AMPLIFICATION OF NUCLEIC ACID FRAGMENTS USING NICKING AGENTS

BACKGROUND OF THE INVENTION

Technical Field

This invention relates to the field of molecular biology, more particularly to methods and compositions involving nucleic acids, and still more particularly to methods and compositions for amplifying nucleic acid fragments using nicking agents.

Description of the Related Art

The chromosomal mapping and nucleic acid sequencing of each of the 80,000 to 100,000 human genes, achieved through the Human Genome Project, provides an opportunity for a comprehensive approach to the identification of nucleotide loci responsible for genetic disease. Many of the 150-200 common genetic diseases and about 600-800 of the rarer genetic diseases are associated with one or more defective genes. Of these, more than 200 human diseases are known to be caused by a defect in a single gene, often resulting in a change of a single amino acid residue. (Olsen, "Biotechnology: An Industry Comes of Age" (National Academic Press, 1986)).

Mutations occurring in somatic cells may induce disease if the mutations affect genes involved in cellular division control, resulting in, for example, tumor formation. In the germline, loss-of-function mutations in many genes can give rise to a detectable phenotype in humans. The number of cell generations in the germline, from one gamete to a gamete in an offspring, may be around 20-fold greater in the male germline than in the female. In the female, an egg is formed after a second meiotic division and lasts for 40 years. Therefore the incidence of different types of germline mutations and chromosomal aberrations depends on the parent of origin.

A majority of mutations, germline or somatic, are of little consequence to the organism since most of the genome appears to lack coding function (about 94%). Even within exon regions, there is some tolerance to mutations both due to the degeneracy of the genetic code and because the amino acid substitutions may have only a slight influence on a protein's function. (*See, e.g.*, Strong et al., *New England Journal of Medicine* 325:1597 (1991)). With the development of increasingly efficient methods to detect mutations in large DNA segments, the need to predict the functional consequences (*e.g.*, the clinical phenotype) of a mutation becomes more important.

Molecular genetic techniques have not been employed to a significant extent in the diagnosis of chromosomal aberrations in genetic and malignant disease; cytogenetics remains the preferred technique to investigate these important genetic mechanisms. In an individual with one mutated copy of a tumor suppressor gene, the remaining normal allele may be replaced by a second copy of the mutant allele in one cell per 10^3-10^4. Mechanisms causing this replacement include chromosomal nondisjunction, mitotic recombination, and gene conversion. In contrast, independent mutations destroying the function of the remaining gene copy are estimated to occur in one cell out of 10^6.

Sensitive mutation detection techniques offer extraordinary possibilities for mutation screening. For example, analyses may be performed even before the implantation of a fertilized egg. (Holding et al., *Lancet 3*:532 (1989)). Increasingly efficient genetic tests may also permit screening for oncogenic mutations in cells exfoliated from the respiratory tract or the bladder in connection with health checkups. (Sidransky et al., *Science 252*:706, 1991). Alternatively, when an unknown gene causes a genetic disease, methods to monitor DNA sequence variants are useful to study the inheritance of disease through genetic linkage analysis. Notwithstanding these unique applications for the detection of mutations in individual genes, the existing methodology for achieving such applications continues to pose technological and economic challenges. While several different approaches have been pursued, none are sufficiently efficient and cost effective for wide scale application.

Conventional methods for detecting mutations at defined nucleotide loci involve time-consuming linkage analyses in families using limited sets of genetic markers that are difficult to "readout." Such methods include, *e.g.*, DNA marker haplotyping (that identifies chromosomes with an affected gene) as well as methods for detecting major rearrangements such as large deletions, duplications, translocations and single base pair mutations. These methods include scanning, screening and fluorescence resonance energy transfer (FRET)-based techniques. (*See,* Cotton, "Mutation Detection" (Oxford University Press, 1997)).

Highly sensitive assays that detect low abundance mutations rely on PCR to amplify the target sequence. However, these assays suffer from high costs associated with PCR reagents. In addition, current methods for detecting and/or characterizing amplified nucleic acid fragments containing mutations are not well developed to allow high degree of multiplex identification of genetic variation.

Due, in part, to the shortcomings in the existing methodology for detecting genetic mutations, there exists an unmet need for rapid and sensitive methods

for detecting mutations at defined nucleotide loci within target nucleic acids. The present invention fulfills this and other related needs by providing methods for the detection of mutations at defined nucleotide loci in target nucleic acids that, *inter alia*, display increased speed and convenience, as well as decreased costs. As disclosed in detail herein below, methods according to the present invention are based on amplifying nucleic acid fragments containing genetic variations at defined locations using nicking agents. In addition, the present invention also provides methods and compositions for detecting the presence, or the absence, of a particular nucleic acid in a biological sample; for preparing single-stranded nucleic acid probes; and for detecting the presence, or the absence, of a junction between two exons in a mRNA molecule or a cDNA molecule.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides a method that includes: (a) Contacting a template nucleic acid molecule with a nicking agent (NA), where the template nucleic acid molecule ("the template") includes a nicking agent recognition sequence (NARS) that is recognizable by the NA. In one embodiment, the template is partially double-stranded, where in one embodiment the NARS is within a single-stranded portion of the template, while in another embodiment the NARS is within a double-stranded portion of the template. In another embodiment, the template is fully double-stranded. (b) If the template does not include a nicking site (NS) nickable by the NA, then the method provides for extending one strand of the template so as to form an extension product of the template, where this extension product includes a NS nickable by the NA. (c) Nicking the template or the extension product thereof at the NS. This will provide 3' and 5' termini at the NS. As described next, multiple copies of the nucleic acid fragment having the 5' terminus at the nicking site will be prepared by the following steps. (d) Extending the nicked product of step (c) from the 3' terminus at the NS. (e) Repeating steps (c) and (d) to thereby amplify a single-stranded nucleic acid fragment. In various embodiments of the invention, this single-stranded nucleic acid fragment has no more than 50, or 45, or 40, or 35, or 30, or 25, or 24, or 23, or 22, or 21, or 20, or 19, or 18, or 17, or 16, or 15, or 14, or 13, or 12, or 11, or 10, or 9, or 8, or 7, or 6, or 5, or 4, or 3, or 2, or 1 nucleotides. In general, the formation of single-stranded nucleic acid fragments having no more than 50, 45, etc. nucleotides is advantageous, relative to forming fragments having more than 50, 45, etc. nucleotides, for at least the following reasons: the rate and efficiency of fragment formation is greater, smaller fragments are more readily characterized to a higher level of distinction

by some characterization methods, *e.g.*, mass spectrometry and liquid chromatography, and the extension step may be performed with a DNA polymerase that does not necessarily have strand displacement activity.

The following criteria are exemplary of additional criteria that may be used to further describe this method of the present invention. Any one, two, three, four, five, six, seven, etc. criteria as disclosed herein, including the following criteria, may be combined in describing this method of the present invention. When two criteria (A and B) are inconsistent with one another, they may be combined in the alternative, that is, the method may be described as including criteria A or criteria B. These criteria include: the NA is a nicking endonuclease (NE); the NA is a restriction endonuclease (RE); step (c) is performed in the presence of a DNA polymerase selected from the following listed DNA polymerases, where the list of potential DNA polymerases may be reduced so as to include any two, three, four, five, six, seven, etc. of the following DNA polymerases: exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm^{TM} DNA polymerase, and T4 DNA polymerase, *e.g.*, step (c) may be performed in the presence of a DNA polymerase selected from the group consisting of exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, or exo$^-$ Vent polymerase; two or more, preferably three or four, and still more preferably all four of steps (b), (c), (d) and (e) are performed under an identical isothermal condition; the method includes the step of characterizing the amplified single-stranded nucleic acid fragment of step (e); the method includes the step of characterizing the amplified single-stranded nucleic acid fragment of step (e) by one or more methods selected from luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, electron ionization, gel electrophoresis, and capillary electrophoresis, for example, the method includes the step of characterizing the amplified single-stranded nucleic acid fragment of step (e) by mass spectrometry, *i.e.*, by mass spectrometry alone or by methods including mass spectrometry; the template nucleic acid has a genetic variation located 3' to the NS and on the same strand of the template that includes the NS, such that the genetic variation is incorporated into the amplified single-stranded nucleic acid fragment; the template nucleic acid may be formed by a method that includes the following steps (i) and (ii), and may further include step (iii): (i) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and a target nucleic acid, where the target nucleic acid includes a genetic variation under investigation, wherein if (a) the

target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, and optionally comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS), however, if (b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a sequence of a sense sequence of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially identical to the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, and optionally comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS). (ii) Extending the first and the second ODNPs to produce a fragment having both the NERS and the RERS. The method optionally includes (iii) Cleaving the extension product of step (ii) with a restriction endonuclease that recognizes the RERS.; the template nucleic acid may be formed by a method that includes the following steps (i) and (ii), as follows: (i) Forming a mixture that includes a first ODNP, a second ODNP, and a target nucleic acid comprising a genetic variation, wherein if (a) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, however, if (b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located

3' to the genetic variation. (ii) Extending the first and the second ODNPs in the presence of deoxyribonucleoside triphosphates and at least one modified deoxyribonucleoside triphosphate to produce a fragment having both the first and the second RERSs; the template nucleic acid is formed by a method that includes (i) and (ii), as follows: (i) Forming a mixture that includes a first ODNP, a second ODNP and a target nucleic acid comprising a genetic variation, wherein if (a) the target nucleic acid is a double-stranded nucleic acid molecule having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, however, if (b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation; where in either event the first and the second ODNPs each further comprise a sequence of a sense strand of a NERS. (ii) Extending the first and the second ODNPs to produce a fragment having two NERSs.; the template nucleic acid molecule further includes a 5' overhang in the strand that contains the NS wherein the overhang includes a nucleic acid sequence that is at least substantially complementary to a target nucleic acid; the template nucleic acid molecule further includes a 3' overhang in the strand that does not contain the NS, wherein the overhang includes a nucleic acid sequence that is at least substantially complementary to a target nucleic acid; in the event that the template contains an overhang, the present invention optionally provides for performing the following additional steps (i), (ii), and (iii): (i) Admixing the template nucleic acid with nucleic acid molecules in a biological sample, where the biological sample may contain the target nucleic acid, under conditions where the target nucleic acid, if present in the biological sample, would hybridize to the overhang of the template nucleic acid. (ii) Removing unhybridized template nucleic acid from the admixture of step (i) before performing step (a). (iii) Combining the hybridized template nucleic acid with a NA; the double-stranded template nucleic acid molecule includes a type IIs restriction endonuclease recognition sequence (TRERS); the double-stranded template nucleic acid molecule is provided by linking a nucleic acid adaptor to a double-stranded target nucleic acid in a manner such that (i) the amplified single-stranded nucleic acid

molecule formed according to the present method includes a portion of the target nucleic acid, (ii) the nucleic acid adaptor includes a type IIs restriction endonuclease recognition sequence (TRERS), (iii) the nucleic acid adaptor includes a NARS in the strand of the adapter that does not link to the strand of the template that contains the NS, (iv) the cleavage site of the type IIs restriction endonuclease that recognizes the TRERS is located (a) within the double-stranded target nucleic acid fragment and (b) 5' to the position corresponding to the NS; the template nucleic acid includes a junction between an upstream exon (Exon A) and a downstream exon (Exon B) located 3' to the NS such that the nucleotides adjacent to the junction at both sides of the junction are incorporated into the amplified single-stranded nucleic acid fragment; the template nucleic acid is formed by a method that includes (i) and (ii) as follows: (i) Admixing a first oligonucleotide primer (ODNP), a second ODNP, and a cDNA, wherein the first ODNP comprises a sequence at least substantially complementary to a portion of the antisense strand of an Exon A near the 5' terminus of Exon A in the antisense strand, the second ODNP comprises a sequence at least substantially complementary to a portion of the sense strand of an Exon B near the 5' terminus of Exon B in the sense strand, and at least one of the first ODNP and the second ODNP further includes a sequence of a sense strand of a nicking agent recognition sequence (NARS). (ii) Extending the first ODNP and the second ODNP to provide the template nucleic acid that comprises the first ODNP and the second ODNP; at some time prior to step (e), the template nucleic acid may be immobilized on a solid support, where in one embodiment the nucleic acid strand that does not contain the nicking site is attached to a solid support, optionally at the 3' end of this strand, or optionally at the 5' end of this strand, while in another embodiment the 5' end of the nucleic acid fragment that includes the 3' termini formed upon nicking at the NS is attached to a solid support, while in yet another embodiment the 3' end of the nucleic acid fragment that includes the 5' termini formed upon nicking at the NS is attached to a solid support; for example, when a cDNA is used to prepare the template nucleic acid, then the cDNA molecule may be attached to a solid support; as another example, when a target nucleic acid is used to prepare the template nucleic acid, then the target nucleic acid may be attached to a solid support. In a related embodiment, the present invention provides a kit that is useful for detecting the presence, or the absence, of a junction between an upstream exon A (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, where this kit includes at least one pair of ODNPs that perform in the method for the detection as described herein. In addition, these kits may further include, or may be further characterized as including, one or more of the following components/conditions: a nicking agent (NA)

that recognizes the NARS; a restriction endonuclease (RE); a buffer for the RE; a nicking endonuclease (NE); a buffer for the NE; N.BstNB I; the first ODNP comprises a NERS, and the second ODNP comprises a RERS; the first ODNP comprises a RERS, and the second ODNP comprises a NERS; the RERS is recognizable by a type IIs restriction endonuclease; a type IIs restriction endonuclease; a buffer for the type-IIs restriction endonuclease; a type IIs restriction endonucelase known as Bpm I; a NE that recognizes the NERS; a DNA polymerase; a buffer for the DNA polymerase; one or more of exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, and exo$^-$ Vent polymerase; instructions for using the kit; a liquid chromatographic column; Buffer A that comprises water and an ammonium salt of a secondary or tertiary amine complexed with an organic or inorganic acid; Buffer B that comprises Buffer A and organic solvent; an ammonium salt wherein the amine component is selected from the group consisting of triethylamine, diallylamine, diisopropylamine, N,N-dimethyl-N-cyclohexylamine, N,N-dimethyl-N-isopropylamine, and N,N-dimethyl-N-butylamine; a secondary or tertiary amine which is complexed with an organic acid, and the organic acid is selected from acetic acid, propionic acid, and halogenated versions thereof; methanol; acetonitrile; a reverse phase chromatography column; trehalose; a deoxyribonucleoside triphosphate; a modified deoxynucleoside triphosphate; oligonucleotide standards; and/or an access code for a software used in designing or ordering the ODNP pair.

In other aspects, the present invention provides methods for identifying a genetic variation at a defined position in a target nucleic acid, where the genetic variation may be, *e.g.,* a single nucleotide polymorphism (SNP). For instance, in one embodiment, the method includes: (a) Providing a template nucleic acid ("template") that includes some or all of the target nucleic acid, and in particular includes a nucleotide sequence that encompasses the genetic variation at the defined position of the target nucleic acid. The template further includes a nicking site (NS) of a nicking agent (NA), where the NS is located 5' to the genetic variation. The template may be partially or fully double-stranded. (b) Amplifying, in the presence of a DNA polymerase and a NA that nicks at the NS, a single-stranded nucleic acid fragment that has a nucleotide sequence that includes the genetic variation. This single-stranded nucleic acid fragment has a 5' terminus that was formed by the nicking that occurred at the NS. (c) Characterizing the single-stranded nucleic acid fragment to thereby identify the genetic variation. In another embodiment, a method of the present invention for identifying a genetic variation at a defined position in a target nucleic acid, where the genetic variation may be, *e.g.,* a SNP, includes: (a) Forming a mixture that includes a

first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation and optionally comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS), however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation and optionally comprises a RERS. (b) Extending the first and the second ODNPs to produce an extension product comprising the first ODNP and the second ODNP. (c) Optionally digesting the extension product of step (b) with a restriction endonuclease that recognizes the RERS to produce a digestion product. (d) Amplifying a single-stranded nucleic acid fragment using one strand of the extension product of step (b) or the digestion product of step (c) as a template in the presence of a nicking endonuclease (NE) that recognizes the NERS. (e) Characterizing the single-stranded fragment of step (d) to thereby identify the genetic variation in the target nucleic acid. As a further example, the present invention provides a method for identifying a genetic variation at a defined location in a target nucleic acid, where the genetic variation may be, *e.g.,* a SPN, where the method includes: (a) Forming a mixture that includes a first ODNP, a second ODNP, and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then

the first ODNP comprises a nucleotide sequence of one strand of a first RERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the complement of the genetic variation, and the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation. (b) Extending the first and the second ODNPs in the presence of deoxyribonucleoside triphosphates and at least one modified deoxyribonucleoside triphosphate to produce an extension product comprising both the first and the second RERSs. (c) Amplifying a single-stranded nucleic acid fragment using one strand of the extension product of step (b) as a template in the presence of restriction endonucleases (REs) that recognize the first RERS and the second RERS, wherein the single-stranded nucleic acid fragment is no more than 35 nucleotides in length. (d) Characterizing the single-stranded fragment of step (c) to thereby identify the genetic variation. In another aspect of the present invention directed to a method for identifying a genetic variation at a defined location in a target nucleic acid, where the genetic variation may be, *e.g.,* a SNP, the method includes: (a) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, where the first and the second ODNPs each further comprise a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS). (b) Extending the first and the second ODNPs to produce an extension product comprising two NERSs. (c) Amplifying a single-stranded nucleic acid fragment using one strand of the extension product of step (b) as a template in the presence of one or more nicking endonucleases (NEs) that recognizes the NERS(s). (d) characterizing the single-stranded fragment of step (c) to thereby identify the genetic variation. In a related aspect, the present invention provides a method for identifying a

genetic variation at a defined location in a single-stranded target nucleic acid, where the method includes: (a) Forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the defined location, the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the complement of the target nucleic acid at a location 3' to the at least substantially complementary nucleotide of the nucleotide at the defined location, the first and the second ODNPs further comprise a first constant recognition sequence (CRS) of a first strand and a second CRS of a second strand of an interrupted restriction endonuclease recognition sequence (IRERS), respectively, but not a complete IRERS, the complete IRERS being a double-stranded nucleic acid having the first and the second strands and comprising the first and the second CRS linked by a variable recognition sequence (VRS), and either the first or the second ODNP further comprises a NERS located at 5' to the first CRS or the second CRS. (b) Extending the first and the second ODNPs to produce a fragment having the NERS and the complete IRERS wherein the genetic variation is within the VRS. (c) Amplifying a single-stranded nucleic acid fragment using one strand of the extension product of step (b) as template in the presence of a nicking endonuclease (NE) that recognizes the NERS. (d) Characterizing the single-stranded fragment to thereby identify the genetic variation. In further describing this method that entails a CRS, VRS, and IRERS, one or more of the following criteria may be utilized, where any two or more of the following criteria may be combined, and these criteria are exemplary only of criteria set forth herein: the single-stranded target nucleic acid is one strand of a denatured double-stranded nucleic acid; the double-stranded nucleic acid is genomic nucleic acid; the double-stranded DNA is cDNA; the nucleotide sequence of the first ODNP at least substantially complementary to the target nucleic acid is at least 12 nucleotides in length; the nucleotide sequence of the second ODNP at least substantially complementary to the complement of the target nucleic acid is at least 12 nucleotides in length; the first ODNP is 15-85 nucleotides in length; the second ODNP is 15-85 nucleotides in length; the first ODNP further comprises one or more nucleotides at least substantially complementary to the target nucleic acid at the 3' terminus of the first CRS; the second ODNP further comprises one or more nucleotides at least substantially complementary to the target nucleic acid at the 3' terminus of the second CRS; all of step (a) through (d) are performed in a single vessel; the IRERS is recognizable by a restriction endonuclease selected from the group consisting of Bsl I, Mwo I and Xcm I; the NERS is

recognizable by N.BstNB I; step (d) is performed at least partially by the use of a technique selected from the group consisting of mass spectrometry, liquid chromatography, fluorescence polarization, electron ionization, gel electrophoresis and capillary electrophoresis;

Each of these methods of identifying a genetic variation at a defined position, including methods where the genetic variation is a SNP, may be further described according to the present invention by one or more of the following criteria, where any two or more of the following criteria may be combined in describing any of the methods, and where these criteria are exemplary only in that other criteria as set forth herein may also be used to further describe a method of identifying a genetic variation at a defined position in a target nucleic acid according to the present invention: the genetic variation is a single nucleotide polymorphism (SNP); the genetic variation is associated with a disease; the genetic variation is associated with a human genetic disease; the genetic variation is associated with drug resistance of a pathogenic microorganism; the target nucleic acid is genomic nucleic acid; the target nucleic acid is cDNA; the target nucleic acid is derived from the genome of a pathogenic virus; the target nucleic acid is derived from the genome of a pathogenic bacterium; the target nucleic acid is derived from the episome of a pathogenic bacterium; the NA is a nicking endonuclease (NE); the NA is N.BstNB I; amplification is performed under isothermal conditions; amplification is performed within the temperature range of 50°C-70°C; amplification is performed at 60°C; amplification is performed at 55 ± 5°C; amplification is performed at 60 ± 5°C; amplification is performed at 65 ± 5°C; amplification is performed at 70 ± 5°C; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 20°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 15°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 10°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 5°C of the lowest temperature; amplifying is performed in the presence of a DNA polymerase, where the DNA polymerase is any of exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm^{TM} DNA polymerase, and T4 DNA polymerase, and where any two or more of these DNA polymerases may form

a group that sets forth the DNA polymerases that may be used for the amplification, *e.g.,* the DNA polymerase is selected from the group consisting of exo⁻ Bst polymerase, exo⁻ Bca polymerase, 9°Nm™ DNA polymerase, and exo⁻ Vent polymerase; the single-stranded nucleic acid fragment is characterized by one or more techniques selected from the group consisting of mass spectrometry, liquid chromatography, fluorescence polarization, and electrophoresis, for example, the single-stranded nucleic acid fragment is characterized by one or more techniques that include liquid chromatography, or as another example the single-stranded nucleic acid fragment is characterized by one or more techniques that include mass spectrometry, or as yet another example, the single-stranded nucleic acid fragment is characterized by one or more techniques that include both of liquid chromatography and mass spectrometry; when the method includes using primers, *i.e.,* ODNPs, in various embodiments of the invention each ODNP used in the method is independently at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides in length; when the method of the invention includes using primers, *i.e.,* ODNPs, then in various embodiments of the invention each ODNP used in the method is independently no more than 50, or no more than 45, or no more than 40, or no more than 35, or no more than 30, or no more than 29, or no more than 28, or no more than 27, or no more than 26, or no more than 25, or no more than 24, or no more than 23, or no more than 22, or no more than 21, or no more than 20, or no more than 19, or no more than 18 nucleotides in length; when the method includes using primers that include a nucleotide sequence at least substantially complementary to a nucleotide sequence in a target nucleic acid, then in various embodiments of the invention this nucleotide sequence in the primer is, independently in each primer, at least 4, at least 5, at least 6, at least 7, at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides in length; when the method of the invention includes using primers that include a nucleotide sequence at least substantially complementary to a nucleotide sequence in a target nucleic acid, then in various embodiments of the invention this nucleotide sequence in the primer is, independently in each primer, no more than 20, or no more than 19, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10, or no more than 9, or no more than 8, or no more than 7, or no more than 6, or no more than 5 nucleotides in length; when the method of the present invention entails the use of a nucleic acid molecule containing a RERS, then in one embodiment the RERS is

recognizable by a Type II restriction endonuclease; when the method of the present invention entails the use of a nucleic acid molecule containing a RERS, then in one embodiment the RERS is recognizable by a Type IIs restriction endonuclease; when the method of the present invention entails the use of a nucleic acid molecule containing a RERS, then in one embodiment the RERS is recognizable by Bpm I; when the method of the present invention entails the use of a nucleic acid molecule containing one or more RERSs, or the method entails the use of two or more molecules each having a RERS, then in various embodiments the RERSs are recognizable by a RE selected from Ava I, Bsl I, BsmA I, BsoB I, Bsr I, BstN I, BstO I, Fnu4H I, Hinc II, Hind II, and Nci I, where a group may be formed from any two or more of these Res; when the method of the present invention entails the use of a nucleic acid molecule containing a RERS, then in one embodiment the RERS is recognizable by an interrupted restriction endonuclease; when a method includes two RERSs, then in one embodiment the nucleotide sequence of the first RERS is not the same as the second RERS, while in another embodiment the nucleotide sequence of the first RERS is the same as the second RERS; when the method of the present invention entails the use of a nucleic acid molecule containing a NERS, then in one embodiment the NERS has a sense strand having the nucleotide sequence 5'GAGTC3'; when the method of the present invention entails extending a nucleic acid molecule, then in one embodiment the extension is preformed by a polymerase chain reaction; the method utilizes a target nucleic acid that is double-stranded; the method utilizes a target nucleic acid that is single-stranded; when the method includes the optional step of digesting an extension product with a restriction endonuclease, then in one embodiment of the invention the method includes the step of digesting an extension product with a restriction endonuclease, while in another embodiment the method omits the step of digesting an extension product with a restriction endonuclease; at some time prior to characterizing the single-stranded nucleic acid fragment to obtain information about the genetic variation, the template nucleic acid may be immobilized on a solid support, where in one embodiment the nucleic acid strand that does not contain the nicking site is attached to a solid support, optionally at the 3' end of this strand, or optionally at the 5' end of this strand, while in another embodiment the 5' end of the nucleic acid fragment that includes the 3' termini formed upon nicking at the NS is attached to a solid support, while in yet another embodiment the 3' end of the nucleic acid fragment that includes the 5' termini formed upon nicking at the NS is attached to a solid support; in some instances, a target nucleic acid and first and second ODNPs are used to form the template, where the template is subjected to an amplification process to provide a

single-stranded nucleic acid fragment, and in these instances, the target nucleic acid, or the first ODNP, or the second ODNP, may be immobilized to a solid support, where the termini of the first or second ODNP that is being extended in order to form the template is preferably not the termini of the ODNP that is linked to the solid support, and where the target nucleic acid may be immobilized either non-specifically, *i.e.,* at neither end, or, more preferably, at the 3' end, or at the 5' end of the target nucleic acid.

