amplification;

V) Optionally, 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;

VI) Amplifying the at least one remaining intact cDNA strand by performing PCR, whereby the other strand is removed, or destroyed or incapable to be amplified due to the at least one modified nucleotide.

The afore-mentioned method further comprises preferably the step of:

VII) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

In the particularly preferred embodiment 1 of the method according to the invention a strand specific cloning of RNAs such as a whole transcriptome of a cell or a tissue shall be carried out. Therefore, at least one single strand RNA is provided, for example by isolation from the respective cell or tissue. The four dNTPs dATP, dCTP,

dGTP and dTTP are added in sufficient amounts, i.e. amounts which allow performance of the inventive method, to the solution containing the at least one RNA strand. Then, a first cDNA strand is generated by using random primers (or oligo(dT), oligo (dU), oligo (U) primers) for annealing to said at least one RNA strand and an enzyme with Reverse Transcriptase activity. In the following at least one of the aforesaid dNTPs is replaced by at least one corresponding modified dNTP. For example, the modified dNTP may be a dNTP which can be specifically degraded by adding a degrading agent in order that the strand containing the such degraded at least one modified dNTP is not suitable anymore for amplification, or the modified dNTP may expose the such marked cDNA strand for being removed by interaction with specific agents interacting with the additional functional group or groups on the at least one modified dNTP. With the four dNTPs including the at least one modified dNTP second strand cDNA synthesis is performed by using the first cDNA strand as a matrix and a suitable enzyme with polymerase activity. In the next step polarity specific adaptors are ligated to the ends of cDNA strands, allowing a polymerase to dock to the adaptor structure for PCR amplification. After the ligation the second cDNA strand containing the at least one modified dNTP is for example degraded by degrading the at least one modified dNTP with a suitable degrading agent, or the at least one second cDNA strand containing the at least one modified dNTP is separated from the at least one first cDNA strand by special molecules or particles attaching to characteristic structures on the modified dNTP. Finally, the remaining intact first cDNA strand is amplified by PCR using a suitable enzyme with polymerase activity.

Thus the steps of this embodiment of the inventive method can be summarized in the following schematic order:

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) Amplifying the at least one remaining intact cDNA strand by performing PCR.

In embodiment 2 strand specific cloning of RNA can be achieved also by varying the steps of the inventive method. In this embodiment the first cDNA strand shall contain the at least one modified dNTP. In the replacement step this at least one modified dNTP is substituted by the corresponding classic dNTP. Thus in this embodiment the at least one first cDNA strand will be either degraded or removed. The amplification step is then performed on the remaining intact second cDNA strand. All other steps are adjusted correspondingly.

The steps of this embodiment of the inventive method can be summarized in the following schematic order:

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) Amplifying the at least one remaining intact cDNA strand by performing PCR.

In embodiment 3 strand specific selection according to the invention can also be achieved using single-stranded DNA as a template. Here at least one DNA single strand is provided. A cDNA strand is synthesized wherein at least one modified dNTP is integrated. In this embodiment only one cDNA strand is synthesized since a second cDNA strand would be identical to the original DNA single strand. Finally, the original single-stranded DNA shall be amplified. Thus all transcripts resulting from the PCR will bear the same reading direction of the sequence as the original single-stranded DNA. The other steps are adjusted accordingly.

This leads to the following schematic order:

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) Amplifying the at least one intact DNA strand by performing PCR.

In embodiment 4 the purpose of the inventive method is not an amplification but a hybridization. This hybridization can occur against any structure carrying a corresponding antisense sequence to the at least one intact first cDNA strand (see below). For a hybridization without amplification no adaptors have to be ligated. Consequently, this step is omitted in this embodiment and the amplification step is replaced by a hybridization step. Thus the order of the steps of the inventive method reads 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 used for the strand specific differentiation;
5) Performing second-strand synthesis on the first cDNA strand;
6) Specifically modifying or removing the second cDNA strand containing the at least one modified dNTP; and
7) Hybridizing the at least one remaining intact cDNA strand to a suitable antisense DNA sequence.

Embodiment 5 varies embodiment 4 in so far that herein the first cDNA strand contains the at least one modified dNTP. Thus the remaining intact second cDNA strand is hybridized. Consequently, the sequence against which the remaining intact second cDNA strand is hybridized must display an antisense sequence of the second strand. The steps of the inventive method are adjusted according to embodiments 2 and 4 and 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, wherein at least one of the four dNTPs is a modified dNTP used for the strand specific differentiation;

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) Specifically modifying or removing the first cDNA strand containing the at least one modified dNTP; and
7) Hybridizing the at least one remaining intact cDNA strand to a suitable antisense DNA sequence.