As mentioned previously, the present invention provides methods for identifying a genetic variation at a defined position in a target nucleic acid, where that genetic variation may be a single nucleotide polymorphism (SNP). For example, in one aspect, the invention provides a method for identifying a SNP at a defined position in a target nucleic acid, where the method includes: (a) Providing a partially or fully double-stranded template nucleic acid that comprises a portion of the target nucleic acid that includes the SNP at the defined position, and further comprises a nicking site (NS) located at 5' to the genetic variation, *i.e.,* 5' to the SNP. (b) Amplifying, in the presence of a DNA polymerase and a nicking endonuclease (NE) that nicks at the NS and under an isothermal condition, a single-stranded nucleic acid fragment that comprises the genetic variation, *i.e.,* the SNP, contains no more than 17 nucleotides, and whose 5' terminus was at the NS. (c) Characterizing the single-stranded nucleic acid fragment using at least liquid chromatography and/or mass spectrometry, to thereby identify the SNP. As another example, in another aspect, the invention provides a method for identifying a SNP at a defined position in a target nucleic acid, where the method includes: (a) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the SNP and further comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS), however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located

3' to the SNP and further comprises a RERS. (b) extending the first and the second ODNPs to produce an extension product comprising the first ODNP and the second ODNP. (c) Digesting the extension product of step (b) with a restriction endonuclease (RE) that recognizes the RERS to produce a digestion product. (d) Amplifying a single-stranded nucleic acid fragment that contains no more than 17 nucleotides under an isothermal condition using one strand of the extension product of step (b) or the digestion product of step (c) as a template in the presence of a nicking endonuclease (NE) that recognizes the NERS. (e) Characterizing the single-stranded fragment of step (d) using at least partially liquid chromatography and/or mass spectrometry to thereby identify the SNP in the target nucleic acid. The invention also provides a method for identifying a single nucleotide polymorphism (SNP) at a defined location in a target nucleic acid, where the method includes: (a) Forming a mixture that includes a first ODNP, a second ODNP, and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the SNP, and the second ODNP comprises a nucleotide sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the SNP, however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence of one strand of a first RERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the complement of the SNP, and the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the SNP. (b) Extending the first and the second ODNPs in the presence of deoxyribonucleoside triphosphates and at least one modified deoxyribonucleoside triphosphate to produce an extension product comprising both the first and the second RERSs. (c) Amplifying a single-stranded nucleic acid fragment that contains no more than 17 nucleotides under an isothermal condition using one strand of the extension product of step (b) as a template in the presence of restriction endonucleases (REs) that recognize the first RERS and the second RERS. (d) Characterizing the single-stranded fragment of step (c) using at least partially liquid chromatography and/or mass spectrometry to thereby identify the SNP. The invention also provides a method for identifying a SNP at a defined location in a target nucleic

acid, where the method includes (a) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP and the target nucleic acid, wherein if (i) the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the SNP, however, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the SNP, where the first and the second ODNPs each further comprise a nucleotide sequence of a sense strand of the same nicking endonuclease recognition sequence (NERS). (b) Extending the first and the second ODNPs to produce an extension product comprising two NERSs. (c) Amplifying a single-stranded nucleic acid fragment that contains no more than 17 nucleotides under an isothermal condition using one strand of the extension product of step (b) as a template in the presence of a nicking endonuclease (NE) that recognizes the NERS. (d) Characterizing the single-stranded fragment of step (c) using at least partially liquid chromatography and/or mass spectrometry to thereby identify the SNP. In yet another aspect, the present invention provides a method for identifying a single nucleotide polymorphism (SNP) at a defined location in a single-stranded target nucleic acid, where the method includes: (a) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the defined location, the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the complement of the target nucleic acid at a location 3' to the at least substantially complementary nucleotide of the nucleotide at the defined location, the first and the second ODNPs further comprise a first constant recognition sequence (CRS) of a first strand and a second CRS of a second strand of an interrupted restriction endonuclease recognition sequence (IRERS), respectively, but not a complete IRERS, the complete IRERS being a double-stranded nucleic acid having the first and the second strands and comprising the first and the second CRS linked by a variable recognition sequence (VRS), and either the first or the second ODNP further comprises

a NERS located at 5' to the first CRS or the second CRS. (b) Extending the first and the second ODNPs to produce a fragment having the NERS and the complete IRERS wherein the SNP is within the VRS. (c) Amplifying a single-stranded nucleic acid fragment under an isothermal condition using one strand of the extension product of step (b) as template in the presence of a nicking endonuclease (NE) that recognizes the NERS, wherein the single-stranded fragment contains no more than 17 nucleotides. (d) Characterizing the single-stranded fragment of step (c) using at least partially liquid chromatography and/or mass spectrometry to thereby identify the SNP.

As a convenient means to perform a method for identifying a genetic variation at a defined location in a target nucleic acid, *e.g.*, a SNP, the present invention provides various kits. In one aspect, a kit of the present invention that is particularly useful when the target nucleic acid is double-stranded, *i.e.*, has first and second strands, includes: (i) A first oligonucleotide primer (ODNP) that includes a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and also includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first strand of the target nucleic acid located 3' to the complement of the genetic variation. (ii) A second ODNP that includes a nucleotide sequence that is at least substantially complementary to the nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation. The second ODNP may optionally also contain the nucleotide sequence of one strand of a restriction endonuclease recognition sequence (RERS). In another aspect, a kit of the present invention that is particularly useful when the target nucleic acid is single-stranded, includes: (i) A first ODNP that includes a nucleotide sequence identical to that of a sense strand of a NERS and further includes a nucleotide sequence that is at least substantially identical to a nucleotide sequence present in the target nucleic acid located 5' to the genetic variation. (ii) A second ODNP that includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid at a location 3' to the genetic variation. The second ODNP may optionally include the nucleotide sequence present in one strand of a RERS. In yet another aspect, the present invention provides a kit for identifying a genetic variation at a defined location in a target nucleic acid, which is particularly useful when the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, where the kit includes: (i) A first ODNP that includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation. (ii) A second ODNP that includes a nucleotide sequence that is at least substantially complementary

to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation. In this kit, the first and the second ODNPs each further include a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS). In still another aspect, the present invention provides a kit for identifying a genetic variation at a defined location in a target nucleic acid, which is particularly useful when the target nucleic acid is a single-stranded, where the kit includes: (i) A first ODNP that includes a nucleotide sequence that is at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation. (ii) A second ODNP that includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation. In this kit, the first and the second ODNPs each further include a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS). The present invention also provides a kit for identifying a genetic variation at a defined location in a target nucleic acid having first and second strands, comprising: A first oligonucleotide primer (ODNP) and a second oligonucleotide primer (ODNP) wherein (a) the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, (b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, (c) the first and the second ODNPs each further comprise a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS); and (d) the distance between the a nicking site (NS) produced by a nicking endonuclease that recognizes the NERS in one strand and the location corresponding to a NS produced by the NE in the other strand is no more than 25 nucleotides. The present invention also provides a kit for identifying a genetic variation at a defined location in a single-stranded target nucleic acid, comprising: A first oligonucleotide primer (ODNP) and a second oligonucleotide primer (ODNP) wherein (a) the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, (b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, (c) the first and the second ODNPs each further comprise a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS), and (d) the distance between the a nicking site (NS) produced by a nicking endonuclease that recognizes the NERS in one strand and the location corresponding to a NS produced by the NE in the other strand is no more than 25 nucleotides.

Optionally, in these kits for identifying a genetic variation, *e.g.*, a SNP, at a defined location in a target nucleic acid, the following criteria and/or components may be present, where any two criteria and/or components may be combined in describing a kit or the contents thereof: in the double-stranded nucleic acid molecule that is amplified according to a method of the invention using the first and the second ODNPs as primers and the target nucleic acid as a template, the distance between a nicking site (NS) produced by a nicking endonuclease (NE) that recognizes the NERS in one strand and the location corresponding to a cleavage site produced by a restriction endonuclease (RE) that recognizes the RERS in the other strand is no more than 40, or no more than 39, or no more than 38, or no more than 37, or no more than 36, or no more than 35, or no more than 34, or no more than 33, or no more than 32, or no more than 31, or no more than 30, or no more than 29, or no more than 28, or no more than 27, or no more than 26, or no more than 25, or no more than 24, or no more than 23, or no more than 22, or no more than 21, or no more than 20, or no more than 19, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10, or no more than 9, or no more than 8, or no more than 7, or no more than 6, or no more than 5 nucleotides; in the double-stranded nucleic acid molecule that is amplified using the first and the second ODNPs as primers and the target nucleic acid as a template, the distance between a nicking site (NS) produced by a nicking endonuclease (NE) that recognizes the NERS in one strand and the location corresponding to a NS produced by the NE in the other strand is no more than 40, or no more than 39, or no more than 38, or no more than 37, or no more than 36, or no more than 35, or no more than 34, or no more than 33, or no more than 32, or no more than 31, or no more than 30, or no more than 29, or no more than 28, or no more than 27, or no more than 26, or no more than 25, or no more than 24, or no more than 23, or no more than 22, or no more than 21, or no more than 20, or no more than 19, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10, or no more than 9, or no more than 8, or no more than 7, or no more than 6, or no more than 5 nucleotides; the NERS has the sequence 5'GAGTC3'; N.BstNB I; a buffer for N.BstNB I; a restriction endonuclease (RE) that recognizes the RERS; a buffer for the RE; a DNA polymerase; a DNA polymerase selected from exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, and exo$^-$ Vent polymerase; a buffer for the DNA polymerase; a buffer for both the NE and the DNA polymease; instructions for using the kit; a liquid chromatography column; a reverse phase liquid chromatography column; Buffer A that

comprises water and an ammonium salt of a secondary or tertiary amine complexed with an organic or inorganic acid, where the amine component of the ammonium salt is optionally selected from the group consisting of triethylamine, diallylamine, diisopropylamine, N,N-dimethyl-N-cyclohexylamine, N,N-dimethyl-N-isopropylamine, and N,N-dimethyl-N-butylamine, where the secondary or tertiary amine is optionally complexed with an organic acid, and the organic acid is optionally selected from acetic acid, propionic acid, and halogenated versions thereof; Buffer B that comprises Buffer A and organic solvent, where the organic solvent is optionally selected from methanol and acetonitrile; a deoxyribonucleoside triphosphate; a modified deoxyribonucleoside triphosphate; a control template nucleic acid molecule and a control ODNP pair; trehalose; an oligonucleotide standard; and/or an access code for software useful in designing or ordering the ODNP pair. For example, a kit may further include a RE that recognizes the RERS and a buffer for the RE, a NE that recognizes the NERS and a buffer for the NE, and a DNA polymerase and a buffer for the DNA polymerase.

In addition to providing a method for identifying a genetic variation at a defined location in a target nucleic acid, the present invention provides a method for the multiplex identification of (multiple) genetic variations at defined locations in (multiple) target nucleic acids. Thus, a single sample may contain two or more nucleic acids, each having a genetic variation. By proper selection of the primers, the presence or absence of a nucleic acid with a particular genetic variation, or the presence of a particular genetic variation within a particular sample of nucleic acid, where the sample contains multiple potential genetic variations, can be analyzed. For instance, in one aspect, the present invention provides a method for the multiplex identification of genetic variations at defined locations in target nucleic acids, where the method includes: (a) For each genetic variation in a target nucleic acid under investigation, (i) forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein a) if the target nucleic acid under investigation is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a sequence of a sense strand of nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises an optional sequence of one strand of a restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, while b) if the target nucleic acid under investigation is a single-stranded

nucleic acid, then the first ODNP comprises a sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises an optional sequence of one strand of a RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation; (ii) extending the first and the second ODNPs to produce a fragment encompassed by the first ODNP and the second ODNP, wherein the extension reaction may be performed either individually for each target nucleic acid or jointly for more than one target nucleic acids. (b) Optionally, *i.e.*, this step may or may not be performed according to the inventive method, cleaving the extension products with a RE that recognizes the RERS so a to provide a digestion product. (c) Amplifying single-stranded nucleic acid fragments using one strand of each extension product of step (a)(ii) or each digestion product of step (c) as template in the presence of a nicking endonuclease that recognizes the NERS. (d) Characterizing the amplified single-stranded fragments to thereby identify the genetic variations. In a related aspect, the present invention provides a method for the multiplex identification of single nucleotide polymorphisms at defined locations in target nucleic acids, where the method includes: (a) for each target nucleic acid, (i) forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, the first ODNP comprises a sequence of a sense strand of nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the SNP, and the second ODNP comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the SNP, or if b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the SNP, and the second ODNP comprises a sequence of one strand of a RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the SNP; (ii) extending the first and the second ODNPs to produce a fragment encompassed by the first ODNP and the second ODNP, wherein the extension reaction may be performed either individually for each target nucleic acid or jointly for more than one target nucleic acids. (b) Cleaving the

extension products with a RE that recognizes the RERS, so as provide a digestion product. (c) Amplifying single-stranded nucleic acid fragments, preferably under isothermal conditions, using one strand of each extension product of step (a)(ii) or each digestion product of step (c) as a template in the presence of a nicking endonuclease that recognizes the NERS, wherein the single-stranded nucleic acid fragments each preferably contain no more than 17 nucleotides. (d) Characterizing the single-stranded fragments using at least partially liquid chromatograph and/or mass spectrometry to thereby identify the genetic variations.

In a related aspect, the present invention provides a method for the multiplex identification of genetic variations at defined locations in target nucleic acids, where the method includes: (a) for each target nucleic acid, (i) forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein a) if and when the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, however, if and when b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, where either event, the first and the second ODNPs each further comprise a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS), (ii) extending the first and the second ODNPs to produce a fragment having two NERSs, wherein the extension reaction may be performed either individually for each target nucleic acid or jointly for more than one target nucleic acids. (b) Amplifying single-stranded nucleic acid fragments using one strand of each extension product as the template in the presence of a nicking endonuclease (NE) that recognizes the NERS. (c) Characterizing the single-stranded fragments of step (b) to thereby identify the genetic variations. In a related aspect, the present invention provides a method for the multiplex identification of single nucleotide polymorphisms (SNPs) at defined locations in target nucleic acids, where the method includes: (a) for each target nucleic acid, (i) forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein if a) the target

nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the SNP, however, if b) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the SNP, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the SNP, where in either event the first and the second ODNPs each further comprise a sequence of a sense strand of a same nicking endonuclease recognition sequence (NERS); and (ii) extending the first and the second ODNPs to produce a fragment having two NERSs, wherein the extension reaction may be performed either individually for each target nucleic acid or jointly for more than one target nucleic acids. (b) Amplifying single-stranded nucleic acid fragments under an isothermal condition using one strand of each extension product as template in the presence of a nicking endonuclease (NE) that recognizes the NERS, preferably wherein the single-stranded nucleic acid fragments each contain no more than 17 nucleotides. (c) Characterizing the single-stranded fragments of step (b) using at least partially liquid chromatography and/or mass spectrometry to thereby identify the genetic variations.

In the methods for the multiplex identification of (multiple) genetic variations at defined locations in (multiple) target nucleic acids, one or more additional criteria may be used to describe the method, where these additional criteria are disclosed herein, and include those criteria set forth above in connection with the method of the invention directed to identifying a genetic variation, *e.g.*, a SNP, at a defined location in a target nucleic acid, where any one or more of these criteria may be utilized in describing the inventive method. For example, some of the criteria that may be used to describe the inventive method for the multiplex identification of (multiple) genetic variations at defined locations in (multiple) target nucleic acids include: at least one of the genetic variations is a SNP; at least one of the genetic variations is associated with a disease, *e.g.*, a human genetic disease; at least one of the genetic variations is associated with drug resistance of a pathogenic microorganism; at least one of the target nucleic acids is genomic nucleic acid; at least one of the target nucleic acids is cDNA; at least one of the target nucleic acids is derived from the genome of a pathogenic virus; at least one of the target nucleic acids is derived from the genome of a pathogenic

bacterium; at least one of the target nucleic acids is derived from the episome of a pathogenic bacterium; the nucleotide sequences of the first ODNPs that are at least substantially complementary to first strands of the target nucleic acids if the target nucleic acids are double-stranded, or identical to the target nucleic acids if the target nucleic acids are single-stranded, are at least 8, or are at least 9, or are at least 10, or are at least 11, or at least 12, or at least 13, or at least 14, or at least 15 nucleotides in length; the nucleotide sequences of the second ODNPs that are at least substantially complementary to second strands of target nucleic acids if the target nucleic acids are double-stranded, or to the target nucleic acids if the target nucleic acids are double-stranded, are at least 8, or are at least 9, or are at least 10, or are at least 11, or at least 12, or at least 13, or at least 14, or at least 15 nucleotides in length; the RERS is recognizable by a Type II restriction endonuclease; the RERS is recognizable by a Type IIs restriction endonuclease; the RERS is recognizable by Bpm I; the RERS is recognizable by an interrupted restriction endonuclease; the NERS is 5'GAGTC3'; extension is performed by a polymerase chain reaction; amplification is performed in the presence of a DNA polymerase, *e.g.*, exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, or exo$^-$ Vent polymerase; amplification is performed under isothermal conditions; the single-stranded nucleic acid fragments of step (c) contains no more than 20, or no more than 19, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10 nucleotides; characterization includes one or more techniques selected from mass spectrometry, liquid chromatography, fluorescence polarization, electron ionization, gel electrophoresis and capillary electrophoresis; characterization is performed by one or more methods that include mass spectrometry; the method further includes the step of combining products of at least two extension reactions of step (a)(ii) before performing step (c); combining products of each extension reaction of step (a)(ii) before performing step (c); the method includes the step of cleaving the extension products with a RE that recognizes the RERS so as to provide a digestion product; the method omits the step of cleaving the extension products with a RE that recognizes the RERS, so that digestion product is not formed.

In another set of related embodiments, the present invention provides kits that may be used in a method for the multiplex identification of genetic variations at locations in target nucleic acids. In one of these embodiments, the kit includes, for each target nucleic acid under investigation: (a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the kit has (i) a first oligonucleotide primer (ODNP) that comprises a sequence of a sense strand of a

nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located at 3' to the complement of the genetic variation, and (ii) a second ODNP that comprises an optional sequence, *i.e.*, that optionally includes a nucleotide sequence, of, *i.e.*, identical to, one strand of a restriction endonuclease recognition sequence (RERS) and has a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation; and, if (b) one or more of the target nucleic acids is single-stranded, then for each single-stranded target nucleic acid, the kit includes (i) a first ODNP that comprises a sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and (ii) a second ODNP that optionally comprises a nucleotide sequence identical to one strand of a RERS, and necessarily comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation. In a related aspect, the present invention provides a kit useful in a method for the multiplex identification of genetic variations of genetic variations at defined locations in target nucleic acids, where the kit includes, for each target nucleic acid under investigation: (a) if a particular target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then for each of this type of target nucleic acid the kit has (i) a first ODNP that comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and (ii) a second ODNP that comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation; however, for (b) each target nucleic acid under investigation that is a single-stranded nucleic acid, the kits includes (i) a first ODNP that comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and (ii) a second ODNP that comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, wherein each set of first and second ODNPs further comprise a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS).

In kits of the present invention that are particularly useful in a method for the multiplex identification of genetic variations at locations in target nucleic acids, one or more of the following criteria may be used in describing the kits, where these criteria may be combined in any combination, and where the kits are designed to be used in the

methods for multiplex identification of genetic variations as described herein, and where these criteria are non-limiting, *i.e.*, other criteria may appear elsewhere herein: the NERS as present in a nucleic acid molecule has the sequence 5'GAGTC3'; N.BstNB I is present in the kit; the kit contains a buffer for N.BstNB I; the kit contains a restriction endonuclease (RE) that recognizes the RERS; the kit contains a buffer for the RE; the kit contains a DNA polymerase; the kit contains one or more of exo⁻ Bst polymerase, exo⁻ Bca polymerase, 9°Nm™ DNA polymerase, or exo⁻ Vent polymerase; the kit contains a buffer for the DNA polymerase; the kit contains a buffer that is suitable for both the nicking endonuclease and the DNA polymerase; the kit contains a buffer suitable for the NE and a different buffer suitable for the DNA polymerase; instructions for using the kit; a liquid chromatography column; a reverse phase liquid chromatography column; Buffer A that comprises water and an ammonium salt of a secondary or tertiary amine complexed with an organic or inorganic acid, where the amine component of the ammonium salt is optionally selected from the group consisting of triethylamine, diallylamine, diisopropylamine, N,N-dimethyl-N-cyclohexylamine, N,N-dimethyl-N-isopropylamine, and N,N-dimethyl-N-butylamine, where the secondary or tertiary amine is optionally complexed with an organic acid, and the organic acid is optionally selected from acetic acid, propionic acid, and halogenated versions thereof; Buffer B that comprises Buffer A and organic solvent, where the organic solvent is optionally selected from methanol and acetonitrile; a deoxyribonucleoside triphosphate; a modified deoxyribonucleoside triphosphate; a control template nucleic acid molecule and a control ODNP pair; trehalose; an oligonucleotide standard; and/or an access code for software useful in designing or ordering the ODNP pair. For example, a kit may further include a RE that recognizes the RERS and a buffer for the RE, a NE that recognizes the NERS and a buffer for the NE, and a DNA polymerase and a buffer for the DNA polymerase.

In another aspect, the present invention provides methods for determining the presence, or the absence, of a target nucleic acid in a sample. These methods are particularly useful in a diagnosis situation, where it is desirable to determine whether a particular sample, *e.g.*, a sample from a patient, contains a particular target nucleic acid, *e.g.*, a nucleic acid that is commonly found in a disease-causing organism. In one embodiment, this method includes: (a) Admixing nucleic acid molecules present in a sample with a partially double-stranded oligonucleotide probe. The nucleic acid molecules and probe are admixed under conditions that allow for hybridization of the probe to target nucleic acid that may be present in the sample. The probe may be characterized as having: (i) A nicking agent recognition sequence

(NARS) that is recognizable by a nicking agent (NA), where the nicking agent cleaves at a nicking site (NS), and (ii) either a 5' overhang in the strand that either contains the NS or may be extended so as to contain the NS, or a 3' overhang in the strand that either does not contain the NS or cannot be extended to contain the NS, where in either case the overhang includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid. (b) Removing unhybridized probe from the admixture of step (a). (c) Performing an amplification reaction in the presence of the NA and a DNA polymerase to provide an amplification product, where the amplification reaction is preferably conducted on probe that has hybridized to target nucleic acid. (d) Detecting the presence, or the absence, of the amplification product of step (c) to thereby determine the presence, or the absence, of the target nucleic acid in the sample. In a related aspect directed to a method for determining the presence, or the absence, of a target nucleic acid in a biological sample, the method includes (a) Admixing the nucleic acid molecules of the sample with a partially double-stranded oligonucleotide probe to allow for hybridization of the probe to the target nucleic acid, where the probe comprises (i) a nicking endonuclease recognition sequence (NERS) that may be recognized by a nicking endonuclease (NE), where after recognition of the NERS, the NE produces a nick in a double stranded nucleic acid at a location termed the nicking site (NS), and (ii) either a 5' overhang in the strand that of the probe that contains the NS or the strand of the probe that is extended and upon extension contains the NS, or a 3' overhang in the strand that does not contain the NS and that will not be extended to form a NS, wherein the overhang (whether 5' or 3') comprises a nucleotide sequence that is at least substantially complementary to a target nucleic acid. (b) Removing unhybridized probe from the admixture of step (a). (c) Performing an amplification reaction, preferably under isothermal conditions, in the presence of the NE and a DNA polymerase to provide an amplification product, where the amplification product preferably contains no more than 17 nucleotides. (d) Detecting the presence, or the absence, of the amplification product of step (c), preferably using, at least partially, liquid chromatography and/or mass spectrometry, where the detecting step allows the determination of whether the target nucleic acid was, or was not, present in the sample. In a related aspect, the present invention provides a method for determining the presence or the absence of a target nucleic acid in a sample, where the method includes: (A) Forming a mixture that includes: (i) nucleic acid molecules from a sample of interest, where these nucleic acid molecules may, or may not, include a target nucleic acid; (ii) a single-stranded nucleic acid probe comprising from 3' to 5': (a) a first nucleotide sequence that is at least substantially complementary to a

nucleotide sequence present in a target nucleic acid, (b) the nucleotide sequence of the antisense strand of a nicking agent recognition sequence (NARS), and (c) a second sequence; (iii) a nicking endonuclease (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s). (B) Maintaining the mixture of step (A) under conditions that amplify a single-stranded nucleic acid fragment using the single-stranded nucleic acid probe as a template if the target nucleic acid is present in the sample. In one embodiment, the single-stranded nucleic acid fragment is at most 25 nucleotides in length. (C) Detecting the presence or the absence of the single-stranded nucleic acid fragment amplified in step (B) to determine the presence, or the absence, of the target nucleic acid in the sample. In a related embodiment, the present invention provides a method for determining the presence or the absence of a target nucleic acid in a sample, where the method includes: (A) forming a mixture that includes: (i) the nucleic acid molecules from the sample; (ii) a single-stranded nucleic acid probe comprising from 3' to 5': (a) a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid under study, and (b) the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS), and (iii) a nicking endonuclease (NA) that recognizes the first NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s). (B) Maintaining the mixture under conditions that allow for amplification of a single-stranded nucleic acid fragment that is at least substantially complementary to the target nucleic acid if the target nucleic acid is present in the sample. In one embodiment, this single-stranded nucleic acid fragment is at most 25 nucleotides in length. (C) Detecting the presence or the absence of the single-stranded nucleic acid fragment amplified in step (B) in order to determine the presence, or the absence, of the target nucleic acid in the sample. In a related aspect, the present invention also provides a method for determining the presence or absence of a target nucleic acid in a sample, where the target nucleic acid includes a nicking agent recognition sequence (NARS). The method includes: (A) Forming a mixture that includes: (i) the nucleic acid molecules of the sample, and (ii) a nicking agent (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s). (B) Maintaining the mixture of (A) under conditions that allow for the amplification of a single-stranded nucleic acid fragment using a portion of the target nucleic acid as a template if the target nucleic acid is present in the sample. (C) Detecting the presence or absence of the single-stranded nucleic acid fragment amplified in step (B) in order to determine the presence or absence of the target nucleic acid in the sample. In a related aspect, the present invention provides a method for

determining the presence or absence of a target nucleic acid, where the target has a nucleotide sequence that includes the nucleotide sequence of a first nicking agent recognition sequence (NARS). The method includes: (A) Forming a mixture that includes: (i) the nucleic acid molecules of the sample, (ii) a single-stranded nucleic acid probe that is substantially identical to one strand of the target nucleic acid and comprise a nucleotide sequence that is the same nucleotide sequence as the antisense strand of the NARS, and (iii) a nicking agent (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s). (B) Maintaining the mixture under conditions that allow for amplification of a single-stranded nucleic acid fragment using a portion of the probe as a template in the event that the target nucleic acid is present in the sample. (C) Detecting the presence or absence of the single-stranded nucleic acid fragment of step (B) to determine the presence or absence of the target nucleic acid in the sample. In a related method, the present invention provides a method for determining the presence or absence of a target nucleic acid in a sample. This method includes: (A) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the nucleic acid molecules of the sample, wherein (i) if the target nucleic acid may be a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises a nucleotide sequence of a sense strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a first portion of the first strand of the target nucleic acid, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a second portion of the second strand of the target nucleic acid and comprises a sequence of the sense strand of a second RERS, the second portion being located 3' to the complement of the first portion in the second strand of the target nucleic acid, however, if (ii) the target nucleic acid may be a single-stranded nucleic acid, then the first ODNP comprises a nucleotide sequence of a sense strand of a first RERS and a nucleotide sequence at least substantially identical to a first portion of the target nucleic acid, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a second portion of the target nucleic acid and comprises a sequence of the sense strand of a second RERS, the second portion being located 5' to the first portion in the target nucleic acid. (B) Subjecting the mixture of (A) to conditions that, if the target nucleic acid is present in the sample, will: (i) extend the first and the second ODNPs to produce an extension product comprising both the first and the second RERSs; and (ii) amplify a single-stranded nucleic acid fragment using one strand of the extension product of step (B)(i) as a template in the presence of one or more restriction endonucleases (REs) that recognize the first and the

second RERSs. In one embodiment, the single-stranded nucleic acid fragment is at most 25 nucleotides in length. (C) Detecting the presence or absence of the single-stranded nucleic acid fragment amplified in step (B) to determine the presence or absence of the target nucleic acid in the sample. Optionally, the first RERS is identical to the second RERS. In a related aspect, the present invention provides a method for determining the presence or absence of a target nucleic acid in a sample. This method includes: (A) Forming a mixture that includes a first oligonucleotide primer (ODNP), a second ODNP, and the nucleic acid molecule of the sample, wherein (i) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then the first ODNP comprises the nucleotide sequence present in the sense strand of a first nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a first portion of the first strand of the target nucleic acid, and the second ODNP comprises a nucleotide sequence at least substantially complementary to a second portion of the second strand of the target nucleic acid and further comprises the nucleotide sequence of the sense strand of a second NERS, where the second portion is located 3' to the complement of the first portion in the second strand of the target nucleic acid, or, if (ii) the target nucleic acid is a single-stranded nucleic acid, then the first ODNP comprises the nucleotide sequence of the sense strand of a first NERS and a nucleotide sequence that is at least substantially identical to a first portion of the target nucleic acid, while the second ODNP comprises a nucleotide sequence at least substantially complementary to a second portion of the target nucleic acid and further comprises the sequence of the sense strand of a second NERS, where the second portion is located 5' to the first portion in the target nucleic acid. (B) Subjecting the mixture to conditions that, if the target nucleic acid is present in the sample, (i) extend the first and the second ODNPs to produce an extension product comprising both the first and the second NERSs; and (ii) amplify a single-stranded nucleic acid fragment using one strand of the extension product of step (B)(i) as a template in the presence of one or more nicking endonucleases (NEs) that recognize the first and the second NERSs. In one embodiment, the single-stranded nucleic acid fragment is at most 25 nucleotides in length. (C) Detecting the presence or absence of the single-stranded nucleic acid fragment amplified in step (B) in order to determine the presence or absence of the target nucleic acid in the sample. In one embodiment, the first and second NERSs are identical. In another related aspect, the present invention provides a method for determining the presence or absence of a target nucleic acid in a sample, where the method includes: (A) Forming a mixture that includes: (i) the nucleic acid molecules

from the sample, and (ii) a single-stranded nucleic acid probe that comprises, from 3' to 5', a nucleotide sequence that is at least substantially complementary to the portion of the target nucleic acid that is located at or near the 5' terminus of the target nucleic acid, and the sequence of the antisense strand of a nicking agent recognition sequence (NARS). (B) Separating the probe molecules that are hybridized to the target, if any, from those that are not hybridized to the target. (C) Performing an amplification reaction in the presence of the probe molecules that are hybridized to the target, if any, and a nicking agent (NA) that recognizes the NARS. (D) Detecting the presence or absence of a single-stranded nucleic acid fragment amplified using the single-stranded nucleic acid probe as a template in step (C), to thereby determine the presence or absence of the target nucleic acid in the sample. In yet another related aspect of the present invention directed to a method for determine the presence or absence of a target nucleic acid in a sample, the invention provides a method that includes: (A) Forming a mixture that includes: (i) the nucleic acid molecules of the sample, (ii) a single-stranded nucleic acid probe immobilized to a solid support via its 5' terminus, where the probe comprises, in the 3' to 5' direction: (a) a nucleotide sequence that is at least substantially complementary to a portion of the target located at or near the 3' end of the target, and (b) the sequence of the sense strand of a nicking agent recognition sequence. (B) Removing the nucleic acid molecules of the sample that are not hybridized to the probe from contact with the solid support. (C) Performing an amplification reaction in the presence of the probe that is immobilized to the solid support of step (B) and a nicking agent that recognizes the nicking agent recognition sequence. (D) Detecting the presence or absence of a single-stranded nucleic acid fragment that was amplified in step (C) using a portion of the target nucleic acid as a template, in order to determine the presence or absence of the target nucleic acid in the sample.

In various embodiments of the invention directed to determining the presence, or the absence, of a target nucleic acid in a sample, one or more of the following criteria may be used to describe the embodiment, where any two, three, four, etc. criteria may be combined to describe the embodiment, and where these embodiments are in addition to other embodiments described elsewhere herein: step (d) is performed at least partially by a technique selected from the group consisting of luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, gel electrophoresis, and capillary electrophoresis, where the characterization may be performed by one, two, three or more of these techniques, and/or the characterization may be performed by one, two, three or more of these techniques in addition to any other characterization techniques

that may be suitable and/or readily available; the target nucleic acid(s) in the sample are all single-stranded; the target nucleic acid(s) in the sample are all double-stranded; the target nucleic acid(s) in the sample are a mixture of single-stranded and double-stranded target nucleic acids; the biological sample that contains the target nucleic acid, or from which the target nucleic acid is derived, is from a mammal, *e.g.*, a human, a cow, a horse, a dog, a sheep, a deer, a pig,, or may otherwise be described as coming from a patient; the target nucleic acid is a nucleic acid molecule originating from a host selected from the group consisting of a bacterium, and/or a virus, and/or a fungus and /or a parasite; the NA is N.BstNB I; amplification is performed in the presence of a DNA polymerase, where the following are exemplary suitable DNA polymerases, and any two or more of the following DNA polymerases may be combined to form a group from which the DNA polymerase used in the invention may be selected: exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm™ DNA polymerase, and T4 DNA polymerase, for example, the polymerase may be selected from the group consisting of exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm™ DNA polymerase, or exo$^-$ Vent polymerase; amplification is carried out under isothermal conditions; amplification is carried out in the presence of a labeled deoxynucleoside triphosphate; amplification is carried out in the presence of a labeled deoxynucleoside triphosphate where the labeled deoxynucleoside triphosphate is a deoxynucleoside triphosphate linked to a label, where the label may be a radiolabel and/or an enzyme and/or a fluorescent dye and/or digoxigenin and/or biotin, where these are exemplary labels; detecting whether or not the amplification product has been formed, and if it has been formed then further detecting some or all of the nucleotide sequence of the amplification product, is accomplished at least partially by one or more techniques that include luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, gel electrophoresis and capillary electrophoresis, for example, the analysis may be carried out by luminescence spectroscopy, fluorescence spectroscopy or mass spectrometry; the amplification product, when present, contains no more than 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or 20, or 21, or 22, or 23, or 24, or 25, or 26, or 27, or 28, or 29, or 30, or 31, or 32, or 33, or 34, or 35, or 36, or 37, or 38, or 39, or 40, or 41, or 42, or 43, or 44, or 45, or 46, or 47, or 48, or 49, or no more than 50 nucleotides; the nucleic acid molecules present in the sample are immobilized onto a solid support; the amplification product is immobilized to a solid

substrate before being detected; the single-stranded nucleic acid probe is immobilized to a solid support.

As a means for conducting a method for the determination of whether a particular target nucleic acid is present or absent in a sample, the present invention provides various kits. In various embodiments, the kit include one or more of the following components, where any two or more of these components may be combined in describing a kit according to the present invention: a partially double-stranded oligonucleotide probe as described in the foregoing methods directed to determining the presence or absence of a target nucleic acid in a sample; two single-stranded oligonucleotides which, upon their annealing with one another, produces a double-stranded oligonucleotide probe as described in the foregoing methods directed to determining the presence or absence of a target nucleic acid in a sample; a kit wherein the NARS is recognized by N.BstNB I; a kit that includes N.BstNB I; a buffer for N.BstNB I; a NA that recognizes the NARS; a DNA polymerase; a buffer for the DNA polymerase; a buffer suitable for both the NE and the DNA polymerase; a buffer suitable for the NE and a different buffer suitable for the DNA polymerase; at least one of exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, and exo$^-$ Vent polymerase; a labeled deoxynucleoside triphosphate, which may be, for example, a deoxynucleoside triphosphate linked to a radiolabel, or linked to an enzyme, or linked to a fluorescent dye, or linked to digoxigenin, or linked to biotin; instructions for using the kit; a liquid chromatography column; a reverse phase liquid chromatography column; Buffer A that comprises water and an ammonium salt of a secondary or tertiary amine complexed with an organic or inorganic acid, where the amine component of the ammonium salt is optionally selected from the group consisting of triethylamine, diallylamine, diisopropylamine, N,N-dimethyl-N-cyclohexylamine, N,N-dimethyl-N-isopropylamine, and N,N-dimethyl-N-butylamine, where the secondary or tertiary amine is optionally complexed with an organic acid, and the organic acid is optionally selected from acetic acid, propionic acid, and halogenated versions thereof; Buffer B that comprises Buffer A and organic solvent, where the organic solvent is optionally selected from methanol and acetonitrile; a deoxyribonucleoside triphosphate; a modified deoxyribonucleoside triphosphate; a control template nucleic acid molecule and a control ODNP pair; trehalose; an oligonucleotide standard; an access code for software useful in designing or ordering probes; a pin coated with a chemical layer bindable to a nucleic acid molecule, or example, a pin coated with poly(ethyleneimine) layer, or a pin coated with PDADMAC, *i.e.*, polydiallyldimethylammonium chloride; and/or trehalose.

The present invention also provides for the multiplex determination of the presence, or the absence, of multiple target nucleic acids in a biological sample. In one embodiment, this method includes: (a) Admixing the nucleic acid molecules from a biological sample with partially double-stranded oligonucleotide probes to allow for hybridization of the probes to any target nucleic acids that may be present in the sample, wherein each probe comprises: (1) a nicking agent recognition sequence (NARS) that is recognizable by a NA, where after recognition of the NARS, the NA will nick a double-stranded portion of nucleic acid at a located termed the nicking site (NS), (2) either a 5' overhang in the strand that contains the NS or will be extended to form a NS, or, a 3' overhang in the strand that does not contain the NS and is not extended according the method of the present invention to form a NS, where the overhang (either 5' or 3') comprises a nucleotide sequence that is at least substantially complementary to a target nucleic acid of interest, and (3) a sequence, which is located in the strand that does not contain the NS and is not extended according to the inventive method so as to form a NS, which is located 5' to the position corresponding to the NS, where this sequence uniquely correlates to the target nucleic acid to which the overhang is at least substantially complementary. The sequence of (3) effectively provides a signal that is unique for a particular target nucleic acid, so that detection of the signal signifies the presence of the particular target nucleic acid that is "coded for" by the sequence of (3). (b) Removing the unhybridized probes from the admixture of step (a). (c) Performing an amplification reaction in the presence of the NA and a DNA polymerase. (d) Detecting the presence, or the absence, of amplification products of step (c) to determine the presence of target nucleic acids correlating to the sequences of the amplification products in the biological sample. In another related aspect, the present invention provides a method for the multiplex determination of the presence, or the absence, of target nucleic acids in a biological sample, where the method includes: (a) admixing the nucleic acid molecules of the biological sample with partially double-stranded oligonucleotide probes, under conditions that allow for hybridization of the probes to the target nucleic acids, wherein each probe comprises: (1) a nicking endonuclease recognition sequence (NERS) that is recognizable by a nicking endonuclease (NE), where the NE may nick a double-stranded nucleic acid molecule at a site termed the nicking site (NS) located near the NERS, (2) either a 5' overhang in the strand that does not contain the NS and will not be extended according to the method of the present invention to contain the NS, or, a 3' overhang in the strand that does not contain the NS and will not be extended according to the method of the present invention to form a NS, wherein the overhang (whether a 5' or 3' overhang) comprises a

nucleotide sequence that is at least substantially complementary to a target nucleic acid of interest, and (3) a nucleotide sequence, located within that strand which does not contain the NS and will not be extended by the method of the present invention to form a NS, where the sequence is located at 5' to the position corresponding to the NS, and the sequence uniquely correlates to a target nucleic acid to which the overhang is at least substantially complementary. The sequence of (3) effectively provides a signal that is unique for a particular target nucleic acid, so that detection of the signal signifies the presence of the particular target nucleic acid that is "coded for" by the sequence of (3). (b) Removing the unhybridized probes from the admixture of step (a). (c) Performing an amplification reaction, preferably under isothermal conditions, in the presence of the NE and a DNA polymerase to provide amplification product(s). In various embodiments, the amplification product(s) each contain no more than 17, 16, 15, 14, 13, or 12 nucleotides. (d) Detecting the presence, or the absence, of the amplification product(s) of step (c) to determine the presence of target nucleic acids correlating to the sequences of the amplification product(s) in the biological sample. Optionally, the detecting step entails using, at least partially, liquid chromatography and/or mass spectrometry.

In further describing the methods of the present invention directed to the multiplex determination of the presence of absence of target nucleic acids in a biological sample, one or more additional criteria may be utilized. In part, the criteria set forth above in connection with the methods of the present invention directed to determining the presence or absence of a target nucleic acid in a sample, may also be used, in any combination, to further characterize the method of the present invention directed to the multiplex determination of the presence of absence of target nucleic acids in a biological sample. For example, the sample may be from a patient. Other criteria appropriate to further characterizing the multiplex determination of the presence of absence of target nucleic acids in a biological sample may appear elsewhere herein. Also, the method may be conducted utilizing nucleic acid molecules that are immobilized on a solid support. For example, the probes may be immobilized to unique locations on a solid support in an array format. Alternatively, the probes may be indiscriminantly immobilized on a solid support. In either event, the sample containing the target nucleic acids would be contacted with the immobilized probes. Alternatively, the target nucleic acids may be immobilized onto a solid support and then the probes added to the support so that they contact the immobilized targets.

The present invention also provides kits that may be utilized in a method as described herein for the multiplex determination of the presence, or absence, of

several target nucleic acids in a biological sample. These kits may include any or all of the components previously described for kits that are useful in a non-multiplexing method, *i.e.,* in a method for determining the presence or absence of a target nucleic acid in a biological sample. The kits that are particularly useful for a multiplexing method may simply include more probes, or more single-stranded oligonucleotides from which the probes are prepared, in order to test for the presence of multiple targets. Also, when the kit contains multiple probes (or probe components) the multiple probes or components thereof are designed so as to provide unique signals in response to different targets. In one aspect, the kits contains 2 different probes (or probe components) while in other aspects the kits contain 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or 20, or 22, or 24, or 26, or 28, or 30, or 35, or 40, or 45, or 50, or 55, or 60, or 65, or 70, or 75, or 80, or 85, or 90, or 95, or 100 different probes or probe components. The kit may include immobilized probes, *e.g.,* probes immobilized at unique locations across a solid surface, or probes immobilized indiscriminantly across a solid surface.

When a nucleic acid is immobilized to a solid support according to a method or composition or kit of the present invention, suitable solid supports include materials formed in whole or part from silicon, glass, paper, ceramic, metal, metalloid, and plastic. In one embodiment, the solid support is coated with a chemical layer that binds a nucleic acid molecule. A suitable chemical layer is formed from poly(ethyleneimine). Another suitable chemical layer is formed from PDADMAC, *i.e.,* polydiallyldimethylammonium chloride.

In another aspect, the present invention provides a method for determining the ratio of the number of first nucleic acid molecules to the number of second nucleic acid molecules in a population of nucleic acid molecules, where the nucleotide sequences of the first and the second nucleic acid molecules are identical except at a defined location. This method includes: (a) Admixing a first oligonucleotide primer (ODNP), a second ODNP and the nucleic acid population, wherein if (i) the first and the second nucleic acids are double stranded, each having a first strand and a second strand, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleic acid sequence present in the first strand of the first nucleic acid molecule, where this sequence is located 3' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleic acid sequence present in the second strand of the first nucleic acid, this nucleotide sequence also being located 3' to the defined location, however, if (ii) the first and the second nucleic acids are single

stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleic acid sequence present in the first nucleic acid, where this nucleotide sequence is located 5' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleic acid sequence present in the first nucleic acid, where this sequence is located 3' to the defined location. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS). (b) Extending the first and the second ODNPs to produce fragments that encompass the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragments in the presence of a DNA polymease and a nicking agent (NA) that recognizes the NARS. (d) Determining the ratio of the number of the single-stranded nucleic acid fragments of step (c) derived from the first nucleic acid relative to the number of single-stranded nucleic acid fragments derived from the second nucleic acid. The present invention also provides a method for determining the ratio of the number of first nucleic acid molecules to the number of second nucleic acid molecules in a nucleic acid population, where the nucleotide sequences of the first and the second nucleic acid molecules are identical except at a defined location, where this method includes: (a) Admixing a first ODNP, a second ODNP and the population of nucleic acids, wherein if (i) the first and the second nucleic acids are double-stranded (*i.e.,* each has a first strand and a second strand) then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the first strand of the first nucleic acid, where this nucleotide sequence is located 3' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the second strand of the first nucleic acid, where this nucleotide sequence is located 3' to the defined location, or, if (ii) the first and the second nucleic acids are each single-stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleic acid sequence present in the first nucleic acid, where this nucleotide sequence is located 5' to the defined location in the first nucleic acid, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first nucleic acid that is located 3' to the defined location of the first nucleic acid. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking endonuclease recognition sequence (NERS). (b) Extending the first and the second ODNPs so as to produce fragments that encompass the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragments, preferably

under isothermal conditions, in the presence of a DNA polymease and a nicking endonuclease (NE) that recognizes the NERS. In various embodiments the single-stranded nucleic acid fragments each contain no more than 20, or 18, or 17, or 16, or 15, or 14, or 13, or 12, or 11, or 10, etc. nucleotides. (d) Determining the ratio of the number of the single-stranded nucleic acid fragments of step (c) that were derived from the first nucleic acid molecules, relative to the number that were derived from the second nucleic acid molecules. This determination may be made using, for example, one or both of liquid chromatography and mass spectrometry. In related methods, the present invention provides for determining the allelic frequency of a target nucleic acid molecule with a genetic variation at a defined location, where this target nucleic acid molecule is present in a nucleic acid population. This method includes: (a) Admixing a first oligonucleotide primer (ODNP), a second ODNP and the nucleic acid population, wherein if (i) the target nucleic acid is double stranded such that it has a first strand and a second strand, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location in the target nucleic acid, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location in the target nucleic acid, however, if (ii) the target nucleic acids are single stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleotide sequence present in the target nucleic acid, where this nucleotide sequence is located 5' to the defined location in the target nucleic acid, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid, where this nucleotide sequence is located 3' to the defined location in the target nucleic acid, and, regardless of whether the target nucleic acid is single- or double-stranded, at least one of the first ODNP and the second ODNP further comprises the sequence of the sense strand of a nicking agent recognition sequence (NARS). (b) Extending the first and the second ODNPs to produce a fragment that encompasses the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragment(s) in the presence of a DNA polymease and a nicking agent (NA) that recognizes the NARS. (d) Determining the allelic frequency of the target nucleic acid molecules that have the genetic variation, relative to the total of the target nucleic acid molecules present in the sample. This determination may be made by determining the percentage of the single-stranded nucleic acid fragment(s) formed in step (c) that contain the genetic variation optionally

present in the target nucleic acid (Product A) relative to all the single-stranded nucleic acid fragment(s) of step (c) that are identical to Product A (other than the genetic variation being examined) to thereby determine the allelic frequency of the target nucleic acid. In a related embodiment, the present invention provides a method for determining the allelic frequency of a target nucleic acid molecule having a genetic variation at a defined location, where the target nucleic acid is contained with a nucleic acid population. This method includes: (a) Admixing a first oligonucleotide primer (ODNP), a second ODNP and the population of nucleic acid molecules, some of which are the target nucleic acid, wherein if (i) the target nucleic acid is double-stranded, having a first strand and a second strand, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first strand of the target nucleic acid, at a location 3' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the second strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location, or, if (ii) the target nucleic acids are single-stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleotide sequence present in the target nucleic acid, where this nucleotide sequence is located 5' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid and located 3' to the defined location in the target nucleic acid. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking endonuclease recognition sequence (NERS). (b) Extending the first and the second ODNPs to produce a fragment that encompasses the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragment(s), preferably under isothermal conditions, in the presence of a DNA polymease and a nicking endonuclease (NE) that recognizes the NERS. In one embodiment, the single-stranded nucleic acid fragment(s) each contain no more than 20, or 19, or 18, or 17, or 16, or 15, or 14, or 13, or 12, or 11 or 10, etc. nucleotides. (d) Determining the allelic frequency of the genetic variation in the target nucleic acid of interest. This determination may be made by, *e.g.*, determining the percentage of the single-stranded nucleic acid fragment(s) of step (c) that contain the genetic variation optionally present in the target nucleic acid (Product A) relative to all the single-stranded nucleic acid fragment(s) of step (c) that are identical to the portions of Product A other than the genetic variation. Such a determination may be made using one or

both of liquid chromatography and mass spectrometry, where these techniques are exemplary only.

In related embodiments, the methods of the present invention directed to (i) determining the ratio of the number of a first nucleic acid molecules present in a population of nucleic acid molecules to the number of a second nucleic acid molecule present in a population of nucleic acid molecules, and (ii) determining the allelic frequency of a target nucleic acid molecule with a genetic variation in a population of nucleic acid molecules, may each be performed simultaneously on more than one nucleic acid molecule of interest, *i.e.,* the methods may be performed via multiplex determination. For example, in one aspect, the present invention provides a method for the multiplex determination of the allelic frequencies of target nucleic acid molecules each having a genetic variation at a defined location, where these target nucleic acid molecules are present in a nucleic acid population, where the method includes: (a) Admixing the nucleic acid population and ODNP pairs, where each pair corresponds to one of the target nucleic acids of interest and comprises a first ODNP and a second ODNP, wherein if (i) a target nucleic acid under investigation is double stranded such that it has a first strand and a second strand, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the second strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location in the target nucleic acid, or, if (ii) a target nucleic acid under investigation is single stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleotide sequence present in the target nucleic acid and located 5' to the defined location within the target nucleic acid, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the target nucleic acid and located 3' to the defined location within the target nucleic acid. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS). (b) Extending the first and the second ODNPs of each ODNP pair to produce a fragment that encompasses the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragments in the presence of a DNA polymease and a nicking agent (NA) that recognizes the NARS, using the extension product(s) of step (b) as the template(s). (d) Determining, for each target nucleic acid, the percentage of the single-

stranded nucleic acid fragments of step (c) that contain the genetic variation in the target nucleic acid (Product A) relative to all of the single-stranded nucleic acid fragments of step (c) that are identical to the portions of Product A other than the genetic variation, to thereby determine the allelic frequency of the target nucleic acid.

In a related aspect, the invention provides a method for the multiplex determination of the allelic frequencies of target nucleic acid molecules having genetic variations at defined locations, where these target nucleic acid molecules are present in a population of nucleic acids, and where this method includes: (a) Admixing the nucleic acid population and ODNP pairs, where each of the ODNP pairs corresponds to one of the target nucleic acids and comprises a first ODNP and a second ODNP, wherein if (i) a target nucleic acid under investigation is double stranded (such that it has a first strand and a second strand), then (a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the first strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the second strand of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location, or, if (ii) a target nucleic acid under investigation is single-stranded, then (a) the first ODNP comprises a nucleotide sequence that is at least substantially identical to a nucleotide sequence of the target nucleic acid that is located 5' to the defined location, and (b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the target nucleic acid, where this nucleotide sequence is located 3' to the defined location. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking endonuclease recognition sequence (NERS). (b) Extending the first and the second ODNPs of each ODNP pair to thereby produce a fragment that encompasses the first and the second ODNPs. (c) Amplifying single-stranded nucleic acid fragments, preferably under isothermal conditions, in the presence of a DNA polymease and a nicking endonuclease (NE) that recognizes the NERS, using the extension product(s) of step (b) as the template(s). Optionally, the single-stranded nucleic acid fragments each contain no more than 20, or 19, or 18, or 17, or 16, or 15, or 14, or 13, or 12, or 11, or 10, etc. nucleotides. (d) Determining, for each target nucleic acid, the percentage of the single-stranded nucleic acid fragments of step (c) that contain the genetic variation optionally present in the target nucleic acid (Product A) relative to all the single-stranded nucleic acid fragments of step (c) that are identical

to the portions of Product A other than the genetic variation to thereby determine the allelic frequency of the target nucleic acid.