In embodiment 6 the hybridization step follows after an amplification step. In various hybridization modes it may be advantageous to hybridize not only one or a few copies of the respective intact cDNA strand, but a large number as resulting from an amplification. Thus the steps of ligating a polarity-specific adaptor and of an amplification have to be integrated again. The order of the steps reads like embodiment 1 and a subsequent hybridization step:

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) Amplifying the at least one remaining intact cDNA strand by performing PCR.
9) Hybridizing the amplified cDNA to a suitable antisense DNA sequence.

Correspondingly, in embodiment 7 the hybridization step can ensue after performing an amplification of the intact second-strand cDNA, 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) Amplifying the at least one remaining intact cDNA strand by performing PCR.
9) Hybridizing the amplified cDNA to a suitable antisense DNA sequence.

Also in the case of single-stranded DNA it may be advantageous to combine an amplification and a hybridization step. Thus in embodiment 8 the order of the steps reads as follows:

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) Amplifying the at least one intact DNA strand by performing PCR.
7) Hybridizing the amplified DNA to a suitable antisense DNA sequence.

Embodiment 9: If the sequence of the original RNA is not known it may be desirable to perform sequencing. Thus such a step is added to the step order of embodiment 1:

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.

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.

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.

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.

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

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:
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.

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.

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.

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.

Embodiment 17: In another preferred embodiment the method according to the invention is performed in the following way:

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, wherein at least one nucleotide is a modified dNTP;
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) Specifically modifying or removing the first cDNA strand containing the at least one modified dNTP; and

Embodiment 18: In another preferred embodiment the method according to the invention is performed in the following way:

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, generating a second cDNA strand, wherein at least one nucleotide is a modified dNTP;
6) Specifically modifying or removing the second cDNA strand containing the at least one modified dNTP.

Embodiment 19: In another preferred embodiment the method according to the invention is performed in the following way:

1) Performing first-strand 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 selection, so that a double-stranded polynucleotide comprising a template a matching strand containing the at least one modified nucleotide is generated;
2) Optionally, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof, if said polynucleotide template on step I) is a single-stranded RNA;
3) Optionally, performing molecular cloning operations with double-stranded DNA; and
4) Strand specific selection using different interaction of modified and unmodified nucleotides with some selection factor.

The inventive method for selection of a particular strand where the preservation of information about the direction is desired after being transformed into a double-stranded form comprises also the following steps:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on the polynucleotide templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) If said polynucleotide templates on step I) are single-stranded RNA, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof is performed;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded RNA from a double stranded polynucleotide comprises the following steps:

I) Providing RNA templates, 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 selection;

II) Performing first-strand synthesis on the RNA templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) Performing second-strand synthesis with unmodified dNTPs or corresponding analogues thereof;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded RNA from a double stranded polynucleotide can also comprise the following steps:

I) Providing RNA templates, 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 selection;

II) Performing first-strand synthesis on the RNA templates with unmodified dNTPs or corresponding analogues thereof;

III) Performing second-strand synthesis wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded DNA from a double stranded polynucleotide comprises the following steps:

I) Providing DNA templates, 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 selection;

II) Performing first-strand synthesis on the DNA templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) Optionally, performing molecular cloning operations with double-stranded DNA; and

VI) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded RNA from a double stranded DNA/RNA hybrid comprises the following steps:

I) Providing DNA/RNA templates, 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 selection;

II) Performing first-strand synthesis on the RNA templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) Performing second-strand synthesis with unmodified dNTPs or corresponding analogues thereof;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded RNA from a double stranded DNA/RNA hybrid comprises the following steps:

I) Providing DNA/RNA templates, 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 selection;

II) Performing first-strand synthesis on the RNA templates with unmodified dNTPs or corresponding analogues thereof;

III) Performing second-strand synthesis wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded DNA from a double stranded DNA/RNA hybrid comprises the following steps:

I) Providing DNA/RNA templates, 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 selection;

II) Performing first-strand synthesis on the DNA templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) Optionally, performing molecular cloning operations with double-stranded DNA; and

VI) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded DNA from a double stranded DNA/RNA hybrid comprises the following steps:

I) Providing DNA/RNA templates, 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 selection;

II) Performing first-strand synthesis on the DNA templates with unmodified dNTPs or corresponding analogues thereof;

III) Performing second-strand synthesis wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

III) Optionally, performing molecular cloning operations with double-stranded DNA; and

VI) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a single stranded DNA from a double stranded polynucleotide can also comprise the following steps:

I) Providing DNA templates, 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 selection;

II) Performing first-strand synthesis on the DNA templates with unmodified dNTPs or corresponding analogues thereof;

III) Performing second-strand synthesis wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a particular strand from a double stranded polynucleotide can further also comprise or consist of the following steps:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on the polynucleotide templates with unmodified dNTPs or corresponding analogues thereof;