In these methods of the present invention directed to (i) determining the ratio of the number of a first nucleic acid molecules present in a population of nucleic acid molecules to the number of a second nucleic acid-molecule present in a population of nucleic acid molecules, and (ii) determining the allelic frequency of a target nucleic acid molecule with a genetic variation in a population of nucleic acid molecules, one or more of the following criteria may be utilized to describe the method, where any two, three, four, five, etc. of the following criteria may be combined to describe the method, and where any one, two, three, four, etc. of the following criteria may be combined with any other criteria set forth herein, in describing the inventive method: the single-stranded nucleic acid fragment that is formed by the amplification step and contains the information needed to make the allelic frequency measurement or to determine the ratio as described above, has at least 1, or at least 2, or at least 3, or at least 4, or at least 5, or at least 6, or at least 7, or at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20; the single-stranded nucleic acid fragment that is formed by the amplification step has no more than 3, or 4, or 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or 20, or 21, or 22, or 23, or 24, or 25, or 26, or 27, or 28, or 29, or 30, or 31, or 32, or 33, or 34, or 35, or 36, or 37, or 38, or 39, or 40, or 41, or 42, or 43, or 44, or 45, or 46, or 47, or 48, or 49, or 50 nucleotides; the NARS utilized in the methods is a restriction endonuclease recognition sequence (RERS); the extension reaction is performed in the presence of at least one modified deoxyribonucleoside triphosphate; the NARS utilized in the methods is a nicking endonuclease recognition sequence (NERS); both the first ODNP and the second ODNP comprise the sequence of the sense strand of the NARS; the NARS is a NERS recognizable by N.BstNB I; the first ODNP comprises the nucleotide sequence of one strand of a RERS, but does not comprise the nucleotide sequence of the sense strand of a NERS; the second ODNP comprises the nucleotide sequence of one strand of a RERS, but does not comprise the nucleotide sequence of the sense strand of a NERS; fragments produced by an extension step do not contain a hemimodified RERS; the RERS is recognizable by a type IIs restriction endonuclease; the fragments of step (b) are digested by a RE that recognizes the RERS before performing step (c); the nucleotide(s) at the defined position is associated with a disease; the population of nucleic acids is taken from a mammal, *e.g.,* a human, or a dog, or a cat, or a pig, or a horse, or a rabbit., the DNA polymerase is any one of exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$

Bst, exo⁻ Pfu, exo⁻ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm™ DNA polymerase, and T4 DNA polymerase, where any two of more of the cited DNA polymerases may be combined to form a group that sets forth the possible DNA polymerases used in the present invention, *e.g.,* the DNA polymerase is exo⁻ Bst polymerase, exo⁻ Bca polymerase, 9°Nm™ DNA polymerase, or exo⁻ Vent polymerase; extension reaction(s) are performed by a polymerase chain reaction; amplification is performed under isothermal conditions; amplification is performed within the temperature range of 50°C-70°C; amplification is performed at 60°C; amplification is performed at 55 ± 5°C; amplification is performed at 60 ± 5°C; amplification is performed at 65 ± 5°C; amplification is performed at 70 ± 5°C; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 20°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 15°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 10°C of the lowest temperature; amplification is performed at temperatures between a highest temperature and a lowest temperature, where the highest temperature is within 5°C of the lowest temperature; in various embodiments of the invention each ODNP used in the method is independently at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides in length; each ODNP used in the method is, independently and optionally, no more than 50, or no more than 45, or no more than 40, or no more than 35, or no more than 30, or no more than 29, or no more than 28, or no more than 27, or no more than 26, or no more than 25, or no more than 24, or no more than 23, or no more than 22, or no more than 21, or no more than 20, or no more than 19, or no more than 18 nucleotides in length; when the method includes using primers that include a nucleotide sequence at least substantially complementary to a nucleotide sequence in a target nucleic acid, then in various embodiments of the invention this nucleotide sequence in the primer is, independently in each primer, at least 4, at least 5, at least 6, at least 7, at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides in length; when the method of the invention includes using primers that include a nucleotide sequence at least substantially complementary to a nucleotide sequence in a target

nucleic acid, then in various embodiments of the invention this nucleotide sequence in the primer is, independently in each primer, no more than 20, or no more than 19, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10, or no more than 9, or no more than 8, or no more than 7, or no more than 6, or no more than 5 nucleotides in length; the measurement performed on the amplification product in order to obtain the desired information includes one, or any two, or any three, or more of the following analytical methods, where additional analytical methods may also be utilized depending on the particular circumstances, *e.g.,* instrument availability: luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, gel electrophoresis, and capillary electrophoresis, *e.g.,* the measurement may be performed at least partially by mass spectrometry, or the measurement may be performed at least partially by liquid chromatography, the measurement may be performed by analytical methods that include both mass spectrometry and liquid chromatography.

In another aspect, the present invention provides a method for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, where this method includes: (a) Admixing a first ODNP, a second ODNP, and the cDNA. The first ODNP includes a nucleotide sequence that is at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand. The second ODNP includes a nucleotide sequence that is at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand. Furthermore, at least one of the first ODNP and the second ODNP further includes the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS). (b) Performing an extension reaction under reaction conditions that produce a fragment encompassed by the first ODNP and the second ODNP if both Exon A and Exon B are present in the cDNA. (c) Performing an amplification reaction in the presence of a DNA polymerase and a nicking agent (NA) that recognizes the NARS. (d) Characterizing an amplification product of step (c), to thereby determine the presence, or the absence, of the junction between Exon A and Exon B. In a related aspect, the present invention provides a method for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, where this method includes: (a) Admixing a first ODNP, a second ODNP, and the cDNA. The first ODNP includes a nucleotide sequence that is at least substantially complementary to a portion of the

antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand. The second ODNP has a nucleotide sequence that includes a sequence that is at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand. Furthermore, at least one of the first ODNP and the second ODNP further includes a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS). (b) Performing an extension reaction under reaction conditions that produce a fragment encompassed by the first ODNP and the second ODNP if both Exon A and Exon B are present in the cDNA. (c) Performing an amplification reaction, preferably under isothermal conditions, in the presence of a DNA polymerase and a nicking endonuclease (NE) that recognizes the NERS. (d) Characterizing an amplification product from step (c) to thereby determine the presence, or the absence, of the junction between Exon A and Exon B. The characterization may be performed using at least partially liquid chromatography and/or mass spectrometry. In optional embodiments of these methods, one or more of the following criteria may be used to describe the method: the NARS is a restriction endonuclease recognition sequence (RERS) and the extension reaction is performed in the presence of at lease one modified deoxyribonucleoside triphosphate; the NARS is a nicking endonuclease recognition sequence (NERS); both the first ODNP and the second ODNP comprise the nucleotide sequence of the sense strand of a NARS; the NARS is a RERS and extension is performed in the presence of at least one modified deoxyribonucleoside triphosphate; the NARS is a NERS; the NARS is recognizable by N.BstNB I; the first ODNP does not contain the nucleotide sequence of the sense strand of a NERS, but does include a nucleotide sequence of one strand of a RERS; the second ODNP does not contain the nucleotide sequence of the sense strand of a NERS, but does include a nucleotide sequence of one strand of a RERS; the fragment produced by the extension reaction does not contain a hemimodified RERS; the RERS is recognizable by a type IIs restriction endonuclease (RE); the extension product of step (b) is digested by a RE that recognizes the RERS before performing step (c); the gene is associated with a disease; the gene is associated with a human disease; the DNA polymerase is exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm^{TM} DNA polymerase, or T4 DNA polymerase, where any two or more of these DNA polymerases may be grouped together to form a group from which the DNA polymerase used in the method of the invention is selected, *e.g.,* the DNA polymerase is exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm^{TM} DNA polymerase, or

exo⁻ Vent polymerase; extension is performed by a polymerase chain reaction; amplification is performed under isothermal conditions; the amplification product, if present, is a single-stranded nucleic acid fragment containing no more than 35, or no more than 17, or no more than 12 nucleotides; characterizing is accomplished, at least in part, by an analytical technique selected from luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, gel electrophoresis, and capillary electrophoresis, *e.g.,* the characterization may be performed at least partially by mass spectrometry.

In another aspect, the present invention provides a method for detecting, in a cDNA molecule or a cDNA population, the presence, or the absence, of a junction between every two exons of a gene that comprises n exons, wherein n is an integer equal or more than 2, Exon A is the most upstream exon, and Exon N is the most downstream exon of the exons present in the gene, where this method includes: (a) Admixing a first set of oligonucleotide primers (ODNPs), a second set of ODNPs, and the cDNA molecule or the cDNA population. The first set of ODNPs comprises n-1 ODNPs, each member of the set corresponding to one of the n exons of the gene, with the exception of Exon A, and furthermore each of the ODNPs comprises a nucleotide sequence that is at least substantially complementary to a portion of the sense strand of the corresponding exon near the 5' terminus of said exon in said strand. The second set of ODNPs comprises n-1 ODNPs corresponding to each of the exons except Exon N, where each of the ODNPs comprises a nucleotide sequence that is at least substantially complementary to a portion of the antisense strand of the corresponding exon near the 5' terminus of said exon in said strand. Each of the ODNPs that is, in a first embodiment of the invention, a member of the first set of ODNPs, or, in a second embodiment of the invention, is a member of the second set of ODNPs, or, in a third embodiment of the invention, is a member of either the first or the second set of ODNPs, further comprises the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS). (b) Amplifying fragments that encompass one ODNP from the first ODNP set and another ODNP from the second ODNP set. (c) Performing an amplification reaction in the presence of a DNA polymerase and a nicking agent (NA) that recognizes the NARS. (d) Detecting the presence, or the absence, of amplification product(s) from step (c) to thereby determine the presence, or the absence, of each potential junction.

In a related aspect, the present invention provides a method for detecting, in a cDNA molecule or a cDNA population, the presence, or the absence, of a junction between every two exons of a gene that comprises n exons, wherein n is an

integer equal or more than 2, Exon A is the most upstream exon, and Exon N is the most downstream exon of the exons in the gene, where this method includes: (a) Admixing a first set of oligonucleotide primers (ODNPs), a second set of ODNPs, and the cDNA molecule or the cDNA population. The first set of ODNPs includes n-1 ODNPs, where each member of this set corresponds to a unique one of the exons except Exon A, and each of the ODNPs also comprises a nucleotide sequence that is at least substantially complementary to a portion of the sense strand of the corresponding exon near the 5' terminus of said exon in said strand. The second set of ODNPs includes n-1 ODNPs, where each member of this set corresponds to a unique one of the exons except Exon N, and each of the ODNPs further comprises a nucleotide sequence that is at least substantially complementary to a portion of the antisense strand of the corresponding exon near the 5' terminus of said exon in said strand. Furthermore, Each of the ODNPs that is, in a first embodiment of the invention, a member of the first set of ODNPs, or, in a second embodiment of the invention, is a member of the second set of ODNPs, or, in a third embodiment of the invention, is a member of either the first or the second set of ODNPs, further comprises the nucleotide sequence of the sense strand of a nicking endonuclease recognition sequence (NERS). (b) Amplifying fragments encompassed by one ODNP from the first ODNP set and another ODNP from the second ODNP set. (c) Performing an amplification reaction, preferably under isothermal conditions, in the presence of a DNA polymerase and a nicking endonuclease (NE) that recognizes the NERS. (d) Detecting the presence, or the absence, of amplification product(s) of step (c). In one embodiment, this detection is done by, liquid chromatography and/or mass spectrometry. The detection allows for the subsequent determination of the presence, or the absence, of each potential junction under investigation. In one embodiment, the amplification product(s) of step (c), if present, contain no more than 17 nucleotides. In various embodiments of the present invention, one or more of the following criteria may be used in describing the embodiment: the NARS is a restriction endonuclease recognition sequence (RERS) and step (b) is performed in the presence of at least one modified deoxyribonucleoside triphosphate; the NARS is a nicking endonuclease recognition sequence (NERS); both the first set of ODNPs and the second set of ODNPs comprise the nucleotide sequence of the sense strand of a NARS; the NARS is a RERS and step (b) is performed in the presence of at least one modified deoxyribonucleoside triphosphate; the NARS is a NERS; the NERS is recognizable by N.BstNB I; the ODNPs in the first set or the ODNPs in the second set comprise a sequence of one strand of a RERS, but do not comprise a sequence of a sense strand of a NERS; the fragments produced by step (b) do not contain a hemimodified RERS; the

RERS is recognizable by a type IIs restriction endonuclease (RE); the amplified fragments of step (b) are digested by a RE that recognizes the RERS before performing step (c); the gene is associated with a disease; the gene is associated with a human disease; step (b) is performed by a polymerase chain reaction; step (c) is performed under an isothermal condition; the amplification product of step (c), if present, is a single-stranded nucleic acid fragment containing no more than 35 nucleotides; the single-stranded nucleic acid fragment contains no more than 17 nucleotides; the single-stranded nucleic acid fragment contains no more than 12 nucleotides; step (d) is performed at least partially by the technique selected from the group consisting of luminescence spectroscopy, fluorescence spectroscopy, mass spectrometry, liquid chromatography, fluorescence polarization, gel electrophoresis, and capillary electrophoresis; step (d) is performed at least partially by mass spectrometry.

The present invention, in other aspects, provides methods for detecting alternative splicing of a gene in a cDNA population. In one embodiment, the method includes: (a) Determining the presence, or the absence, of each potential junction between any two exons of the gene according to a method of the present invention described herein. (b) Indicating the presence of alternative splicing of the gene in the cDNA population, if more than one junctions are present for at least one exon of the gene at at least one terminus of the exon. In a related embodiment, the method provides for the multiplex detection of alternative splicing of genes in a cDNA population, where this method include, for each gene: (a) Determining the presence, or the absence, of each potential junction between any two exons of the gene according to the method according to a method of the present invention as described herein. (b) Indicating the presence of alternative splicing of the gene in the cDNA population, if more than one junctions are present for at least one exon of the gene at at least one terminus of the exon. In a related embodiment, the present invention provides a method for detecting alternative splicing of a gene between two biological samples, where the method includes: (a) Determining, for each biological sample, the presence, or the absence, of each potential junction between any two exons of the gene by performing a method according to the present invention. (b) Indicating the presence of alternative splicing of the gene between the two biological samples, if at least one junction is present in one of the biological samples, but not in the other biological sample. In a related embodiment, the present invention provides a method for the multiplex detection of alternative splicing of genes between two biological samples, which includes, for each gene: (a) Determining, for each biological sample, the presence, or the absence, of each potential junction between any two exons of the gene according to a method of the

present invention as described herein. (b) Indicating the presence of alternative splicing of the gene between the two biological samples, if at least one junction is present in one of the biological samples, but not in the other biological sample. In a related embodiment, the present invention provides an oligonucleotide primer (ODNP) pair for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, where: (a) The first ODNP comprises a nucleotide sequence that is at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand. The second ODNP comprises a nucleotide sequence that is at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand. Furthermore, at least one of the first ODNP and the second ODNP further comprises the nucleotide sequence of the sense strand of a nicking agent recognition sequence (NARS). In various embodiments of this ODNP pair of the invention, the following criteria may be used to described the pair, where any two or more of the following criteria may be combined, optionally with other criteria described elsewhere herein: the NARS is a restriction endonuclease recognition sequence (RERS); the NARS is a nicking endonuclease recognition sequence (NERS); both the first ODNP and the second ODNP comprise the nucleotide sequence of the sense strand of a NARS; the NARS is a RERS; the NARS is a NERS; the NERS is recognizable by N.BstNB I; the first ODNP comprises the nucleotide sequence of a sense strand of a NERS, and the second ODNP comprises the nucleotide sequence of one strand (the sense strand, in one embodiment, the antisense strand in another embodiment) of a RERS; the first ODNP comprises the nucleotide sequence of one strand of a RERS, and the second ODNP comprises the nucleotide sequence of a sense strand of a NERS; the RERS is recognizable by a type IIs restriction endonuclease (RE); the gene is associated with a disease; the first ODNP is at least 12 nucleotides in length; the second ODNP is at least 12 nucleotides in length.

In another aspect, the present invention provides a method of making a template nucleic acid molecule that is useful for synthesizing a single-stranded cDNA molecule, where the method includes: (a) Providing a double-stranded cDNA molecule. (b) Ligating the double-stranded cDNA molecule with a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS) so that the single-stranded nucleic acid molecule amplified in the presence of a nicking agent (NA) that recognizes the NARS and a DNA polymerase using the ligation product as a template comprises at least a portion of one strand of the double-stranded cDNA molecule. In another aspect, the present invention provides a method for making a cDNA library, where this method

includes: (a) Providing mRNA molecules isolated from a biological sample. (b) Preparing double-stranded cDNA molecules using the isolated mRNA molecules as templates. (c) Ligating the double-stranded cDNA molecules to a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS) so that the single-stranded nucleic acid molecules that will be amplified in the presence of (i) a nicking agent that recognizes the NARS and (ii) a DNA polymerase, using the ligation products as templates, comprise at least a portion of one strand of the double-stranded cDNA molecules. In optional embodiments, these methods of the present invention may be further characterized by one or more of the following criteria, where any two or more of these criteria may be combined in characterizing this method of the invention, and where these criteria are non-exclusive: said portion being at least 40, or at least 38, or at least 36, or at least 34, or at least 32, or at least 30, or at least 28, or at least 26, or at least 24, or at least 22, or at least 20, or at least 19, or at least 18, or at least 17, or at least 16, or at least 15, or at least 14, or at least 13, or at least 12, or at least 11, or at least 10, or at least 9, or at least 8, or at least 7, or at least 6, or at least 5, or at least 4 nucleotides in length; said portion being no more than 100, or nor more than 90, or no more than 80, or no more than 70, or no more than 60, or no more than 50, or no more than 40, or no more than 38, or no more than 36, or no more than 34, or no more than 32, or no more than 30, or no more than 28, or no more than 26, or no more than 24, or no more than 22, or no more than 20, or no more than 18, or no more than 17, or no more than 16, or no more than 15, or no more than 14, or no more than 13, or no more than 12, or no more than 11, or no more than 10, or no more than 9, or no more than 8, or no more than 7, or no more than 6, or no more than 5 nucleotides in length; the double-stranded cDNA molecule is immobilized via the terminus that is not involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor; the nucleic acid adaptor is immobilized via the terminus that is not involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor; the NA is a nicking endonuclease (NE); the NA is N.BstNB I; the NARS is a hemimodified restriction endonuclease recognition sequence; when an adaptor is involved, then either the double-stranded cDNA molecules or the nucleic acid adaptor is immobilized via the terminus of the immobilized material that is not involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor.

In another aspect, the present invention provides a method that includes: (a) Providing a double-stranded nucleic acid molecule that includes (i) a type IIs restriction endonuclease recognition sequence (TRERS) as an optionally present sequence, (ii) a nicking agent recognition sequence (NARS), and (iii) a fragment of a

target nucleic acid. In this double-stranded nucleic acid, and in particular in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, (1) the target nucleic acid fragment is located 5' to the NARS, and (2) if the TRERS is present, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS. (b) As an optional step, cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS. (c) Amplifying a single-stranded nucleic acid fragment in the presence of a NA that recognizes the NARS. In a related aspect, the present invention provides a method that includes: (a) Providing a double-stranded nucleic acid molecule that includes: (i) a type IIs restriction endonuclease recognition sequence (TRERS), (ii) a nicking endonuclease recognition sequence (NERS), and (iii) a target nucleic acid fragment. In this double-stranded nucleic acid molecule, and in particular in the strand that does not contain the nicking site (NS) of a nicking endonuclease (NE) that recognizes the NERS, (1) the target nucleic acid fragment is located 5' to the NERS, and (2) the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS. (b) Cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS. (c) Amplifying a single-stranded nucleic acid fragment, optionally under isothermal conditions, in the presence of a NA that recognizes the NARS. In one embodiment the single-stranded nucleic acid fragment contains no more than 17 nucleotides. In another aspect, the present invention provides a method for making a plurality of single-stranded nucleic acid probes, where this method includes: (a) Providing double-stranded nucleic acid molecules, where each double-stranded nucleic acid molecule includes: (i) a type IIs restriction endonuclease recognition sequence (TRERS), where this is an optional component of the double-stranded nucleic acid molecules; (ii) a nicking agent recognition sequence (NARS), and (iii) a target nucleic acid fragment. Furthermore, in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, (1) the target nucleic acid fragment is located 5' to the NARS, and (2) if the TRERS is present, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS. (b) the method includes the optional step of cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS. (c) Amplifying single-stranded nucleic acid fragments in the presence of a NA that recognizes the NARS. In a related aspect, the present invention provides a method for

making a plurality of single-stranded nucleic acid probes, where the method includes: (a) Providing a plurality of double-stranded nucleic acid molecules, where each double-stranded nucleic acid molecule includes: (i) a type IIs restriction endonuclease recognition sequence (TRERS), (ii) a nicking endonuclease recognition sequence (NERS), and (iii) a target nucleic acid fragment. In these double-stranded nucleic acid molecules, and in particular in the strand that does not contain the nicking site (NS) of a nicking endonuclease (NE) that recognizes the NERS, (1) the target nucleic acid fragment is located 5' to the NERS, and (2) the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS. (b) Cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS. (c) Amplifying single-stranded nucleic acid fragments, preferably under isothermal conditions, in the presence of a NE that recognizes the NERS. Optionally, the single-stranded nucleic acid fragments each contain no more than 25, or no more than 17, nucleotides. These methods of the present invention may be further described by one or more of the following criteria, where any two or more of the following criteria may be combined in describing these methods of the present invention: the NA is a nicking endonuclease (NE); the NA is is N.BstNB I; the NARS is a hemimodified restriction endonuclease recognition sequence; the NARS is recognizable by a restriction endonuclease (RE) selected from the group consisting of Ava I, Bsl I, BsmA I, BsoB I, Bsr I, BstN I, BstO I, Fnu4H I, Hinc II, Hind II, and Nci I; the TRERS is recognizable by Bpm I or Mme I; providing the double-stranded nucleic acid molecule is performed, at least in part, by linking an adaptor molecule as described herein, and incorporating a NARS, to a double-stranded target nucleic acid fragment, where optionally the double-stranded target nucleic acid fragment is cleaved by a restriction endonuclease (RE) before being linked to the adaptor, and where also optionally the double-stranded target nucleic acid fragment is a cDNA fragment, where the cDNA fragment may optionally be immobilized to a solid support, and where the immobilized cDNA may optionally be produced by a method comprising: (i) isolating mRNA from a biological sample using an immobilized oligonucleotide primer; and (ii) synthesizing a double-stranded cDNA using the mRNA as a template; amplifying is performed in the presence of a DNA polymerase, where the DNA polymerase is one of, or any group consisting of two or more of, exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm^{TM} DNA polymerase, and T4 DNA

polymerase, *e.g.*, the polymerase is exo⁻ Bst polymerase, exo⁻ Bca polymerase, 9°Nm™ DNA polymerase, or exo⁻ Vent polymerase; amplification is performed under isothermal conditions; the single-stranded nucleic acid fragment that is amplified contains no more than 25 nucleotides; the single-stranded nucleic acid fragment is 15-50 nucleotides in length; the single-stranded nucleic acid fragment is 51-20,000 nucleotides in length.

In another aspect, the present invention provides a method for amplifying a single-stranded nucleic acid molecule, where this method includes: (A) Forming a mixture that includes: (i) a target nucleic acid; and (ii) an oligonucleotide primer that (a) includes the nucleotide sequence of the sense strand of a double-stranded nicking agent recognition sequence (NARS) recognizable by a nicking agent (NA) that nicks at a nicking site (NS) located outside the recognition sequence, and (b) includes a nucleotide sequence that is at least substantially complementary to the target nucleic acid. (B) Amplifying a single-stranded nucleic acid molecule using a portion of the target nucleic acid as a template in the presence of the nicking agent. Optionally, the double-stranded nicking agent recognition sequence is recognizable by N.BstNB I. Optionally, one nucleotide within the nucleotide sequence of the sense strand of the double-stranded nicking agent recognition sequence does not form a conventional base pair with another nucleotide of the target nucleic acid when the oligonucleotide primer anneals to the target nucleic acid. Optionally, two or more nucleotides within the nucleotide sequence of the sense strand of the double-stranded nicking agent recognition sequence do not form conventional base pairs with nucleotides of the target nucleic acid when the oligonucleotide primer anneals to the target nucleic acid. Optionally, the target nucleic acid is immobilized. Optionally, the oligonucleotide primer is immobilized at its 5' terminus.

In another aspect, the present invention provides a method that includes: (A) Forming a mixture that includes: (i) a target nucleic acid, (ii) an oligonucleotide primer that (a) includes the nucleotide sequence of the sense strand of a NARS recognizable by a nicking agent that nicks outside the recognition sequence, and (b) is at least substantially complementary to a first region of the target nucleic acid; and (iii) a partially double-stranded nucleic acid that (a) comprises a double-stranded type IIs restriction endonucelase recognition sequence, and (b) a 3' overhang that is at least substantially complementary to a second region of the target nucleic acid located 5' to the first region. These components are combined under conditions that allow for hybridization between the oligonucleotide primer and the first region of the target and between the 3' overhang of the partially double-stranded nucleic acid and the second region of the target. (B) Digesting the target nucleic acid hybridized to the

oligonucleotide primer and the partially double-stranded nucleic acid in the second region. (C) Amplifying a single-stranded nucleic acid molecule using a portion of the target nucleic acid digested in step (B) as a template in the presence of the nicking agent. In a related embodiment, the present invention provides a method that includes: (A) Forming a mixture comprising: (i) a target nucleic acid; (ii) an oligonucleotide primer that (a) comprises the nucleotide sequence of the sense strand of a NARS recognizable by a nicking agent that nicks outside the recognition sequence, and (b) is at least substantially complementary to a first region of the target nucleic acid; and (iii) a single-stranded nucleic acid molecule that (1) is at least substantially complementary to a second region of the target nucleic acid located 5' to the first region. The components are combined under conditions that allow for hybridization between the oligonucleotide primer and the first region of the target, and between the single-stranded nucleic acid molecule and the second region of the target. (B) Amplifying a single-stranded nucleic acid molecule using a portion of the target nucleic acid as a template in the presence of the nicking agent and DNA polymerase. Preferably, the DNA polymerase does not have strand displacement activity so that in effect, the single-stranded nucleic acid molecule blocks further extension reaction by the DNA polymerase. According to these aspects of the present invention, multiple single stranded nucleic acid molecules are produced, preferably under isothermal conditions, and preferably of a length that is less than the full length that could be produced based on the length of the target nucleic acid molecule. In one embodiment, there is one mismatched base pair in the NARS, while in another embodiment there are two mismatched base pairs in the NARS, while in another embodiment all of the base pairs that form the NARS are mismatched, while in another embodiment, n-1 of the base pairs that form the NARS are mismatched, where n base pairs form the NARS. In another embodiment, there is an unmatched nucleotide in the NARS. In another embodiment, all of the nucleotides that form the sense sequence of the NARS are unmatched.

In another aspect, the present invention provides a method that includes: (a) Forming a mixture of (i) a partially or fully double-stranded template nucleic acid molecule that comprises a nicking agent recognition sequence (NARS), (ii) a nicking agent (NA) that recognizes the NARS, (iii) a DNA polymerase, and (iii) one or more deoxyribonucleoside triphosphates. (b) Maintaining the mixture of step (a) under conditions that allows for the amplification of a single-stranded nucleic acid fragment, where under the conditions the single-stranded nucleic acid fragment is capable of spontaneously dissociating from the template nucleic acid in the absence of any strand

displacement activity of the DNA polymerase or a strand displacement facilitator. In a related aspect, the present invention provides a method that includes: (a) Contacting a partially or fully double-stranded template nucleic acid with a nicking agent (NA), where the template comprises a nicking agent recognition sequence (NARS). (b) If the template does not comprise a nicking site (NS) nickable by the NA, then extending the template to provide the NS. (c) Nicking the template or the extension product thereof at the NS to provide a nicked product that comprises a new 3' terminus at the NS and a nicked product that comprises a new 5' terminus at the NS; wherein the melting temperature of the duplex formed between the nicked product comprising the new 5' terminus at the NS and the strand of the template or the extension product thereof that comprises the sequence of the antisense strand of the NARS is higher than that of the duplex formed between the nicked product comprising the new 3' terminus at the NS and the same strand of the template or the extension product thereof. (d) Extending the nicked product of step (c) that comprises the new 3' terminus at the NS. (e) Repeating steps (c) and (d) to amplify a single-stranded nucleic acid fragment. Optionally, the nicking, extension and/or amplification are preformed under isothermal conditions. Optionally, the temperature at which steps (b), (c), (d) and (e) is performed is between the melting temperature of the duplex formed between the nicked product comprising the new 5' terminus at the NS and the strand of the template or the extension product thereof that comprises the sequence of the antisense strand of the NARS and the melting temperature of the duplex formed between the nicked product comprising the new 3' terminus at the NS and the same strand of the template or the extension product thereof.

Whenever a single-stranded nucleic acid fragment is formed by an amplification step in any method of the present invention, and particularly (but not necessarily) when this single-stranded nucleic acid fragment contains information of interest, *e.g.*, has a particular nucleotide at a particular location, or has a particular nucleotide sequence, where this information allows a determination to be made, then in various embodiments of the methods of the invention, the single-stranded nucleic acid fragment has at least 1, or at least 2, or at least 3, or at least 4, or at least 5, or at least 6, or at least 7, or at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides. In addition, the single-stranded nucleic acid fragment that is formed by an amplification step of the present invention may be characterized as containing no more than 3, or no more than 4, or no more than 5, or no more than 6, or no more than 7, or no more than 8, or no more than 9, or no more than

10, or no more than 11, or no more than 12, or no more than 13, or no more than 14, or no more than 15, or no more than 16, or no more than 17, or no more than 18, or no more than 19, or no more than 20, or no more than 21, or no more than 22, or no more than 23, or no more than 24, or no more than 25, or no more than 26, or no more than 27, or no more than 28, or no more than 29, or no more than 30, or no more than 31, or no more than 32, or no more than 33, or no more than 34, or no more than 35, or no more than 36, or no more than 37, or no more than 38, or no more than 39, or no more than 40, or no more than 41, or no more than 42, or no more than 43, or no more than 44, or no more than 45, or no more than 46, or no more than 47, or no more than 48, or no more than 49, or no more than 50 nucleotides.

When a digestion product is formed in any method of the present invention, then in various embodiments the digestion product has at least 8, or at least 10, or at least 12, or at least 14, or at least 16, or at least 18, or at least 20, or at least 22, or at least 24, or at least 26, or at least 28, or at least 30, or at least 32, or at least 34, or at least 36, or at least 38, or at least 40, or at least 42, or at least 44, or at least 46, or at least 48, or at least 50, or at least 52, or at least 54, or at least 56, or at least 58, or at least 60, or at least 62, or at least 64, or at least 66, or at least 68, or at least 70, or at least 72, or at least 74, or at least 76, or at least 78, or at least 80, or at least 82, or at least 84, or at least 86, or at least 88, or at least 90, or at least 92, or at least 94, or at least 96, or at least 98, or at least 100 nucleotides; when a digestion product is formed, in various embodiments the digestion product has no more than 200, or no more than 180, or no more than 160, or no more than 140, or no more than 120, or no more than 110, or no more than 100, or no more than 90, or no more than 80, or no more than 70, or no more than 60, or no more than 50, or no more than 40, or no more than 38, or no more than 36, or no more than 34, or no more than 32, or no more than 30, or no more than 28, or no more than 26, or no more than 24, or no more than 22, or no more than 20, or no more than 18, or no more than 16, or no more than 14, or nor more than 12, or no more than 10 nucleotides.

In any of the methods or compounds or compositions of the present invention that include a NARS, the NARS may contain a, *i.e.*, one or more, mismatched nucleotides. In other words, one or more of the nucleotide base pairs that form the NARS may not be hybridized according to the conventional Watson-Crick base pairing rules. However, when mismatched nucleotides are present in the NARS, then at least all of the nucleotides that form the sense strand of the NARS are present.

When any method of the present invention entails the use of a modified deoxyribonucleoside triphosphate then in one embodiment the modified deoxyribonucleoside triphosphate is an α-thio deoxyribonucleoside triphosphate.

When a DNA polyerase is utilized in any method of the invention, the DNA polymerase is any one or more of exo⁻ Vent, exo⁻ Deep Vent, exo⁻ Bst, exo⁻ Pfu, exo⁻ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm™ DNA polymerase, and T4 DNA polymerase, where a group of suitable DNA polymerases may be formed by combining any two, or any three, or any four, etc. of these listed polymerases, *e.g.,* the DNA polymerase may be selected from the group consisting of exo⁻ Bst polymerase, exo⁻ Bca polymerase, 9°Nm™ DNA polymerase, or exo⁻ Vent polymerase.

In addition to the methods, kits and compositions summarized above and described in greater detail herein, the present invention also provides the following compounds, compositions and kits that are useful in, or may be generated by, the methods and kits of the present invention:

A partially double-stranded oligonucleotide probe that includes (a) a nicking agent recognition sequence (NARS) that is recognizable by a nicking agent (NA), where the NA will nick either the probe, or an extension product thereof, at a nicking site (NS), (b) either (i) a 5' overhang in the strand that contains the NS wherein the overhang includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence in a target nucleic acid, or (ii) a 3' overhang in the strand that does not contain the NS wherein the overhang includes a nucleotide sequence that is at least substantially complementary to a nucleotide sequence in a target nucleic acid; and (c) a nucleotide sequence within that strand of the probe that neither contains the NS nor may be extended so as to contain the NS, where this nucleotide sequence is located 5' to the position corresponding to the NS, and this nucleotide sequence is uniquely correlated to the target nucleic acid to which the overhang is at least substantially complementary. This probe is particularly useful in a method wherein, sequentially, a test sample questionably containing the target nucleic acid is combined with the probe; probes that have not hybridized to a nucleic acid from the sample are separated from probes that have hybridized to a nucleic acid from the sample; probes that have hybridized to a nucleic acid from the sample are then treated with nicking agent and polymerase so as to amplify a single-stranded nucleic acid fragment that contains a nucleotide sequence that uniquely correlates to the nucleic acid from the sample; and the amplified nucleic acid fragments are characterized so as to determine their

nucleotide sequence at least insofar as that sequence is the sequence that uniquely correlates to the presence of the nucleic acid from the sample. For instance, the target or probe may be immobilized, then non-immobilized target (when the probe is immobilized) or non-immobilized probe (when the target is immobilized) is combined with the immobilized material, then anything that is not immobilized, or is not hybridized to something that is immobilized, is washed away, then the immobilized material is subjected to repeated nicking and extension according to the present invention to provide an amplified single-stranded nucleic acid fragment, then the amplified single-stranded fragment is characterized to at least identify the complement of the nucleotide sequence present in the probe that uniquely correlates to the target, then a determination is made regarding whether the target of interest, *i.e.,* the target to which the probe was designed to "uniquely correlate" was present in the sample. In essence, the nucleotide sequence in the probe that unique correlates to the target nucleic acid is effectively a signal that is specific for the particular target nucleic acid. In one aspect, the present invention provides a compound wherein a target nucleic acid is hybridized to the probe. In another aspect, the present invention provides a compound wherein the target nucleic acid is immobilized to a solid support and is also hybridized to the probe. In another aspect, the present invention provides that the probe is immobilized to a solid support. In another aspect, the present invention provides that the probe is immobilized to a solid support and the target nucleic acid is hybridized to the probe. In various additional embodiments of the present invention, one or more of the following criteria may be used in describing the probe and/or the hybridization product of the probe and the target: the target nucleic acid is one strand of a denatured double-stranded nucleic acid, where the double-stranded nucleic acid may be, for example, genomic nucleic acid and/or may be cDNA; the target nucleic acid is derived from the genome of a pathogenic virus; the target nucleic acid is derived from the genome of a pathogenic bacterium; the target nucleic acid is derived from the episome of a pathogenic bacterium; the double-stranded region of the probe is 10-22, or 11-21, or 12-20, or 13-19, or 8-25, or 8-30, or 8-35, or 10-25, or 10-30, or 10-35 nucleotides in length; the probe has a 3' overhang in the strand that does not contain the NS; the probe has a 5' overhang in the strand that contains the NS; the 5' overhang or the 3' overhang is 5-25, or 6-22, or 8-20, or 10-18 nucleotides in length; the probe contains less than 200 nucleotides; the probe contains less than 150 nucleotides; the probe contains less than 100 nucleotides; the probe contains less than 90, or less than 80, or less than 70, or less than 60, or less than 50 nucleotides; the probe contains at least 20, or at least 21, or at least 22, or at least 23, or at least 24, or at least 25, or at least 26, or at least 27, or at

least 28, or at least 29, or at least 30 nucleotides; the NS is nickable by a nicking endonuclease; the NS is nickable by N.BstNB I; the present invention also provides a composition that includes a plurality of unique probes as describe above, where a probe is unique, *i.e.,* un-like any other probe in the composition, by virtue of it having a sequence according to feature (c) of the probe, that uniquely correlates with a target nucleic acid in a manner unlike that which any other probe correlates with that particular target nucleic acid, or else the probe has a sequence that uniquely correlates to a target nucleic acid where no other probe in the composition also correlates with that particular target nucleic acid. In one aspect, this plurality of probes is organized in an array format, *e.g.,* the plurality is immobilized to a solid support at different, defined locations on the support.

A pair of oligonucleotide primers (ODNPs), or a kit or composition that includes a pair of ODNPs, the pair being defined by a first ODNP and a second ODNP, wherein (a) the first ODNP is formed, at least in part, by a nucleotide sequence that is at least substantially complementary to a nucleotide sequence located 3' to a defined position in a single-stranded target nucleic acid, (b) the second ODNP is formed, at least in part, by a nucleotide sequence that is at least substantially complementary to a nucleotide sequence present in the complement of the target nucleic acid, where this nucleotide sequence present in the complement is located at a position 3' to the complement of the defined position; (c) the first and the second ODNPs further include a first constant recognition sequence (CRS) of a first strand and a second CRS of a second strand of an interrupted restriction endonuclease recognition sequence (IRERS), respectively, but not a complete IRERS, where the complete IRERS is a double-stranded nucleic acid sequence formed from the first and second strands, where the first and the second CRSs are linked by a variable recognition sequence (VRS), and (d) either the first or the second ODNP further comprises a NERS located 5' to the first CRS or the second CRS. The following criteria, in addition to other criteria as may be set forth herein, may be further utilized to describe the ODNPs alone, or the ODNPs as present in a composition or kit, where any two or more of these criteria may be combined in describing the ODNPs: the nucleotide sequence of the first ODNP which is at least substantially complementary to the target nucleic acid is at least 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16 or 17, or 18, or 19, or 20 nucleotides in length; the nucleotide sequence of the second ODNP which is at least substantially complementary to the complement of the target nucleic acid is at least 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16 or 17, or 18, or 19, or 20 nucleotides in length; the first ODNP is 15-85, or 16-70, or 17-60, or 18-50, or 18-40, or 20-30, or 15-30, or 15-

40, or 15-50, nucleotides in length; the second ODNP is 15-85, or 16-70, or 17-60, or 18-50, or 18-40, or 20-30, or 15-30, or 15-40, or 15-50 nucleotides in length; the first ODNP comprises one or more nucleotides that together are at least substantially complementary to the nucleotides present in the target nucleic acid at the 3' terminus of the first CRS; the second ODNP comprises one or more nucleotides that together are at least substantially complementary to the nucleotides present in the target nucleic acid at the 3' terminus of the second CRS; the single-stranded target nucleic acid is one strand of a double-stranded nucleic acid; where the double-stranded nucleic acid may be, for example, genomic nucleic acid and/or may be cDNA; the target nucleic acid may be derived from the genome of a pathogenic virus; the target nucleic acid may be derived from the genome of a pathogenic bacterium; the target nucleic acid may be derived from the episome of a pathogenic bacterium. The composition may further comprise the target nucleic acid, and/or the complement of the target nucleic acid. The composition or kit may further include one or more of: a restriction endonuclease that recognizes the IRERS; Bsl I; a nicking endonuclease (NE) that recognizes the NERS; N.BstNB I; a DNA polymerase; appropriate buffers to run an assay with the ODNPs; and instructions for using the kit. For instance, the present invention provides a genotyping kit that includes the just-described primer pair, along with a NE that recognizes the NERS, a RE that recognizes the IRERS, a DNA polymerase, and an instruction for using the kit.

A nucleic acid adaptor comprising a nicking agent recognition sequence (NARS). The present invention provides the adaptor, kit(s) containing the adaptor, methods of using the adaptor, and compositions containing the adaptor. In various embodiments of the invention, the following criteria may be used to further characterize the adaptor, where any two or more of the criteria may be combined, where these criteria are exemplary only and supplement any criteria set forth elsewhere herein: it contains both the sense and antisense sequence of the NARS; it contains a double-stranded nicking site (NS) where one strand only is cleaved by a nicking agent that recognizes the NARS; it contains less than 200, or less than 175, or less than 150, or less than 125, or less than 100, or less than 90, or less than 80, or less than 70, or less than 65, or less than 60, or less than 55, or less than 50, or less than 45, or less than 40, or less than 35, or less than 30, or less than 25, or less than 20 nucleotides; it contains more than 10, or more than 15, or more than 20, or more than 25, or more than 30, or more than 35, or more than 40, or more than 45, or more than 50 nucleotides; it contains a 3' overhang for ligating with a target nucleic acid; it contains a 5' overhang for ligating with a target nucleic acid; it contains a blunt end for ligating with a target nucleic acid;

it is immobilized to a solid support, it is immobilized to a solid support through a 3' end of the adaptor; it is immobilized to a solid support through a 5' end of the adaptor; the NARS is recognized by a nicking endonuclease; the NARS is recognized by N.BstNB I; the NARS is a hemimodified restriction endonuclease recognition sequence; the NARS is recognizable by a restriction endonuclease (RE) selected from the group consisting of Ava I, Bsl I, BsmA I, BsoB I, Bsr I, BstN I, BstO I, Fnu4H I, Hinc II, Hind II, and Nci I; the adaptor is designed such that when the adaptor is linked to a double-stranded target nucleic acid fragment, then a strand of target nucleic acid is located 5' to that portion of the NARS located in the strand of the adaptor that does not contain the NS; it contains a type IIs restriction endonuclease recognition sequence (TRERS); it contains a type IIs restriction endonuclease recognition sequence (TRERS) and the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both (i) within the double-stranded target nucleic acid fragment, and (ii) 5' to the position corresponding to the NS that is cleaved by the NA; it contains a type IIs restriction endonuclease recognition sequence (TRERS) and the TRERS is located 5' to the NARS as measured within the strand containing the NS; it contains a type IIs restriction endonuclease recognition sequence (TRERS) and the TRERS is located 3' to the NARS as measured within the strand containing the NS; it contains a type IIs restriction endonuclease recognition sequence (TRERS) that is recognizable by Bpm I; it contains a type IIs restriction endonuclease recognition sequence (TRERS) that is recognizable by Mme I. The present invention also provides compositions and kits including the adaptor. The compositons and kits may be useful in, *e.g.,* synthesizing an oligonucleotide probe as described herein. In addition to the adaptor, optional components of the compositions and kits include one or more of: a nicking agent that recognizes the NARS; a nicking endonuclease (NE) that recognizes the NARS; a buffer for the NE; a Type IIs restriction endonuclease that recognizes the TRERS; a buffer for the Type IIs restriction endonuclease; a DNA polymerase (*e.g.,* a DNA polymerase selected from exo$^-$ Bst polymerase, exo$^-$ Bca polymerase, 9°Nm™ DNA polymerase, and/or exo$^-$ Vent polymerase); a buffer for the DNA polymerase; a ligase that can ligate the adaptor to a target nucleic acid; a buffer for the DNA ligase; instructions for using the kit; a primer immobilized to a solid support that comprises a poly (dT) oligonucleotide; a reverse transcriptase; a strand displacement facilitator, *e.g.,* trehalose. The present invention also provides a composition comprising the adaptor and a double-stranded target nucleic acid. Optionally, the double-stranded target nucleic acid fragment is a cDNA fragment. Optionally, double-stranded target nucleic acid is immobilized to a solid support. Optionally, the composition further includes a ligase

and/or a polymerase and/or a restriction endonuclease, and/or a nicking endonuclease. For instance, the present invention provides a nucleic acid adaptor comprising (a) a type IIs restriction endonuclease recognition sequence (TRERS); and (b) a nicking agent recognition sequence (NARS), wherein when the adaptor is linked to a double-stranded target nucleic acid fragment so that, in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, the target nucleic acid fragment is located 5' to the NARS, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS in said strand, if the TRERS is present, is located both (i) within the double-stranded target nucleic acid fragment, and (ii) 5' to the position corresponding to the NS.

A template nucleic acid molecule, which may be used to synthesize a single-stranded cDNA molecule. The template nucleic acid molecule includes: (a) a portion that may be characterized as being derived from a double-stranded cDNA molecule or a fragment thereof; and (b) a portion that may be characterized as being derived from a nucleic acid adaptor, where the (a) portion is ligated to the (b) portion. The nucleic acid adaptor portion (b) includes a nicking agent recognition sequence (NARS). In the combined presence of (i) the template nucleic acid molecule, (ii) a nicking agent (NA) that recognizes the NARS and (iii) a DNA polymerase, a single-stranded molecule is amplified, where at least part of the single-stranded molecule has a nucleotide sequence that is identical to at least part of the nucleotide sequence of one strand of the double-stranded cDNA. Preferably, said part is at least 6, or at least 7, or at least 8, or at least 9, or at least 10, or at least 11, or at least 12, or at least 13, or at least 14, or at least 15, or at least 16, or at least 17, or at least 18, or at least 19, or at least 20 nucleotides in length. The following criteria may be used, alone or in any combination, to further characterize the template nucleic acid molecule of the present invention: the NA that recognizes the NARS is a nicking endonuclease (NE); the NA that recognizes the NARS is N.BstNB I; the NARS is a hemimodified restriction endonuclease recognition sequence. Optionally, at least one strand of the template nucleic acid molecule is immobilized to a solid support. The immobilization may be, *e.g.*, via a terminus of either the double-stranded cDNA molecule or the nucleic acid adaptor that is not, was not, and will not be involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor. Optionally, a linker molecule may be present between the solid support and the terminus of the template nucleic acid molecule which is immobilized to the solid support. Optionally, a linker moiety may be present between the solid support and the terminus of the template nucleic acid molecule which is most immediately immobilized to the solid support, where the linker

comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS). A plurality of these template nucleic acid molecules may be used to form a cDNA library. In one embodiment, each of the plurality of the template nucleic acid molecules that form the library is immobilized to a solid support. Optionally, in this library, the NA that recognizes the NARS is a nicking endonuclease (NE), *e.g.*, N.BstNB I. Alternatively, as already mentioned, the NARS may be a hemimodified restriction endonuclease recognition sequence.

These and related aspects and embodiments of the present invention are set forth in more detail below.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic diagram of major steps in a method of the present invention for identifying a genetic variation at a defined location in a target nucleic acid using an exemplary ODNP pair containing a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a sequence of one strand of type IIs restriction endonuclease recognition sequence (TRERS), respectively.

Figure 2 is a schematic diagram of major steps in a method of the present invention for identifying a genetic variation at a defined location in a target nucleic acid using an exemplary ODNP pair each containing a sequence of a sense strand of a NERS.

Figure 3 is a diagram of major steps in a method of the present invention for identifying a nucleotide at a defined position in a target nucleic acid using an ODNP pair, Bsl I as an exemplary IRERS, and N.BstNB I as an exemplary NERS.

Figure 4 is a schematic diagram of the major components of an interrupted restriction endonuclease recognition sequence. $A_1A_2...A_m$ is a specific nucleotide sequence consisting of m nucleotides, whereas $A'_1A'_2...A'_m$ is the complement sequence of $A_1A_2...A_m$. The double-stranded fragment comprised of $A_1A_2...A_m$ and $A'_1A'_2...A'_m$ forms the first CRS. $N_1N_2...N_n$ is a variable nucleotide sequence consisting of n nucleotides where any one of the nucleotide can contain any of the four bases (A, C, T, or G). $N'_1N'_2...N'_n$ is the complement of $N_1N_2...N_n$ and forms a VRS in combination of $N_1N_2...N_n$. $C_1C_2...C_i$ is a specific nucleotide sequence consisting of i nucleotides, whereas $C'_1C'_2...C'_i$ is the complement of $C_1C_2...C_i$. The double-stranded fragment comprised of $C_1C_2...C_i$ and $C'_1C'_2...C'_i$ forms the second CRS.

Figure 5 is a schematic diagram of major steps in a method of the present invention for identifying a genetic variation at a defined location in a target nucleic acid

using an exemplary ODNP pair each containing a sequence of a sense strand of a hemimodified RERS. The broken lines represent extension/amplification products that have incorporated modified deoxynucleotides.

Figure 6 is a schematic diagram of major steps in a method of the present invention for detecting the presence, or the absence, of a target nucleic acid in a biological sample.

Figure 7 is a chromatogram obtained from electrospray-liquid-chromatography/mass spectrometry-time of flight (ES-LC/MS-TOF) analysis of a 1:10 dilution of a 4-mer ODN (5′-ACGA-3′).

Figure 8 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:100 dilution of a 4-mer ODN (5′-ACGA-3′).

Figure 9 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:1000 dilution of a 4-mer ODN (5′-ACGA-3′).

Figures 10A and 10B are chromatograms obtained from (ES-LC/MS-TOF) analysis of a 1:10 dilution of a 6-mer ODN (5′-ACGATG-3′).

Figure 11 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:100 dilution of a 6-mer ODN (5′-ACGATG-3′).

Figure 12 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:1000 dilution of a 6-mer ODN (5′-ACGATG-3′).

Figures 13A and 13B are chromatograms obtained from (ES-LC/MS-TOF) analysis of a 1:10 dilution of a 8-mer ODN (5′-ACGATGCA-3′).

Figure 14 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:100 dilution of a 8-mer ODN (5′-ACGATGCA-3′).

Figures 15A and 15B are chromatograms obtained from (ES-LC/MS-TOF) analysis of a 1:10 dilution of a 10-mer ODN (5′-GAACATCCAT-3′).

Figure 16 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:100 dilution of a 10-mer ODN (5′-GAACATCCAT-3′).

Figure 17 is a chromatogram obtained from (ES-LC/MS-TOF) analysis of a 1:1000 dilution of a 10-mer ODN (5′-GAACATCCAT-3′).

Figure 18 is a high pressure liquid chromatography (HPLC) chromatogram of a set of 4, 6, 8 and 10 nucleotide ODNPs.

Figures 19A and 19B show HPLC fractionation and detection of three 8-mers (Figure 19A) and three 10-mers (Figure 19B).

Figures 20A and 20B show the HPLC separation of one 4-mer, one 6-mer, three 8-mers and three 10-mers (Figure 20A) and the elution of two 6-mers (Figure 20B).

Figure 21 is a UV chromatogram of products of a PCR reaction after being digested with both N.BstNB I and Bsl I. The arrow indicates the single-stranded nucleic acid digestion product.

Figures 22A and 22B are single ion chromatogram and total ion chromatogram, respectively, of products of a PCR reaction after being digested with both N.BstNB I and Bsl I.

Figure 23 shows the mass spectrum of a single-stranded oligonucleotide excised from a double-stranded PCR product with double digestion of N.BstNB I and Bsl I.

Figures 24A and 24B show a UV chromatogram (Figure 24A) and an extracted ion current chart (Figure 24B) for the sequence 5′-GAGCACAGGATG-3′ (the top strand) and for 5′-CATCCTGTGCTC-3′ (the bottom stand) amplified from a genomic DNA of one individual.

Figures 25A and 25B show a UV chromatogram (Figure 25A) and an extracted ion current chart (Figure 25B) for the sequence 5′-GAGCACAGGATG-3′ (the top strand) and for 5′-CATCCTGTGCTC-3′ (the bottom stand) amplified from a genomic DNA of another individual.

Figures 26A and 26B show a UV chromatogram (Figure 26A) and an extracted ion current chart (Figure 26B) for a no template control.

Figures 27A and 27B show the mass spectra for the top fragment (m/z =3 @ 1265 amu) (Figure 27B) and for the bottom fragment (m/z=3 @1216 amu) (Figure 27A).

Figure 28 is a schematic diagram of major steps in an exemplary method of the present invention for preparing single-stranded nucleic acid probe. Immobilized poly(dT) probe is used to isolate mRNA and to function as a primer for synthesizing cDNA. The synthesized cDNA is then ligated with an exemplary adaptor of the present invention, which comprises a nicking endonuclease recognition sequence (NERS) and a type IIs restriction endonuclease recognition sequence (TRERS). The ligated nucleic acid fragment is then digested with a type IIs restriction endonuclease and used as the template for synthesizing short single-stranded nucleic acid probes in the presence of a DNA polymerase and a nicking endonuclease that recognizes the NERS in the nucleic acid adaptor.

Figures 29A-H are schematic diagrams of ligation products of exemplary adaptors and a target nucleic acid. The target nucleic acid is represented by broken lines. The arrows above or below the TRERS (the abbreviation for "type IIs restriction endonuclease recognition sequence") indicate the directions from the TRERS to the

cleavage site of a restriction endonuclease (RE) that recognizes the TRERS. Likewise, the arrows above or below the NERS (the abbreviation for "nicking endonuclease recognition sequence") indicate the directions from the NERS to the nicking site of a nicking endonuclease (NE) that recognizes the NERS.

Figures 30A and 30B illustrate the major steps of the present method for determining the presence (Fig. 30A), or the absence (Fig. 30B)), of a particular exon-exon junction (the junction of Exon A and Exon B) in a target cDNA molecule.

Figures 31A-C illustrate the major steps of the present method for determining the presence, or the absence, of every exon-exon junction in a cDNA derived from a gene of interest. Figure 31A is a schematic diagram of a gene of interest and two ODNP sets useful in the present method. The broken lines represent the portion of the gene that is not shown. Figure 31B is a schematic diagram of a target cDNA molecule (Target cDNA I) and amplified template nucleic acid fragments. Figure 31B is a schematic diagram of another target cDNA (Target cDNA II) and amplified template nucleic acid fragments.