III) Performing second-strand synthesis wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

The inventive method for selection of a particular strand from a double stranded polynucleotide can further also comprise or consist of the following steps:

I′) Providing polynucleotide templates, 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 selection;

IIa´) Performing first-strand synthesis on the polynucleotide templates with unmodified dNTPs or corresponding analogues thereof;

or

IIb´) Performing first-strand synthesis on the polynucleotide templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

IIIa`) If step IIa´ was performed, second-strand synthesis is performed wherein at least one of the four dNTPs is replaced by a corresponding modified dNTP used for strand-specific selection, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the template contains the at least one modified nucleotide;

IIIb´) If step IIb´ was performed and said polynucleotide templates on step I) are single-stranded RNA, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof is performed

and

IV´) Optionally, performing molecular cloning operations with double-stranded DNA; and

V´) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

In a further embodiment the inventive method comprises selecting in ds (double stranded) polynucleotides the strand with at least one modified nucleotide to distinguish the template from the matching strand, wherein the selected strand with the at least one modified nucleotide in ds (double stranded) polynucleotides can be removed or destroyed and the removed strand with the modified nucleotide can be

used separately, whereby if the strand with the modified nucleotide is not removed and destroyed, it can remain in the same solution with the matching strand, wherein the modified nucleotide in the strand prevents the amplification of the strand comprising the following steps:

I) Providing polynucleotide templates, 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 selection;

II) Performing first-strand synthesis on the polynucleotide templates, so that double-stranded polynucleotides each consisting of a template and a matching strand are generated, wherein the matching strand contains the at least one modified nucleotide;

III) If said polynucleotide templates on step I) are single-stranded RNA, second-strand synthesis with unmodified dNTPs or corresponding analogues thereof is performed;,;

IV) Optionally, performing molecular cloning operations with double-stranded DNA; and

V) Specific selection of at least one of the strands of the said polynucleotide duplex using different interaction of modified and unmodified nucleotides with at least one selection factor.

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.

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 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.

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.

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.

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.

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.

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.

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).

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.

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

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.

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.

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 selection;
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.

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 selection;
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.

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 selection;
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.

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. Further kits are also preferred:

KIT A:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific selection;
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.

KIT B:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific selection;
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.

KIT C:

A) All four dNTPs dATP, dCTP, dGTP and dTTP or corresponding analogues thereof, additionally at least one modified dNTP used for the strand specific selection;
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.
G) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

KIT D:

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;

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;

E) One or more DNA polymerases with reverse transcriptase activity;

Optionally comprising:

F) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

KIT E:

A kit for ds cDNA synthesis from RNA preserving information about direction of original RNA strand comprising:

A) All four dNTPs dATP, dCTP, dGTP and dTTP/dUTP and at least one modified nucleotide corresponding to at least one from dATP,ATP, dCTP, CTP, dGTP, GTP, dTTP, TTP, dUTP and UTP;

B) Primers;

C) One or more DNA polymerases with reverse transcriptase activity;

D) One or more DNA polymerases; and

E) One or more selection factors suitable to specifically interact with said at least one modified nucleotide and/or one or more selective media for removing the strand containing the modified nucleotide;

Or additionally comprising

F) Components for specific removal of at least one modified or at least one unmodified nucleotide by gel-filtration or purification on silica-column.

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. Actinomycin D is included in the kit to make the reaction more reproducible.

Thus in a preferred embodiment the kit also includes an inhibitor of DNA-dependent synthesis characterized in that said inhibitor does inhibit DNA-dependent, but not RNA-dependent, DNA synthesis.

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.

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.

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. In particular, there is no need for the use of promoters or primer constructs containing promoter regions. The essence of the invention is the preservation of information about the direction of the original single-stranded molecules after being transformed into a double-stranded form or of particular strands in the original double stranded molecule and selection of the original strand according to the method of the invention. . In a preferred embodiment the single-stranded molecule is a RNA and the double-stranded form a DNA.

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.

A further advantage of the method according to the invention is that the RNA is transcribed in cDNA and all the following steps are carried out with the cDNA which is much more resistant to degradation compared to RNA. In particular, strand specific differentiation is achieved by modifying a nucleotide in the cDNA and not in the RNA.

Figures

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.

(A3) adaptors are ligated to double-stranded DNA with distinguishable strands, with one of adaptors bearing biotin in one of the strands. Streptavidine coated beads are used to fish out fragments, containing biotin. Uridine-containing strand is destroyed, remaining strand is amplified.

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.

Fig. 3B: Preparation of single-stranded DNA according to the inventive method using biotinylated 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.

Examples

Example 1: Transcriptome of Saccharomyces cerevisiae

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:
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:
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):
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:

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):

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:

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:

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

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:

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

Data analysis:

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:

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:

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