Figure 32 shows mass spectrometric analyses of amplified single-stranded fragments containing an biallelic SNP (*i.e.*, A or T) or its complementary nucleotide derived from Individual No. 1 who is homozygous for the A allele. The top panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the T allele (referred to as "the bottom strand of the T allele"). The second panel shows the extracted ion current chromatogram for a fragment containing the T allele (referred to as "the upper strand of the T allele"). The third panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the A allele (referred to as "the bottom strand of the A allele"). The fourth panel shows the extracted ion current chromatogram for a fragment contain the A allele (referred to as "the top strand of the A allele"). The bottom panel shows the chromatogram for the total ion current.

Figure 33 shows mass spectrometric analyses of amplified single-stranded fragments containing an biallelic SNP (*i.e.*, A or T) or its complementary nucleotide derived from Individual No. 2 who is heterozygous for the A and T alleles. The top panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the T allele (referred to as "the bottom strand of the T allele"). The second panel shows the extracted ion current chromatogram for a fragment containing the T allele (referred to as "the upper strand of the T allele"). The third panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the A allele (referred to as "the bottom strand of the A

allele"). The fourth panel shows the extracted ion current chromatogram for a fragment contain the A allele (referred to as "the top strand of the A allele"). The bottom panel shows the chromatogram for the total ion current.

Figure 34 shows mass spectrometric analyses of amplified single-stranded fragments containing an biallelic SNP (*i.e.*, A or T) or its complementary nucleotide derived from a sample pooled from 200 individuals that are homozygous for the A allele and one individual that are homozygous for the T allele. The top panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the T allele (referred to as "the bottom strand of the T allele"). The second panel shows the extracted ion current chromatogram for a fragment containing the T allele (referred to as "the upper strand of the T allele"). The third panel shows the extracted ion current chromatogram for a fragment containing a complementary nucleotide of the A allele (referred to as "the bottom strand of the A allele"). The fourth panel shows the extracted ion current chromatogram for a fragment contain the A allele (referred to as "the top strand of the A allele"). The bottom panel shows the chromatogram for the total ion current.

Figure 35 shows the diode array trace (UV trace) for the amplified oligonucleotide fragments same as those in Figure 34 between 2.5 minutes and 5 minutes (the upper panel) and the total ion current (the bottom panel).

Figure 36 shows a schematic diagram of the major steps for preparing a template nucleic acid molecule by annealing a trigger nucleic acid derived from a genomic DNA to a single-stranded nucleic acid probe and subsequent amplification of a single-stranded nucleic acid molecule.

Figure 37 shows a schematic diagram of the major steps for preparing a template nucleic acid molecule from a genomic DNA and subsequent amplification of a single-stranded nucleic acid molecule. The genomic DNA comprises a nicking agent recognition sequence. The template molecule is produced by annealing one strand of the genomic DNA fragment to a single-stranded nucleic acid probe that is a portion of the other strand of the genomic DNA fragment.

Figure 38 shows a schematic diagram of the major steps for preparing a template nucleic acid molecule from a genomic DNA and subsequent amplification of a single-stranded nucleic acid molecule. The genomic DNA comprises a nicking agent recognition sequence and a restriction endonuclease recognition sequence. A nicking endonuclease recognition sequence recognizable by a nicking endonuclease that nicks outside its recognition sequence is used as an exemplary nicking agent recognition sequence.

Figure 39 shows a schematic diagram of the major steps for preparing a template nucleic acid molecule from a target nucleic acid using two oligonucleotide primers and subsequent amplification of a single-stranded nucleic acid molecule. One primer comprises a sequence of a sense strand of a NERS while the other comprises one strand of a Type-IIs restriction endonuclease recognition sequence (TRERS).

Figure 40 shows a schematic diagram of the major steps for preparing template nucleic acid molecules using two ODNPs and subsequent amplification of single-stranded nucleic acid molecules. In this exemplary embodiment, both ODNPs comprise a sequence of the sense strand of a NERS.

Figure 41 shows a schematic diagram of the major steps for preparing a template nucleic acid molecule in an exemplary embodiment using two ODNPs and subsequent amplification of a single-stranded nucleic acid molecule. Both ODNPs comprise a sequence of one strand of a RERS. The amplification is performed in the presence of an α-thio deoxynucleoside triphosphate, which is used as an exemplary modified deoxynucleoside triphosphate.

Figure 42 shows a schematic diagram of a method for detecting an immobilized target nucleic acid using a single-stranded nucleic acid probe that comprises a sequence of the antisense strand of a NARS.

Figure 43 shows a schematic diagram of a method for amplifying a single-stranded nucleic acid molecule using an oligonucleotide primer that comprises a sequence of the sense strand of a nicking agent recognition sequence.

Figure 44 shows a schematic diagram of a method for detecting the presence of a target nucleic acid in using an immobilized single-stranded nucleic acid probe that comprises a sequence of the sense strand of a NARS and a sequence that is at least substantially complementary to the 3' portion of the target nucleic acid.

Figure 45 shows a schematic diagram of a method for detecting the presence of a target nucleic acid in using an immobilized single-stranded nucleic acid probe that comprises a sequence of the sense strand of a NARS and is at least substantially complementary to the target nucleic acid.

Figure 46 shows a schematic diagram of a method for amplifying a single-stranded nucleic acid molecule using an oligonucleotide primer that comprises a sequence of the sense strand of a nicking agent recognition sequence and a partially double-stranded nucleic acid molecule that comprise a double-stranded type IIs restriction endonuclease recognition sequence (TRERS).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides methods, compositions and kits for amplifying single-stranded nucleic acid fragments where the amplification method is useful in, *e.g.*, identifying genetic variations; preparing oligonucleotide probes; and detecting the presence, or the absence, of nucleic acids in biological samples or of junctions between exons in cDNAs and mRNAs. The amplification is performed by combining (a) a template double-stranded nucleic acid that comprises a nicking agent recognition sequence (NARS) and (b) a nicking site (NS) with a nicking agent (NA) that recognizes the NARS. The double-stranded nucleic acid template is then nicked at the NS by the nicking agent to thereby produce a 3′ terminus at the NS. In the presence of a 5′→3′ exonuclease deficient DNA polymerase, the 3′ terminus at the NS is extended, replacing the downstream single-stranded DNA fragment (*i.e.*, the single-stranded DNA fragment having a 5′ terminus which was at the NS). The extension product is then nicked by the nicking agent and the thus-formed 3′ terminus at the NS is again extended by the DNA polymerase to form another extension product. This nicking-extension process is repeated multiple times, resulting in the amplification of the single-stranded DNA fragment having a 5′ terminus which is created by the nicking agent at the NS.

In various aspects, the present invention uses the above amplification method in various useful applications, including identification of genetic variations in target nucleic acids, preparation of oligonucleotide probes, and detecting the presence, or the absence, of a particular nucleic acid in a biological sample or of a particular exon-exon junction in a cDNA or mRNA molecule or population. Such applications are useful in many areas, including genetic analysis for hereditary diseases, tumor or pathogen diagnosis, identifying disease predisposition, forensics, paternity determination, creating and monitoring crop cultivation or animal superior breeding programs, expression profiling of cell function and/or disease marker genes, and identification and/or characterization of infectious organisms that cause infectious diseases in plants or animals and/or that are related to food safety.

A. Conventions/Definitions

Prior to providing a more detailed description of the present invention, it may be helpful to an understanding thereof to define some convention(s) and term(s) as used herein, as follows. Additional definitions may be found throughout the description of the present invention.

The terms "3'" and "5'" are used herein to describe the location of a particular site within a single strand of nucleic acid. When a location in a nucleic acid is "3' to" or "3' of" a reference nucleotide or a reference nucleotide sequence, this means that the location is between the 3' terminus of the reference nucleotide or the reference nucleotide sequence and the 3' hydroxyl of that strand of nucleic acid. Likewise, when a location in a nucleic acid is "5' to" or "5' of" a reference nucleotide or a reference nucleotide sequence, this means that the location is between the 5' terminus of the reference nucleotide or the reference nucleotide sequence and the 5' phosphate of that strand of nucleic acid. Further, when a subject nucleotide sequence is "directly 3' to" or "directly 3' of" a reference nucleotide or a reference nucleotide sequence, this means that the subject nucleotide sequence is immediately next to the 3' terminus of the reference nucleotide or the reference nucleotide sequence. Similarly, when a subject nucleotide sequence is "directly 5' to" or "directly 5' of" a reference nucleotide or a reference nucleotide sequence, this means that the subject nucleotide sequence is immediately next to the 5' terminus of the reference nucleotide or the reference nucleotide sequence.

As used herein, "nicking" refers to the cleavage of only one strand of a double-stranded nucleic acid molecule or a double-stranded portion of a partially double-stranded nucleic acid molecule at a specific position relative to a nucleotide sequence that is recognized by the enzyme that performs the nicking. The specific position where the nucleic acid is nicked is referred to as the "nicking site" (NS).

A "nicking agent" (NA) is an enzyme that recognizes a particular nucleotide sequence of a completely or partially double-stranded nucleic acid molecule and cleaves only one strand of the nucleic acid molecule at a specific position relative to the recognition sequence. Nicking agents include, but are not limited to, a nicking endonuclease (*e.g.*, N.BstNB I) and a restriction endonuclease (*e.g.*, Hinc II) when the double-stranded (or partially double-stranded) nucleic acid molecule contains a hemimodified recognition/cleavage sequence in which one strand contains at least one derivatized nucleotide(s) that prevents cleavage of that strand (*i.e.*, the strand that contains the derivatized nucleotide(s)) by the restriction endonuclease.

A "nicking endonuclease" (NE), as used herein, refers to an endonuclease that recognizes a nucleotide sequence of a completely or partially double-stranded nucleic acid molecule and cleaves only one strand of the nucleic acid molecule at a specific location relative to the recognition sequence. Unlike a restriction endonuclease (RE), which requires its recognition sequence to be modified by containing at least one derivatized nucleotide to prevent cleavage of the derivatized

nucleotide-containing strand of a fully or partially double-stranded nucleic acid molecule, a NE typically recognizes a nucleotide sequence composed of only native nucleotides and cleaves only one strand of a fully or partially double-stranded nucleic acid molecule that contains the nucleotide sequence.

As used herein, "native nucleotide" refers to adenylic acid, guanylic acid, cytidylic acid, thymidylic acid or uridylic acid. A "derivatized nucleotide" is a nucleotide other than a native nucleotide.

The nucleotide sequence of a completely or partially double-stranded nucleic acid molecule that a NA recognizes is referred to as the "nicking agent recognition sequence" (NARS). Likewise, the nucleotide sequence of a completely or partially double-stranded nucleic acid molecule that a NE recognizes is referred to as the "nicking endonuclease recognition sequence" (NERS). The specific sequence that a RE recognizes is referred to as the "restriction endonuclease recognition sequence" (RERS). A "hemimodified RERS," as used herein, refers to a double-stranded RERS in which one strand of the recognition sequence contains at least one derivatized nucleotide (*e.g.*, α-thio deoxynucleotide) that prevents cleavage of that strand (*i.e.*, the strand that contains the derivatized nucleotide within the recognition sequence) by a RE that recognizes the RERS.

In certain embodiments, a NARS is a double-stranded nucleotide sequence where each nucleotide in one strand of the sequence is complementary to the nucleotide at its corresponding position in the other strand. In such embodiments, the sequence of a NARS in the strand containing a NS nickable by a NA that recognizes the NARS is referred to as a "sequence of the sense strand of the NARS" or a "sequence of the sense strand of the double-stranded NARS," while the sequence of the NARS in the strand that does not contain the NS is referred to as a "sequence of the antisense strand of the NARS" or a "sequence of the antisense strand of the double-stranded NARS. "

Likewise, in the embodiments where a NERS is a double-stranded nucleotide sequence of which one strand is exactly complementary to the other strand, the sequence of a NERS located in the strand containing a NS nickable by a NE that recognizes the NERS is referred to as a "sequence of a sense strand of the NERS" or a "sequence of the sense strand of the double-stranded NERS," while the sequence of the NERS located in the strand that does not contain the NS is referred to a "sequence of the antisense strand of the NERS" or a "sequence of the antisense strand of the double-stranded NERS. " For example, the recognition sequence and the nicking site of an exemplary nicking endonuclease, N.BstNB I, are shown below with "▼" to indicate the cleavage site and N to indicate any nucleotide:

▼
5′-GAGTCNNNNN-3′
3′-CTCAGNNNNN-5′

The sequence of the sense strand of the N.BstNB I recognition sequence is 5′-GAGTC-3′, whereas that of the antisense strand is 5′-GACTC-3′.

Similarly, the sequence of a hemimodified RERS in the strand containing a NS nickable by a RE that recognizes the hemimodified RERS (*i.e.*, the strand that does not contain any derivatized nucleotides) is referred to as "the sequence of the sense strand of the hemimodified RERS" and is located in "the sense strand of the hemimodified RERS," while the sequence of the hemimodified RERS in the strand that does not contain the NS (*i.e.*, the strand that contains derivatized nucleotide(s)) is referred to as "the sequence of the antisense strand of the hemimodified RERS" and is located in "the antisense strand of the hemimodified RERS."

In certain other embodiments, a NARS is an at most partially double-stranded nucleotide sequence that has one or more nucleotide mismatches, but contains an intact sense strand of a double-stranded NARS as described above. According to the convention used herein, in the context of describing a NARS, when two nucleic acid molecules anneal to one another so as to form a hybridized product, and the hybridized product includes a NARS, and there is at least one mismatched base pair within the NARS of the hybridized product, then this NARS is considered to be only partially double-stranded. Such NARSs may be recognized by certain nicking agents (*e.g.*, N.BstNB I) that require only one strand of double-stranded recognition sequences for their nicking activities. For instance, the NARS of N.BstNB I may contain, in certain embodiments, an intact sense strand, as follows,

5′-GAGTC-3′
3′-NNNNN-5′

where N indicates any nucleotide, and N at one position may or may not be identical to N at another position, however there is at least one mismatched base pair within this recognition sequence. In this situation, the NARS will be characterized as having at least one mismatched nucleotide.

In certain other embodiments, a NARS is a partially or completely single-stranded nucleotide sequence that has one or more unmatched nucleotides, but

contains an intact sense strand of a double-stranded NARS as described above. According to the convention used herein, in the context of describing a NARS, when two nucleic acid molecules (*i.e.,* a first and a second strand) anneal to one another so as to form a hybridized product, and the hybridized product includes a nucleotide sequence in the first strand that is recognized by a NA, *i.e.,* the hybridized product contains a NARS, and at least one nucleotide in the sequence recognized by the NA does not correspond to, *i.e.,* is not across from, a nucleotide in the second strand when the hybridized product is formed, then there is at least one unmatched nucleotide within the NARS of the hybridized product, and this NARS is considered to be partially or completely single-stranded. Such NARSs may be recognized by certain nicking agents (*e.g.*, N.BstNB I) that require only one strand of double-stranded recognition sequences for their nicking activities. For instance, the NARS of N.BstNB I may contain, in certain embodiments, an intact sense strand, as follows,

5′-GAGTC-3′
3′-N_{0-4}-5′

(where "N" indicates any nucleotide, 0-4 indicates the number of the nucleotides "N," a "N" at one position may or may not be identical to a "N" at another position), which contains the sequence of the sense strand of the double-stranded recognition sequence of N.BstNB I. In this instance, at least one of G, A, G, T or C is unmatched, in that there is no corresponding nucleotide in the complementary strand. This situation arises, *e.g.,* when there is a "loop" in the hybridized product, and particularly when the sense sequence is present, completely or in part, within a loop.

An amplified single-stranded nucleic acid fragment is, in some aspects of the present invention a single-stranded nucleic acid fragment that is, at most 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, or 100 nucleotides in length. The short length of such a fragment facilitates its detection and/or characterization via various techniques, including mass spectrometry and liquid chromatography, as described in detailed below. However, in certain embodiments of some aspects (*e.g.*, making of single-stranded nucleic acid probes), the amplified single-stranded nucleic acid fragment may be relatively long, for example, it may be 51-1000 nucleotides in length.

A nucleotide in one strand (the first strand) of a double-stranded nucleic acid that is located at a position "corresponding to" a position (*i.e.*, a defined position) in the other strand (the second strand) of the double-stranded nucleic acid refers to the nucleotide in the first strand that hydrogen bonds with and is complementary to the nucleotide at the defined position in the second strand when the first and second strands

hybridize to one another so as to form the double-stranded nucleic acid. Likewise, a position in one strand (the first strand) of a double-stranded nucleic acid "corresponding to" a nicking site within the other strand (the second strand) of the double-stranded nucleic acid refers to the position between the two nucleotides in the first strand that are complementary to the two nucleotides in the second strands that surround the nicking site.

B. Detection of Genetic Variations

In one aspect, the present invention provides methods, compounds and kits for detecting genetic variations in target nucleic acids.

1. Methods for Detecting Genetic Variations

The present invention provides methods for detecting genetic variations at defined locations in target nucleic acids. The methods comprise (a) providing a double-stranded template nucleic acid that comprises a fragment of a target nucleic acid and a nicking site (NS) located 5′ to a genetic variation; (b) amplifying, in the presence of a DNA polymerase and a nicking agent that nicks at the NS, a single-stranded fragment that comprises the genetic variation and of which 5′ terminus is at the NS; and (c) characterizing the single-stranded fragment and thereby identifying the genetic variation. In one embodiment, the single-stranded fragment is amplified in the presence of a nicking endonuclease. In another embodiment, the fragment is amplified in the presence of a restriction endonuclease.

As to the method using a nicking endonuclease, the double-stranded template nucleic acid may be provided by first forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP and a target nucleic acid that contains a genetic variation at a defined location. In the cases where the target nucleic acid is double stranded, as illustrated in Figure 1, the first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of one strand of the target nucleic acid located 3′ to the complement of the genetic variation. The second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence 3′ to the genetic variation located on the other strand of the target nucleic acid, and optionally comprises a nucleotide sequence of one strand of a restriction endonuclease recognition sequence (RERS).

In the cases where the target nucleic acid is single stranded, the first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide

sequence at least substantially identical to the target nucleic acid located 5′ to the genetic variation, while the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3′ to the genetic variation and optionally comprises a nucleotide sequence of one strand of a RERS.

The extension of the first and the second ODNPs using the target nucleic acid as template produces an extension product having a NERS and optionally having a RERS where the genetic variation is located between the NERS and RERS. This extension product is the desired double-stranded template nucleic acid that comprises a fragment of a target nucleic acid and a NS located 5′ to a genetic variation. The extension product, if it contains a RERS, may then be cleaved with a restriction endonuclease that recognizes the RERS, and used as template for amplifying a single-stranded nucleic acid molecule in the presence of a NE that recognizes the NERS. Otherwise, the extension product may be directly used as a template for amplifying a single-stranded nucleic acid molecule. An exemplary process of the method is illustrated in Figure 1, using TRERS (Type IIs restriction endonuclease recognition sequence) as an exemplary RERS.

Alternatively, as illustrated in Figure 2, the double-stranded template nucleic acid may be provided by first forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP and a double-stranded target nucleic acid that contains a genetic variation at a defined location. The first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of one strand of the target nucleic acid located 5′ to the complement of the genetic variation. The second ODNP comprises a sequence of a sense strand of a NERS instead of a RERS as described above. The extension of the first and the second ODNPs using the target nucleic acid as template produces an extension product having two NERSs. This extension product is the desired double-stranded template nucleic acid that comprises a fragment of a target nucleic acid and a NS located 5′ to a genetic variation. The extension product may be then used as the template for amplifying a single-stranded nucleic acid fragment that comprises the genetic variation and of which 5′ terminus is at the NS, by using a NE that recognizes the NERS.

As to the method using a restriction endonuclease for nicking a double-stranded nucleic acid, the double-stranded template nucleic acid may be provided by first forming a mixture of a first ODNP, a second ODNP, and a double-stranded target nucleic acid (Figure 5). In the cases where the target nucleic acid is double-stranded,

the first ODNP comprises a sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of one strand of the target nucleic acid located at 3′ to the complement of the genetic variation, while the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3′ to the defined location. In the cases where the target nucleic acid is single stranded, the first ODNP comprises a sequence of one strand of a first RERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5′ to the complement of the genetic variation, while the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3′ to the genetic variation. The extension of the first and second ODNPs in the presence of deoxyribonucleoside triphosphates and at least one modified deoxyribonucleoside triphosphate produces a fragment having both the first and the second hemimodified RERSs. Such an extension product may be used as a template for amplifying a single-stranded nucleic acid molecule in the presence of REs that recognize the first RERS and the second RERS. In certain embodiments, the first RERS is the same as the second RERS. Thus, the extension product may be amplified in the presence of a RE that recognizes both the first and the second RERSs.

In all the above-described embodiments, the amplification of single-stranded nucleic acid molecules is carried out in the presence of a DNA polymerase in addition to a nicking agent. The amplified single-stranded nucleic acid molecules contain genetic variations of the target nucleic acids and are then characterized. The genetic variation(s) of the target nucleic acid is thereby identified.

a. Target Nucleic Acids

The target nucleic acid of the present invention related to identifying genetic variations is any nucleic acid molecule that may contain a genetic variation using a wild type nucleic acid sequence as a reference and that can serve as a template for a primer extension reaction, *i.e.*, can base pair with an oligonucleotide primer.

The term "nucleic acid" refers generally to any molecule, preferably a polymeric molecule, incorporating units of ribonucleic acid or an analog thereof. The target nucleic acid can be either single-stranded or double-stranded. A single-stranded target nucleic acid may be one strand of a denatured double-stranded DNA. Alternatively, it may be a single-stranded nucleic acid not derived from any double-

stranded DNA. In one aspect, the target nucleic acid is DNA. In another aspect, the target is RNA. Suitable nucleic acid molecules are DNA, including genomic DNA, ribosomal DNA and cDNA. Other suitable nucleic acid molecules are RNA, including mRNA, rRNA and tRNA. The nucleic acid molecule may be naturally occurring, as in genomic DNA, or it may be synthetic, *i.e.*, prepared based on human action, or may be a combination of the two.

In one aspect, the target nucleic acid either is or is derived from naturally occurring nucleic acid. A naturally occurring target nucleic acid is obtained from a biological sample. Preferred biological samples include one or more mammalian tissues, preferably human tissues, (for example blood, plasma/serum, hair, skin, lymph node, spleen, liver, etc.) and/or cells or cell lines. The biological samples may comprise one or more human tissues and/or cells. Mammalian and/or human tissues and/or cells may further comprise one or more tumor tissues and/or cells.

Methodology for isolating populations of nucleic acids from biological samples is well known and readily available to those skilled in the art of the present invention. Exemplary techniques are described, for example, in the following laboratory research manuals: Sambrook *et al.*, "Molecular Cloning" (Cold Spring Harbor Press, 3rd Edition, 2001) and Ausubel *et al.*, "Short Protocols in Molecular Biology" (1999) (incorporated herein by reference in their entireties). Nucleic acid isolation kits are also commercially available from numerous companies, and may be used to simplify and accelerate the isolation process.

A synthetic target nucleic acid is produced by human intervention. At this time, many companies are in the business of making and selling synthetic nucleic acids that may be useful as the target nucleic acid molecule in the present invention. These companies may or will also prepare primers that are useful in the present invention. *See, e.g.*, Applied Bio Products Bionexus (www.bionexus.net); Commonwealth Biotechnologies, Inc. (Richmond, VA; www.cbi-biotech.com); Gemini Biotech (Alachua, Florida; www.geminibio.com); INTERACTIVA Biotechnologie GmbH (Ulm, Germany; www.interactiva.de); Microsynth (Balgach, Switzerland; www.microsynth.ch); Midland Certified Reagent Company (Midland, TX; www.mcrc.com); Oligos Etc. (Wilsonville, OR; www.oligosetc.com); Operon Technologies, Inc. (Alameda, CA; www.operon.com); Scandanavian Gene Synthesis AB (Köping, Sweden; www.sgs.dna); Sigma-Genosys (The Woodlands, Texas; www.genosys.com); Synthetic Genetics (San Diego, CA; www.syntheticgenetics.com, owned by Epoch Biosciences, Inc. (Bothell, WA; www.epochbio.com); and many others.

The synthetic nucleic acid target may be prepared using an amplification reaction. The amplification reaction may be, for example, the polymerase chain reaction (PCR).

The synthetic nucleic acid target may be prepared using recombinant DNA means through production in one or more prokaryotic or eukaryotic organism such as, *e.g.*, *E. coli*, yeast, Drosophila or mammalian tissue culture cell line.

The target nucleic acid molecule may, and typically will, contain one or more of the natural bases present in nucleotides, *i.e.*, adenine (A), guanine (G), cytosine (C), thymine (T) and, in the case of an RNA, uracil (U). In addition, and particularly when the nucleic acid is a synthetic molecule, the target nucleic acid may include "unnatural" nucleotides. Unnatural nucleotides are chemical moieties that can be substituted for one or more natural nucleotides in a nucleotide chain without causing the nucleic acid to lose its ability to serve as a template for a primer extension reaction. The substitution may include either sugar and/or phosphate substitutions, in addition to base substitutions.

Such moieties are very well known in the art, and are known by a large number of names including, for example, abasic nucleotides, which do not contain a commonly recognized nucleotide base, such as adenine, guanine, cytosine, uracil or thymine (*see, e.g.*, Takeshita et al. "Oligonucleotides containing synthetic abasic sites" *The Journal of Biological Chemistry*, vol. 262, pp. 10171-10179 1987; Iyer et al. "Abasic oligodeoxyribonucleoside phosphorothioates: synthesis and evaluation as anti-HIV-1 agents" *Nucleic acids Research*, vol. 18, pp. 2855-2859 1990; and U.S. Patent 6,117,657); base or nucleotide analogs (*see, e.g.*, Ma et al., "Design and Synthesis of RNA Miniduplexes via a Synthetic Linker Approach. 2. Generation of Covalently Closed, Double-Stranded Cyclic HIV-1 TAR RNA Analogs with High Tat-Binding Affinity," *Nucleic Acids Research 21*:2585 (1993). Some bases are known as universal mismatch base analogs, such as the abasic 3-nitropyrrole); convertides (*see, e.g.*, Hoops et al., *Nucleic Acids Res. 25*:4866-4871 (1997)); modified nucleotides (*see, e.g.*, Millican et al., "Synthesis and biophysical studies of short oligodeoxynucleotides with novel modifications: A possible approach to the problem of mixed base oligodeoxynucleotide synthesis," *Nucleic Acids Research 12*:7435-7453 (1984); nucleotide mimetics; nucleic acid related compounds; spacers (*see, e.g.*, Nielsen et al., *Science 254*:1497-1500 (1991); and specificity spacers (*see, e.g.*, PCT International Publication No. WO 98/13527).

Additional examples of non-natural nucleotides are set forth in: Jaschke et al., *Tetrahedron Lett. 34*:301 (1993); Seela and Kaiser, *Nucleic Acids Research*

15:3113 (1990) and *Nucleic Acids Research 18*:6353 (1990); Usman et al., PCT International Patent Application No. PCT/US 93/00833; Eckstein, PCT International Patent Application No. PCT/EP91/01811; Sproat et al., U.S. Patent No. 5,334,711, and Buhr and Matteucci, PCT International Publication No. WO 91/06556; Augustyns, K. A. et al., Nucleic Acids Res., 1991, 19, 2587-2593); and U.S. Patent Nos. 5,959,099 and 5,840,876.

When the target nucleic acid molecule, and/or the primer used in the present method, contains a non-natural nucleotide, then a base-pair mismatch will occur between the target and the primer. The term "base-pair mismatch" refers to all single and multiple nucleotide substitutions that perturb the hydrogen bonding between conventional base pairs (*e.g.*, G:C, A:T, and A:U) by substitution of a nucleotide with a moiety that does not hybridize according to the standard Watson-Crick model to a corresponding nucleotide on the opposite strand of the oligonucleotide duplex. Such base-pair mismatches include, *e.g.*, G:G, G:T, G:A, G:U, C:C, C:A, C:T, C:U, T:T, T:U, U:U and A:A. Also included within the definition of base-pair mismatches are single or multiple nucleotide deletions or insertions that perturb the normal hydrogen bonding of a perfectly base-paired duplex. In addition, base-pair mismatches arise when one or both of the nucleotides in a base pair have undergone a covalent modification (*e.g.*, methylation of a base) that disrupts the normal hydrogen bonding between the bases. Base-pair mismatches also include non-covalent modifications such as, for example, those resulting from incorporation of intercalating agents such as ethidium bromide and the like that perturb hydrogen bonding by altering the helicity and/or base stacking of an nucleic acid duplex.

The target, in addition to containing nucleic acids or analogs thereof, also contains one or more natural bases of unknown identity (*i.e.*, potential genetic variations). The present invention provides compositions and methods whereby the identity of the unknown nucleotide(s) becomes known and thereby the genetic variation becomes identified. The base(s) of unknown identity is present at the "nucleotide locus" (or the "defined position" or the "defined location"), which refers to a specific nucleotide or region encompassing one or more nucleotides having a precise location on a target nucleic acid.

In certain embodiments, the target nucleic acids are immobilized to a solid support. Any solid supports to which nucleic acid molecules may directly or indirectly attach can be used. Such supports include nitrocellulose, nylon, glass and metal. Target nucleic acids may be immobilized to a solid support without being first purified or isolated. Alternatively, they may be first purified from a sample that

contains the target nucleic acids and then immobilized to a solid support. Methods for immobilizing nucleic acid molecules to a solid support are well known in the art, such as tissue printing and the use of poly(ethyleneimine).

The terms "polymorphism" and "genetic variation," as used herein, refer to the occurrence of two or more genetically determined alternative sequences or alleles in a small region (*i.e.*, one to several (*e.g.*, 2, 3, 4, 5, 6, 7, or 8) nucleotides in length) in a population. The allelic form occurring most frequently in a selected population is referred to as the wild type form. Other allelic forms are designated as variant forms. Diploid organisms may be homozygous or heterozygous for allelic forms.

Genetic variations may or may not have effects on gene expression, including expression levels and expression products (*i.e.*, encoded peptides). Genetic variations that affect gene expression are also referred to as "mutations," including point mutations, frameshift mutations, regulatory mutations, nonsense mutations, and missense mutation. A "point mutation" refers to a mutation in which a wild-type base (*i.e.*, A, C, G, or T) is replaced with one of the other standard bases at a defined nucleotide locus within a nucleic acid sample. It can be caused by a base substitution or a base deletion. A "frameshift mutation" is caused by small deletions or insertions that, in turn, cause the reading frame(s) of a gene to be shifted and, thus, a novel peptide to be formed. A "regulatory mutation" refers to a mutation in a non-coding region, *e.g.*, an intron, a region located 5′ or 3′ to the coding region, that affects correct gene expression (*e.g.*, amount of product, localization of protein, timing of expression). A "nonsense mutation" is a single nucleotide change resulting in a triplet codon (where mutation occurs) being read as a "STOP" codon causing premature termination of peptide elongation, *i.e.*, a truncated peptide. A "missense mutation" is a mutation that results in one amino acid being exchanged for a different amino acid. Such a mutation may cause a change in the folding (3-dimensional structure) of the peptide and/or its proper association with other peptides in a multimeric protein.

In one aspect of the invention, the genetic variation is a "single-nucleotide polymorphism" (SNP), which refers to any single nucleotide sequence variation, preferably one that is common in a population of organisms and is inherited in a Mendelian fashion. Typically, the SNP is either of two possible bases and there is no possibility of finding a third or fourth nucleotide identity at an SNP site.

The genetic variation may be associated with or cause diseases or disorders. The term "associated with," as used herein, refers to the presence of a positive correlation between the occurrence of the genetic variation and the presence of a disease or a disorder in the host. Such diseases or disorders may be human genetic

diseases or disorders and include, but are not limited to, bladder carcinoma, colorectal tumors, sickle-cell anemia, thalassemias, al-antitrypsin deficiency, Lesch-Nyhan syndrome, cystic fibrosis/mucoviscidosis, Duchenne/Becker muscular dystrophy, Alzheimer's disease, X-chromosome-dependent mental deficiency, and Huntington's chorea, phenylketonuria, galactosemia, Wilson's disease, hemochromatosis, severe combined immunodeficiency, alpha-1-antitrypsin deficiency, albinism, alkaptonuria, lysosomal storage diseases, Ehlers-Danlos syndrome, hemophilia, glucose-6-phosphate dehydrogenase disorder, agammaglobulimenia, diabetes insipidus, Wiskott-Aldrich syndrome, Fabry's disease, fragile X-syndrome, familial hypercholesterolemia, polycystic kidney disease, hereditary spherocytosis, Marfan's syndrome, von Willebrand's disease, neurofibromatosis, tuberous sclerosis, hereditary hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos syndrome, myotonic dystrophy, osteogenesis imperfecta, acute intermittent porphyria, and von Hippel-Lindau disease.

Target nucleic acids may be amplified before being combined with ODNPs as described below. Any of the known methods for amplifying nucleic acids may be used. Exemplary methods include, but are not limited to, the use of Qbeta Replicase, Strand Displacement Amplification (Walker *et al.*, *Nucleic Acid Research 20*: 1691-6, 1995), transcription-mediated amplification (Kwoh *et al.*, PCT Int'l. Pat. Appl. Pub. No.WO88/10315), RACE (Frohman, *Methods Enzymol. 218*:340-56, 1993), one-sided PCR (Ohara *et al.*, *Proc. Natl. Acad. Sc. 86*: 5673-7, 1989), and gap-LCR (Abravaya *et al.*, *Nucleic Acids Res. 23*: 675-82, 1995). The cited articles and the PCT international patent application are incorporated herein by reference in their entireties.

b. Making of Double-Stranded Template Nucleic Acids

As noted above, a single-stranded nucleic acid molecule that contains a genetic variation from a target nucleic acid may be amplified in the presence of a nicking agent (*e.g.*, a NE or a RE) and a DNA polymerase. For the amplification of the single-stranded nucleic acid molecule using a NE as the nicking agent, the template nucleic acid for the amplification must contain a NERS (sometimes referred to as the first NERS). In addition, the template nucleic acid may also contain a RERS or a second NERS. The presence of either the RERS or the second NERS in addition to the first NERS allows for a cleavage in the strand of the template that does not contain the NS, near the site corresponding to the nicking site produced by a NE that binds to the first NERS. Such a cleavage allows the single-stranded nucleic acid molecule amplified from the template to be relatively short and thereby facilitates the

characterization of the genetic variation incorporated into the single-stranded nucleic acid molecule. For the amplification of the single-stranded nucleic acid molecule that uses a RE as the nicking agent, the template nucleic acid for the amplification must contain a hemimodified RERS and may contain an additional RERS. If two hemimodified RERSs are present, they may, or may not, be the same. Similar to the template that contains two NERSs, the presence of the second hemimodified RERS also allows the amplified single-stranded nucleic acid molecule to be relatively short.

i. Design of Oligonucleotide Primers (ODNPs)

The template nucleic acid of the present invention may be provided by amplifying a fragment of a target nucleic acid containing a genetic variation using specifically designed ODNPs. The term "oligonucleotide" (ODN) refers to a nucleic acid fragment (typically DNA or RNA) obtained synthetically as by a conventional automated nucleic acid (*e.g.*, DNA) synthesizer. Oligonucleotide is used synonymously with the term polynucleotide. The term "oligonucleotide primer" (ODNP) refers to any polymer having two or more nucleotides used in a hybridization, extension, and/or amplification reaction. The ODNP may be comprised of deoxyribonucleotides, ribonucleotides, or an analog of either. As used herein for hybridization, extension, and amplification reactions, ODNPs are generally between 8 and 200 bases in length. More preferred are ODNPs of between 12 and 50 bases in length and still more preferred are ODNPs of between 18 and 32 bases in length.

In one aspect, the present invention provides an ODNP pair (referred to as "the first ODNP pair") useful for producing a template nucleic acid containing a NERS, and optionally containing an RERS, where the genetic variation is located between the NERS and the optionally present RERS. For convenience, a double-stranded target nucleic acid is used in the following description of the ODNP pair. One ODNP of the first ODNP pair (referred to as "the first ODNP") comprises (1) a sequence of a sense strand of a NERS and (2) a nucleotide sequence at least substantially complementary to a nucleotide sequence of the strand of the target nucleic acid that contains the complement nucleotide(s) of the genetic variation. The nucleotide sequence of the target to which a portion of the first ODNP is complementary is located 3′ to the complement nucleotide(s) of the genetic variation. Such a design allows the extension product of the first ODNP to incorporate the genetic variation. The phrase "at least substantially complementary" refers to a degree of complementarity between a portion of the ODNP and a target nucleic acid sufficient to allow the ODNP to specifically anneal to the target and to function as a primer for extension/amplification.

In a preferred embodiment, the complementarity is exact. The sequence of the ODNP that complements the target can be located either 3′ or 5′ to the sequence of the sense strand of the NERS. Preferably, there exist sequences located both 5′ and 3′ to the sequence of the sense strand of the NERS that are at least substantially complementary to the target. The presence of a substantially or exactly complementary sequence located 3′ to the sequence of the sense strand of the NERS facilitates annealing of the primer to the template at a pre-defined location and increases extension/amplification efficiency. The presence of a substantially or exactly complementary sequence located 5′ to the sequence of the sense strand of the NERS reduces the number of nucleotides located 3′ to the sequence of the sense strand of the NERS that are needed for successful and efficient annealing and extension, and also shortens the length of the subsequently amplified single-stranded nucleic acid molecule as described in detail below. The complete sequence of the sense strand of the NERS of the ODNP may or may not be complementary to the corresponding region of the target. In one aspect, the complete sequence of the sense strand of the NERS of the ODNP is not exactly complementary to the corresponding region of the target. Generally, the ODNP contains at least 6, preferably 8, more preferably 10, most preferably 12, 14, or 16 nucleotides that are exactly complementary to the target nucleic acid.

The other ODNP of the first ODNP pair (referred to as "the second ODNP") comprises (1) a nucleotide sequence at least substantially complementary to a nucleotide sequence of the strand of the target nucleic acid that contains the genetic variation; and, optionally, (2) a sequence of one strand of a RERS. The nucleotide sequence of the ODNP to which a portion of the target is at least substantially complementary is designed to be located 3′ to the genetic variation. Such a design allows the extension product of the second ODNP to incorporate the complement of the genetic variation. Similar to the first ODNP, the complementarity between the annealing portion of the second ODNP and the corresponding portion of the target need not be exact, but must be sufficient to allow the second ODNP to specifically anneal to the target at a desired location and to function as a primer for extension/amplification. The sequence of the second ODNP that is at least substantially complementary to the target can be located either 3′ or 5′ to the sequence of the one strand of the RERS if a RERS is present. Preferably, there exist sequences located both 5′ and 3′ to the sequence of the RERS that are at least substantially complementary to the target. The sequence of the RERS of the second ODNP may or may not be complementary to the corresponding region of the target, and typically it will not be exactly complementary to the corresponding region of the target. Generally, the second ODNP contains at least 6,

preferably 8, more preferably 10, most preferably 12, 14, or 16 nucleotides that are exactly complementary to the target nucleic acid.

While the first ODNP pair has been described above in connection with a target nucleic acid that is double stranded, one of ordinary skill in the art could, with the guidance provided herein, readily design this ODNP pair as well as other ODNP pairs (*i.e.*, the second ODNP pair and the third ODNP pair as described below) for situations where the target nucleic acid is single stranded. Briefly, if the target nucleic acid is single stranded, one first ODNP comprises a nucleotide sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5′ to the genetic variation, while the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 5′ to the genetic variation and optionally comprises a sequence of one strand of a RERS. The phrase "at least substantially identical" refers to as degree of identity between a portion of the nucleotide sequence of an ODNP and a corresponding portion of a target nucleic acid sufficient to allow the ODNP to specifically anneal to a nucleic acid comprising a sequence that is exactly complementary to the corresponding portion of the target nucleic acid, and to function as a primer for extension/amplification with the nucleic acid as a template. In a certain preferred embodiment, a portion of the nucleotide sequence of the first ODNP is exactly identical to a portion of a target nucleic acid.

In another aspect, the present invention provides an ODNP pair (referred to as "the second ODNP pair") useful for producing a template nucleic acid containing one or two NERSs, where the genetic variation from the target is located between the two NERSs if a second NERS is present. The first ODNP of the second ODNP pair is the same as that the first ODNP of the first ODNP pair described above. The second ODNP of the second ODNP pair is the same as the second ODNP of the first ODNP pair except that the sequence of one strand of the RERS in the second ODNP of the first ODNP pair is replaced with a sequence of a sense strand of a NERS in the second ODNP of the second ODNP pair.

In yet another aspect, the present invention provides an ODNP pair (referred to as "the third ODNP pair") useful for producing a template nucleic acid containing one or two hemimodified RERSs, with a genetic variation located between the two hemimodified RERSs if a second hemimodified RERS is present. The first ODNP of the third ODNP pair is the same as the first ODNP of the first ODNP pair except that the sequence of the sense strand of the NERS in the first ODNP of the first ODNP pair is replaced with a sense strand of the hemimodified RERS. The second

ODNP of the third ODNP pair is the same as the second ODNP of the first ODNP pair except that the sequence of one strand of the RERS in the second ODNP of the first ODNP pair may be replaced with a sequence of a sense strand of a hemimodified RERS in the second ODNP of the third ODNP pair. The sequence of the sense strand of the hemimodified RERS in the first ODNP may or may not be identical to that in the second ODNP of the third ODNP pair. Preferably, they are the same.

In a further aspect, the present invention provide an ODNP pair (referred to as "the fourth ODNP pair") useful for producing a portion of a single-stranded nucleic acid containing a genetic variation to be identified at a defined location (Figure 3). For convenience, a single-stranded target nucleic acid is used in the following description of the ODNP pair. However, it is within the skill of an ordinary artisan to design ODNP pairs wherein the target nucleic acid is double stranded using the guidance provided herein.

One primer of the fourth ODNP pair ("the first ODNP") comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of a target nucleic acid at a location 3′ to the defined position ("the first region of the target nucleic acid"), whereas the other primer of the fourth ODNP pair ("the second ODNP") comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the complement of the target nucleic acid at a location 3′ to the complementary nucleotide of the nucleotide at the defined position ("the first region of the complement"). The complementarity between the ODNPs and their corresponding target nucleic acid, or the complement thereof, need not be exact, but must be sufficient for the ODNPs to selectively hybridize with the target nucleic acid, or the complement thereof, such that the ODNPs are able to function as primers for extension and/or amplification using the target nucleic acid, or the complement thereof, as a template. Generally, each ODNP contains at least 6, preferably 8, more preferably 10, most preferably 12, 14, or 16 nucleotides that are exactly complementary to the target nucleic acid or the complement thereof. Because each ODNP of the fourth ODNP pair hybridizes to a target nucleic acid, or the complement thereof, at a location 3′ to the defined position in the target or the complementary position in the complement of the target, the resulting extension and/or amplification products from the fourth ODNP pair incorporates the nucleotide to be identified at the defined position or the complement thereof.

Each ODNP in the fourth ODNP pair of the present invention further comprises a partial sequence of one strand of an interrupted restriction endonuclease recognition sequence (IRERS), but not a complete sequence of that strand of the

IRERS, located 3′ to, and preferably located directly 3′ to, the nucleic acid sequence described above (*i.e.*, the sequence complementary to the target nucleic acid or the complement thereof). As described in more detail below, a complete IRERS is a double-stranded nucleotide sequence comprising a first constant recognition sequence (CRS) and a second CRS linked with a variable recognition sequence (VRS) (Figure 4). Generally, the first ODNP and the second ODNP comprise the first CRS of the first strand of the IRERS and the second CRS of the second strand of the IRERS, respectively. In addition, the first ODNP and the second ODNP are so spaced that (1) the extension and/or amplification product with the fourth ODNP pair as primers and the target nucleic acid as a template contains a complete IRERS and (2) the nucleic acid to be identified is incorporated within the VRS. In other words, the number of nucleotides between the first and the second CRS is the exact number of nucleotides in the VRS so that the extension and/or amplification product from the fourth ODNP pair can be digested by a RE that recognizes the complete IRERS. The partial IRERS in each ODNP may or may not be complementary to the target nucleic acid.

In a preferred embodiment, each ODNP of the fourth ODNP pair further contains one or more nucleotides complementary to the target nucleic acid or the complement thereof ("the second region of the target nucleic acid" and "the second region of the complement," respectively) at a location 3′ to, or preferably the 3′ terminus of, the CRS. Such nucleotides are a portion of the VRS (Figures 3 and 4). The number of nucleotides between the first and second regions of the target nucleic acid or the complement thereof may be larger or smaller, but preferably are equal to, the number of nucleotides of the ODNPs located between the corresponding first and second regions that are complementary to the target nucleic acids or the complement thereof. In addition, in one aspect the first ODNP further contains a sequence of a sense strand of a NERS at a location 5′ to the first CRS of the first strand of an IRERS. Alternatively, the second ODNP further contains a sequence of a sense strand of a NERS at a location 5′ to the second CRS of the second strand of the IRERS. The presence of such a NERS allows the production of short single-stranded oligonucleotides upon digestion of extension/amplification products from the ODNPs using a NE that recognizes the NERS and a RE that recognizes the IRERS as described below.

General techniques for designing sequence-specific primers are well known. For instance, such techniques are described in books, such as *PCR Protocols: Current Methods and Application* edited by Bruce A. White, 1993; *PCR Primer: A Laboratory Manual* edited by Carl W. Dieffenbach and Gabriela S. Dveksler, 1995;

PCR (Basics: From Background to Bench) by McPherson et al.; *PCR Applications: Protocols for Functional Genomics* edited by Michael A. Innis, 1999; *PCR: Introduction to Biotechniques Series* by Newton and Graham, 1997; *PCR Protocols: A Guide to Methods and Applications* by Gelfand et al., 1990; *PCR Strategies* by Michael A. Innis; *PCR Technology: Current Innovations*, by Griffin and Griffin, 1994; and *PCR: Essential Techniques*, edited by J. F. Burke. In addition, software programs for designing primers are also available, including Primer Master (*see*, Proutski and Holmes, "Primer Master: A new program for the design and analysis of PCR primers" *Comput. Appl. Biosci. 12*: 253-5, 1996) and OLIGO Primer Analysis Software from Molecular Biology Insights, Inc. (Cascade, CO, USA). The above reference books and description of software programs are incorporated herein by reference in their entireties.

ODNPs according to the invention can be synthesized by any method known in the art for oligonucleotide synthesis such as methods disclosed in U.S. Patent Nos. 6,166,198, 6,043,353, 6,040,439, and 5,945,524 (incorporated herein in their entireties by reference). For instance, solid phase oligonucleotide synthesis can be performed by sequentially linking 5′ blocked nucleotides to a nascent oligonucleotide attached to a resin, followed by oxidizing and unblocking to form phosphate diester linkages. ODNPs of the present invention may then be isolated. The term "isolated" as used herein refers to a molecule that is substantially free of undesired contaminants, such as molecules having other sequences.

ii. Immobilization of Oligonucleotide Primers

In certain embodiments, the ODNPs of the present invention may be immobilized. Generally, an ODNP may be immobilized to a substrate in the following two ways: (1) synthesizing the ODNP directly on the substrate (often termed "*in situ* synthesis"), or (2) synthesizing the ODNP separately and then position and bind it to the substrate (sometimes termed "post-synthetic attachment"). For *in situ* synthesis, the primary technology is photolithography. Briefly, the technology involves modifying the surface of a solid support with photolabile groups that protect, for example, oxygen atoms bound to the substrate through linking elements. This array of protected hydroxyl groups is illuminated through a photolithographic mask, producing reactive hydroxyl groups in the illuminated areas. A 3'-*O*-phosphoramidite-activated deoxynucleoside protected at the 5'-hydroxyl with the same photolabile group is then presented to the surface and coupling occurs through the hydroxyl group at illuminated areas. Following further chemical reactions, the substrate is rinsed and its surface is illuminated through a second mask to expose additional hydroxyl groups for coupling.

A second 5'-protected, 3'-*O*-phosphoramidite-activated deoxynucleoside is present to the surface. The selective photo-de-protection and coupling cycles are repeated until the desired set of products is obtained. Detailed description of using photolithography in array fabrication may be found in the following patents or published patent applications: U.S. Patent Nos. 5,143,854; 5,424,186; 5,856,101; 5,593,839; 5,908,926; 5,737,257; and Published PCT Patent Application Nos. WO99/40105; WO99/60156; WO00/35931.

The post-synthetic attachment approach requires a methodology for attaching pre-existing oligonucleotides to a substrate. One method uses the biotin-streptavidin interaction. Briefly, it is well known that biotin and streptavidin form a non-covalent, but very strong, interaction that may be considered equivalent in strength to a covalent bond. Alternatively, one may covalently bind pre-synthesized oligonucleotide to a substrate. For example, carbodiimides are commonly used in three different approaches to couple DNA to solid supports. In one approach, the support is coated with hydrazide groups that are then treated with carbodiimide and carboxy-modified oligonucleotide. Alternatively, a substrate with multiple carboxylic acid groups may be treated with an amino-modified oligonucleotide and carbodiimide. Epoxide-based chemistries are also used with amine modified oligonucleotides. Detailed descriptions of methods for attaching pre-existing oligonucleotides to a substrate may be found in the following references: U.S. Patent Nos. 6,030,782; 5,760,130; 5,919,626; published PCT Patent Application No. WO00/40593; Stimpson et al. *Proc. Natl. Acad. Sci.* 92:6379-6383 (1995); Beattie et al. *Clin. Chem.* 41:700-706 (1995); Lamture et al. *Nucleic Acids Res.* 22:2121-2125 (1994); Chrisey et al. *Nucleic Acids Res.* 24:3031-3039 (1996); and Holmstrom et al., *Anal. Biochem.* 209:278-283 (1993).

The primary post-synthetic attachment technologies include ink jetting and mechanical spotting. Ink jetting involves the dispensing of oligonucleotides using a dispenser derived from the ink-jet printing industry. The oligonucleotides are withdrawn from the source plate up into the print head and then moved to a location above the substrate. The oligonucleotides are then forced through a small orifice, causing the ejection of a droplet from the print head onto the surface of the substrate. Detailed description of using ink jetting in array fabrication may be found in the following patents: U.S. Patent Nos: 5,700,637; 6,054,270; 5,658,802; 5,958,342; 6,136,962 and 6,001,309.

Mechanical spotting involves the use of rigid pins. The pins are dipped into an oligonucleotide solution, thereby transferring a small volume of the solution

onto the tip of the pins. Touching the pin tips onto the substrate leaves spots, the diameters of which are determined by the surface energies of the pins, the oligonucleotide solution, and the substrate. Mechanical spotting may be used to spot multiple arrays with a single oligonucleotide loading. Detailed description of using mechanical spotting in array fabrication may be found in the following patents or published patent applications: U.S. Patent Nos. 6,054,270; 6,040,193; 5,429,807; 5,807,522; 6,110,426; 6,063,339; 6,101,946; and published PCT Patent Application Nos. WO99/36760; 99/05308; 00/01859; 00/01798.

One of ordinary skill in the art would appreciate that besides the techniques described above, other methods may also be used in immobilizing ODNPs to a substrate. Descriptions of such methods can be found in, but are not limited to, the following patent or published patent applications: U.S. Patent Nos. 5,677,195; 6,030,782; 5,760,130; 5,919,626; and published PCT Patent Application Nos. WO98/01221; WO99/41007; WO99/42813; WO99/43688; WO99/63385; WO00/40593; WO99/19341; WO00/07022. These patents and patent applications, as well as all the references (including patents, patent applications, and journal articles) cited above in this section (*i.e.*, Immobolization of ODNPs) are incorporated herein by reference in their entirety.

iii. Hybridization

Methods, kits and compositions of the present invention may involve or include ODNPs that are hybridized to the target nucleic acid, where the ODNP facilitates the production and/or amplification of a defined nucleotide locus within the target nucleic acid. The ODNPs and target nucleic acid are thus preferably combined under base-pairing conditions which allow for hybridization and/or amplification. Selection of suitable nucleic acid hybridization and/or amplification conditions is within the skill of one of ordinary skill in the art, and may be assisted by reference to, for example, the following laboratory research manuals: Sambrook et al., "Molecular Cloning" (Cold Spring Harbor Press, 1989) and Ausubel et al., "Short Protocols in Molecular Biology" (1999) (incorporated herein by reference in their entirety).

Depending on the application envisioned, the artisan may vary conditions of hybridization to achieve desired degrees of selectivity of an ODNP towards a target sequence. For applications requiring high selectivity, relatively stringent conditions may be employed to form the hybrids, such as low salt and/or high temperature conditions (*e.g.*, from about 0.02 M to about 0.15 M salt at temperatures of

from about 50°C to about 70°C). Such selective conditions are relatively intolerant of large mismatches between the ODNP and the target nucleic acid.

Alternatively, hybridization of the ODNPs may be achieved under moderately stringent conditions such as, for example, in 10 mM Tris, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$ at 60°C, which conditions permit the hybridization of an ODNP comprising nucleotide mismatches with the target nucleic acid. The design of alternative hybridization conditions is well within the expertise of the skilled artisan.

iv. Nucleic Acid Extension/Amplification

To obtain template nucleic acid that comprises both a portion of a target nucleic acid containing a genetic variation at a defined position, and various combinations of recognition sequences (*e.g.*, a NERS and a RERS (including an IRERS), two NERSs or two hemimodified RERSs), a pair of primers is hybridized to the target nucleic acid, and each primer of the ODNP pair is extended using various methodologies known in the art, such as the polymerase chain reaction (PCR) and modified ligase chain reaction (LCR). Generally, at least three runs of extension reactions from the ODNP pairs described above need be carried out. Briefly, the first run of extension is for the first primer having a sequence of a sense strand of an enzyme recognition sequence (ERS) ("the first ERS") (*e.g.*, hemimodifed RERS, NERS) to incorporate the genetic variation of the target nucleic acid into the first extension product. The second primer, optionally having an additional sequence of one strand of another ERS ("the second ERS"), which may or may not be the same as the first ERS, then hybridizes to and extends using the first extension product as a template and thereby incorporates the complement of the genetic variation and the complement of the first ERS into a second extension product. An unextended first primer then hybridizes to and extends using the second extension product as a template, to thereby form, in combination with the second extension product, a double-stranded nucleic acid fragment that contains the genetic variation and the complement thereof, as well as two complete ERSs.

While three runs of extension reactions are sufficient to produce a fragment containing a genetic variation of a target nucleic acid and two ERSs, preferably, more than three extension reactions are conducted to amplify the template. As one of ordinary skill in the art will appreciate, in the subsequent runs of extension, the first primer can hybridize to and extend using any of the target nucleic acid, the second extension product, and the complement of the third extension product as a template. Similarly, in the subsequent runs of extension, the second primer can

hybridize to and extend using either the first extension product or the third extension product as a template. However, because the third extension product and the complement thereof are shorter than any of the target nucleic acid, the first extension product and the second extension product, they are the preferred templates for subsequent extension reactions from either the first or the second ODNPs. This is because the extension efficiency with a short fragment as a template is higher than that with a large fragment as a template. With an increase in the number of extension reactions, the double stranded fragment containing both a genetic variation and the first ERS and the optional second ERS accumulates more quickly than other molecules in the reaction mixture. Such accumulation increases the sensitivity of subsequent characterization of a single-stranded nucleic acid amplified using the above template nucleic acid as a template.

The extension/amplification reaction can be carried out by any method known in the art. For instance, PCR methods described in U.S. Patent Nos. 4,683,195, 4,683,202, and 4,800,159 may be used. Other PCR methods may also be used as described in books, *e.g.*, Gelfand et al., "PCR Protocols: A Guide to Methods and Application" (1990); Burke (ed), "PCR: Essential Techniques"; and McPherson et al. "PCR (Basic: From Background to Bench)". Each of the above references is incorporated herein by reference in its entirety. Briefly, in PCR, two ODNPs are prepared that are complementary to regions on opposite complementary strands of the target nucleic acid sequence. An excess of deoxyribonucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* or *Pfu* polymerase). If the target nucleic acid sequence is present in a sample, the ODNPs will bind to the target and the polymerase will cause the ODNPs to be extended along the target nucleic acid sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended ODNPs will dissociate from the target to form reaction products, excess ODNPs will bind to the target and to the reaction product and the process is repeated.

Exemplary PCR conditions according to the present invention may include, but are not limited to, the following: 100 µl PCR reactions comprise 100 ng target nucleic acid; 0.5 µM of each of the first ODNP and the second ODNP; 10 mM Tris, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$; 200 µM each dNTP; 4 units Taq™ DNA Polymerase (Boehringer Mannheim; Indianapolis, IN), and 880 ng TaqStart™ Antibody (Clontech, Palo Alto, CA). Exemplary thermocycling conditions may be as follows: 94°C for 5 minutes initial denaturation; 45 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute; final extension at 72°C for 5 minutes. Exemplary nucleic

acid polymerases may include one of the thermostable DNA polymerases that are readily available in the art such as, *e.g.*, Taq™ , Vent™ or PFU™. Depending on the particular application contemplated, it may be preferred to employ one of the nucleic acid polymerases having a defective 3′ to 5′ exonuclease activity.

An alternative way to make and/or amplify a fragment containing a nucleotide to be identified is by a modified ligase chain reaction, referred to herein as the gap-LCR (Abravaya et al., *Nucleic Acids Res. 23*:675-682 (1995)). Briefly, in the presence of the target sequence, each pair of the set will bind to the target, or the complement thereof, located 5′ and 3′ of (on either side of) the nucleotide of interest in the target nucleic acid. In the presence of a polymerase and a ligase, the gap between the two ODNPs of each pair will be filled in and the ODNPs of each pair ligated to form a single unit. By temperature cycling, as in PCR, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess ODNP pairs. Thus, LCR uses both a nucleic acid polymerase enzyme and a nucleic acid ligase enzyme to drive the reaction. Exemplary nucleic acid polymerases may include one of the thermostable DNA polymerases that are readily available in the art such as, *e.g.*, Taq™, Vent™ or PFU™. Exemplary nucleic acid ligases may include T4 DNA ligase, or the thermostable Tsc or Pfu DNA ligases. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding ODNP pairs to a target sequence.

Exemplary gap-LCR conditions may include, but are not limited to, the following: 50 µl LCR reactions comprise 500 ng DNA; a buffer containing 50 mM EPPS, pH 7.8, 30 mM $MgCl_2$, 20 mM K^+, 10 µM NAD, 1-10 µM gap filling nucleotides, 30 nM each oligonucleotide primer, 1 U *Thermus flavus* DNA polymerase, lacking 3′→5′ exonuclease activity (MBR, Milwaukee, WI), and 5000 U *T. thermophilus* DNA ligase (Abbott Laboratories). Cycling conditions may consist of a 30 second incubation at 85°C and a 30 second incubation at 60°C for 25 cycles and may be carried out in a standard PCR machine such as a Perkin Elmer 9600 thermocycler.

To prevent the template nucleic acid that contains two RERSs from being cleaved in both strands, extension/amplification from the third ODNP pair may be performed in the presence of a modified deoxyribonucleoside triphosphate (*e.g.*, α-thio deoxyribonucleoside triphosphate). The incorporation of the modified deoxynucleotide into one strand of the RERS blocks the cleavage by the RE that recognizes the RERS in the modified strand. Consequently, only the sense strand of the RERS that is a portion of the ODNP pair may be cleaved.

Any modified deoxyribonucleoside triphosphate that contributes to the inhibition of cleavage of one of the two DNA strands comprising the RERS may be used in the present invention. Exemplary modified deoxyribonucleoside triphosphates include, but are not limited to, 2′-deoxycytidine 5′-O-(1-thiotriphosphate) [i.e., dCTP(.alpha.S)], 2′-deoxyguanosine 5′-O-(1-thiotriphosphate), thymidine-5′-O-(1-thiotriphosphate), 2′-deoxycytidine 5′-O(1-thiotriphosphate), 2′-deoxyuridine 5′-triphosphate, 5-methyldeoxycytidine 5′-triphosphate, and 7-deaza-2′-deoxyguanosine 5′-triphosphate.

In certain embodiments, more than one modified deoxyribonucleoside triphosphate may be used. In some other embodiments, post-synthesis modification of certain appropriate nucleotide(s) provides an alternative to the use of a modified deoxyribonucleoside triphosphate.

In some embodiments where the target nucleic acid is a RNA molecule (*e.g.*, mRNA), a DNA polymerase using a RNA molecule (*e.g.*, a reverse transcriptase) as a template may be required. The single-stranded DNA molecule synthesized by the DNA polymerase may be further used as a template to synthesize and/or amplified double-stranded DNA molecules in the presence of a DNA polymerase using a DNA molecule as a template.

The present invention may also be used for multiplex determination of genetic variations in target nucleic acids at defined positions. In such an application, multiple target nucleic acids and ODNP pairs specific to these target nucleic acids may be combined together under amplification conditions. This combination reduces the expenses associated with reagents and DNA polymerase required for the amplification. The resulting amplification products will subsequently be used as templates for amplification of single-stranded nucleic acid molecules, and therefore the amplification of the templates generally need not be extensive since the secondary amplification of the single-stranded nucleic acid molecules increases the amount of sample that is detected by the subsequent detection method.

In some embodiments, the template nucleic acids of the present invention may be obtained using the same DNA polymerase as that used in the amplification of single-stranded nucleic acid molecules described in detail below.

v. Restriction Endonucleases and Digestion Conditions

The amplified template nucleic acids as described above, if containing a RERS (not including a hemimodified RERS), are typically digested by a RE that recognizes the RERS. The term "restriction endonuclease" (RE) refers to the class of

nucleases that recognize unique double-stranded nucleic acid sequences and that generate a cleavage in the double-stranded nucleic acid, resulting in either blunt double-stranded ends, or single-stranded ends with either a 5′ or a 3′ overhang. REs are usually classified into three types: type I, type II and type III, among which type II are most commonly used in molecular biological manipulation. The RE that may be used in the present invention includes any type II RE, such as (1) the vast majority of type II REs that recognize specific sequences that are four, five, or six nucleotides in length and display twofold symmetry (*e.g.*, EcoR I, BamH I), (2) type IIs REs (TSREs), and (3) interrupted REs (IREs). A "type IIs restriction endonuclease" is a type II restriction endonuclease that cleaves outside its recognition sequence. Exemplary TSREs include, but are not limited to, Fok I, BsgI and BpmI. An "interrupted restriction endonuclease" is a type II restriction endonuclease that recognizes an interrupted restriction endonuclease recognition sequence (IRERS). An IRERS is defined as a restriction endonuclease recognition sequence that is comprised of a "first constant recognition sequence (CRS)," a "second CRS," and a "variable recognition sequence (VRS)" that links the first and second CRSs (Figure 4). All the three components of an IRERs are double-stranded. According to the present invention, "first CRS" is defined as that region of the IRERS that contains the constant (not variable) nucleotides of the IRERS immediately adjunct to the VRE of the IRERS at one side, wherein "second CRS" is defined as that region of the IRERS that contains the constant (not variable) nucleotides of the IRERS immediately adjunct to the VRE of the IRERS at the other side. According to the present invention, the "VRE" is defined as the stretch of one or more variable nucleotides that are located between the first and second CRSs. Exemplary IREs include, but are not limited to, Bsl I, EcoN I, Dra III and Dde I.

The RE useful for the present invention may be purchased from various companies such as, *e.g.*, New England Biolabs Inc. (Beverly, MA; www.neb.com); Stratagene (La Jolla, CA; www.stratagene.com), Promega (Madison, WI: www.promega.com), and Clontech (Palo Alto, CA; www.clontech.com). Non-commercially available restriction enzymes may be isolated and/or purified based on the teaching available in the art. For instances, the following articles describe the isolation and/or purification of several non-commercially available restriction enzymes suitable for the present invention and are incorporated herein by reference in their entireties by reference: for restriction enzyme ApaB I, Grones and Turna, *Biochim. Biophys. Acta 1162*:323-325 (1993), Grones and Turna, *Biologia (Bratisl) 46*:1103-1108 (1991); for EcoH I, Glatman et al., *Mol. Gen. Mikrobiol. Virusol. 3*:32 (1990); for Fmu I, Rebentish et al. *Biotekhnologiya 3*:15-16 (1994); for HpyB II, *FEMS Microbiol. Lett.*

179:175-180 (1999); for Sse8647 I, Nomura et al., European Patent Application No. 0698663 A1, Ishino et al., *Nucleic Acids Res. 23*: 742-744 (1995); for Unb I, Kawalec et al., *Acta Biochim. Pol. 44*:849-852 (1997); for VpaK11A I, Miyahara et al. *J. Food. Hyg. Sci. Japan 35*:605-609 (1994).

Descriptions of conditions for the storage and the use of restriction endonucleases used according to the present invention are readily available in the art and are found, for example, in laboratory manuals such as Sambrook et al., *supra* and Ausubel et al., *supra*. Briefly, the number of units of RE added to a reaction may be calculated and adjusted according to the varying cleavage rates of nucleic acid substrates. 1 unit of restriction endonuclease will digest 1 μg of substrate nucleic acid in a 50 μl reaction in 60 minutes. Generally, fragments (*e.g.*, amplicons) may require more than 1 unit/μg to be cleaved completely. The restriction enzyme buffer is typically used at 1X concentration in the reaction. Some restriction endonucleases require bovine serum albumin (BSA) (usually used at a final concentration of 100 μg/ml for optimal activity). Restriction endonucleases that do not require bovine serum albumin (BSA) for optimal activity are not adversely affected if BSA is present in the reaction.

Most restriction endonucleases are stable when stored at -20°C in the recommended storage buffer. Exposure to temperatures above -20°C should be minimized whenever possible. All restriction endonucleases should be kept on ice when not otherwise being stored in the freezer. Enzymes should always be the last component added to a reaction.

The recommended incubation temperature for most restriction endonucleases is about 37°C. Restriction endonucleases isolated from thermophilic bacteria require higher incubation temperatures, typically ranging from 50°C to 65°C. Incubation time may often be shortened if an excess of restriction endonuclease is added to the reaction. Longer incubation times are often used to allow a reaction to proceed to completion with fewer units of restriction endonuclease.

When two REs are used simultaneously, reaction conditions including incubation temperature and a reaction buffer need to be optimized to be suitable for both enzymes. However, if no digestion condition can be found that is suitable for both of two different REs, digestion with one RE may be carried out prior to digestion with the other RE. In the case that double digestion is carried out using a NE and a RE, the resulting products from an extension/amplification product using the primer pair of the present invention (*e.g.*, the fourth ODNP pair) are two double-stranded oligonucleotides and a single-stranded oligonucleotide (*see*, *e.g.*, Figure 3).

c. Amplification of Single-Stranded Nucleic Acids

According to the present invention, template nucleic acids that comprise a NARS, a NS, and a genetic variation 3′ to the NS are used as templates for amplifying single-stranded nucleic acids that comprise the genetic variation or the complement thereof. The amplification is performed by repetitively nicking the template and then extending from the nicking site to reproduce the template. Thus, the template nucleic acid is first nicked by a NA that recognizes the NARS(s) of the template to thereby produce a 3′ terminus at the NS. A DNA polymerase that has a strand displacement ability, but lacks a 5′→3′ exonuclease activity, is then used to extend the nicked template from the 3′ terminus at the NS, thus displacing the downstream single stranded nucleic acid. This extension creates a new nicking site for the NA. Repetition of the nicking-extending process amplifies a single-stranded nucleic acid having a 5′ terminus at the NS created by the NA. Because the genetic variation is located 3′ to the NS in the template, the genetic variation is incorporated into the amplified single-stranded nucleic acid. Thus, subsequent characterization of the amplified single-stranded nucleic acid allows for identification of the genetic variation that originated from a target nucleic acid.

In certain embodiments of the present invention, such as where the template nucleic acid comprises both a sequence of a sense strand of NERS and a sequence of one strand of a RERS (*see*, *e.g.*, Figure 1), the nicking of the template nucleic acid after being digested by a RE that recognizes the RERS produces a short nucleic acid fragment that readily dissociates from the remaining portion of the template nucleic acid. In these embodiments, a DNA polymerase that does not have a strand displacement ability and/or that is not 5′ to 3′ exonuclease deficient may be used to amplify the short single-stranded nucleic acid fragment. Such a DNA polymerase may also be used where the template nucleic acid contains two NARSs (*e.g.*, two NERSs or two hemimodified RERSs or one NERS with one hemimodified RERS) sufficiently close to each other so that (1) the initial nicking of the template nucleic acid by NA(s) that recognizes the two NARSs produces two partially double-stranded nucleic acid fragments that dissociate from each other under the nicking reaction condition, and (2) the subsequent nicking of the extension product of the nicked template nucleic acid produces a short nucleic acid fragment that readily dissociates from the other portion of the extension product (*see*, *e.g.*, Figures 2 and 5). In these embodiments where the strand replacement activity of a DNA polymerase is not required for amplifying single stranded nucleic acid fragments, the amplification is performed using the single stranded portion of digested or nicked template nucleic acids

as the templates, which is distinct from traditional strand displacement amplification wherein the templates for the amplification of single stranded nucleic acid fragments are double stranded.

i. Nicking of Template Nucleic Acids

The template nucleic acid of the present invention can be nicked by any nicking agent that recognizes the nicking agent recognition sequences (NARSs) in the template nucleic acid. Preferably, the nicking agent is an enzyme that recognizes a specific sequence of a double-stranded nucleic acid and cleaves only one strand of the nucleic acid. Such an enzyme can be a nicking endonuclease that recognizes a specific sequence that consists of native nucleotides or a restriction endonuclease that recognizes a hemimodified recognition sequence.

A nicking endonuclease may or may not have a nicking site that overlaps with its recognition sequence. An exemplary NE that nicks outside its recognition sequence is N.BstNB I, which recognizes a unique nucleic acid sequence composed of 5′-GAGTC-3′, but nicks four nucleotides beyond the 3′ terminus of the recognition sequence. The recognition sequence and the nicking site of N.BstNB I are shown below with "▼" to indicate the cleavage site where the letter N denotes any nucleotide:

```
          ▼
5′-GAGTCNNNNN-3′
3′-CTCAGNNNNN-5′
```

N.BstNB I may be prepared and isolated as described in U.S. Pat. No. 6,191,267. Buffers and conditions for using this nicking endonuclease are also described in the '267 patent. An additional exemplary NE that nicks outside its recognition sequence is N.AlwI, which recognizes the following double-stranded recognition sequence:

```
          ▼
5′-GGATCNNNNN-3′
3′-CCTAGNNNNN-5′
```

The nicking site of N.AlwI is also indicated by the symbol "▼". Both NEs are available from New England Biolabs (NEB). N.AlwI may also be prepared by mutating a type IIs RE AlwI as described in Xu *et al.* (*Proc. Natl. Acad. Sci. USA 98*:12990-5, 2001).

Exemplary NEs that nick within their NERSs include N.BbvCI-a and N.BbvCI-b. The recognition sequences for the two NEs and the NSs (indicated by the symbol "▼") are shown as follows:

N.BbvCI-a

▼
5′-CCTCAGC-3′
3′-GGAGTCG-5′

N.BbvCI-b

▼
5′-GCTGAGG-3′
3′-CGACTCC-5′

Both NEs are available from NEB.

Additional exemplary nicking endonucleases include, without limitation, N.BstSE I (Abdurashitov *et al.*, *Mol. Biol. (Mosk) 30*: 1261-7, 1996), an engineered EcoR V (Stahl *et al.*, *Proc. Natl. Acad. Sci. USA 93*: 6175-80, 1996), an engineered Fok I (Kim *et al.*, *Gene 203*: 43-49, 1997), endonuclease V from *Thermotoga maritima* (Huang *et al.*, *Biochem. 40*: 8738-48, 2001), Cvi Nickases (*e.g.*, CviNY2A, CviNYSI, Megabase Research Products, Lincoln, Nebraska) (Zhang *et al.*, *Virology 240*: 366-75, 1998; Nelson *et al.*, *Biol. Chem. 379*: 423-8, 1998; Xia *et al.*, *Nucleic Acids Res. 16*: 9477-87, 1988), and an engineered Mly I (*i.e.*, N.Mly I) (Besnier and Kong, *EMBO Reports 2*: 782-6, 2001). Additional NEs may be obtained by engineering other restriction endonuclease, especially type IIs restriction endonucleases, using methods similar to those for engineering EcoR V, AlwI, Fok I and/or Mly I.

The nicking of the template nucleic acid of the present invention can be performed by any restriction endonuclease that nicks a double-stranded nucleic acid at its hemimodified recognition sequences. A "hemimodified recognition sequence" is a double-stranded recognition sequence of a RE in which one strand of the recognition sequence contains at least one derivatized nucleotide (*e.g.*, α-thio deoxynucleotide) that prevents cleavage of one strand (*i.e.*, the strand that contains the derivatized nucleotide or the other strand that does not contain the derivatized nucleotide) by the restriction endonuclease. Exemplary REs that nick their double-stranded hemimodified recognition sequences include, but are not limited to, Ava I, Bsl I, BsmA I, BsoB I, Bsr I, BstN I, BstO I, Fnu4H I, Hinc II, Hind II and Nci I. Additional REs that nick a hemimodified recognition sequence may be screened by the strand protection assays

described in U.S. Patent No. 5,631,147, incorporated herein by reference in its entirety. Briefly, a single stranded template comprising the restriction endonuclease/cleavage site and a primer complementary to a portion of the template other than the recognition/cleavage site are synthesized. The template and the primer are then labeled, *e.g.*, radiolabeled. The primer and the template are annealed and modified dNTPs are incorporated by extension of the primer, resulting in a fully double-stranded molecule containing a hemimodified restriction endonuclease recognition/cleavage site. This molecule is then combined with the restriction endonuclease under conditions suitable for double-stranded cleavage. The digestion products are then analyzed by denaturing electrophoresis to determine, by the size of the fragments generated, whether the recognition/cleavage site was nicked, cleaved or uncut. The size of the digestion products is also used to determine which of the two strands of the recognition/cleavage site (*i.e.*, modified or unmodified) was protected from cleavage.

REs that nick a hemimodified restriction endonuclease recognition sequence may be purchased from various companies such as, *e.g.*, New England Biolabs Inc. (Beverly, MA; www.neb.com); Stratagene (La Jolla, CA; www.stratagene.com), Promega (Madison, WI: www.promega.com), and Clontech (Palo Alto, CA; www.clontech.com). Non-commercially available restriction enzymes may be isolated and/or purified based on the teaching available in the art. Conditions for using the restriction endonuclease for nicking a template nucleic acid may or may not be the same as those for cleaving a double-stranded nucleic acid. These conditions can be optimized for the nicking activities of the RE using ordinary skill in the art (*see*, *e.g.*, Walker, *PCR Methods Appl.* *3*:1-6, 1993, incorporated herein by reference in its entirety).

In certain embodiments, a nicking agent may recognize a nucleotide sequence in a DNA-RNA duplex and nicks in one strand of the duplex. In certain other embodiments, a nicking agent may recognize a nucleotide sequence in a double-stranded RNA and nicks in on strand of the RNA.

ii. Extension

The nicking of the template nucleic acid produces a 3′ terminus at the nicking site, from which extension may be performed in the presence of a DNA polymerase, using the unnicked strand as a template. When the DNA polymerase lacks a 5′→3′ exonuclease activity, but has a strand displacement activity, the extension of the nicked template nucleic acid at the nicking site displaces the downstream single-

stranded nucleic acid fragment. Such displacement allows the accumulation, thus amplification, of the single-stranded nucleic acid fragment.

Any DNA polymerase that is 5′→3′ exonuclease deficient but has a strand displacement activity may be used to extend from a nicked template nucleic acid and to subsequently amplify a single-stranded nucleic acid in the continuous presence of a nicking agent. Such DNA polymerases include, but are not limited to, exo⁻ Deep Vent, exo⁻ Bst, exo⁻ Pfu, and exo⁻ Bca. Additional DNA polymerase useful in the present invention may be screened for or created by the methods described in U.S. Patent No. 5,631,147, incorporated herein by reference in its entirety. Briefly, the polymerase screening system is an extension assay that tests the ability of the polymerase to displace a downstream strand initiating at a nicking site in a double-stranded template. It also tests for the presence, or the absence, of 5′→3′ exonuclease activity. 5′→3′ exonuclease activity can be inactivated by routine methods known in the art (*e.g.*, PCT published patent application WO 92/06200). One way to inactivate the exonuclease activity of a polymerase is to clone the gene for the polymerase, identify the portion of the gene that encodes the protein domain responsible for exonuclease activity, and inactivate it by *in vitro* mutagenesis. Alternatively, exonuclease activity may be inactivated by treating the polymerase with protease and thereafter isolating fragments that exhibit only the desired polymerization and displacing activities. The strand displacement activity of a DNA polymerase may be further enhanced by the presence of a strand displacement facilitator as described below.

A "strand displacement facilitator" is any compound or composition that facilitates strand displacement during nucleic acid extensions from a 3′ terminus at a nicking site catalyzed by a DNA polymerase. Stand displacement facilitators useful in the present invention include, but are not limited to, BMRF1 polymerase accessory subunit (Tsurumi *et al.*, *J. Virology 67*: 7648-53, 1993), adenovirus DNA-binding protein (Zijderveld and van der Vliet, *J. Virology 68*: 1158-64, 1994), herpes simplex viral protein ICP8 (Boehmer and Lehman, *J. Virology 67*: 711-5, 1993; Skaliter and Lehman, *Proc. Natl. Acad. Sci. USA 91*: 10665-9, 1994), single-stranded DNA binding protein (Rigler and Romano, *J. Biol. Chem.* 270: 8910-9, 1995), phage T4 gene 32 protein (Villemain and Giedroc, *Biochemistry 35*: 14395-4404, 1996), calf thymus helicase (Siegel *et al.*, *J. Biol. Chem. 267*: 13629-35, 1992) and trehalose.

In addition, as described above, in certain embodiments wherein the nicking of a template nucleic acid or a digestion product thereof, or, the re-nicking of an extension product of a nicked template nucleic acid by a nicking agent produces a short

single-stranded nucleic acid fragment that readily dissociates from the remaining portion of the template nucleic acid, a DNA polymerase that does not have a strand displacement activity and/or is not 5′ to 3′ exonuclease deficient may also be used. DNA polymerases that do not have a strand displacement activity include, but are not limited to, exo⁻ Vent, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, and Phi29 DNA polymerase.

The short single stranded nucleic acid fragments may contain, for example, at most 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleotides. With the increase of the lengths of such nucleic acid fragments, the interaction between these fragments and their corresponding complementary strands of the template nucleic acids via hydrogen bonds may become sufficiently strong that they are unable to dissociate from their complementary strands. In such a situation, the amplification of the single-stranded nucleic acid fragments requires a DNA polymerase having a strand displacement ability, the presence of a strand displacement facilitator, or both. The maximum length of the amplified single stranded nucleic acid fragments that is capable of readily dissociating from the other portion of the template nucleic acids may also depend on the nicking-extension reaction conditions. For example, the maximum length of a single stranded nucleic acid fragment amplified under a relatively high nicking-extension reaction temperature (*e.g.*, 55°C) that can readily dissociate from its complementary strand is larger than that amplified under a relatively low reaction temperature (*e.g.*, 37°C or 22°C).

Additional exemplary DNA polymerases useful in the present invention include, but are not limited to, phage M2 DNA polymerase (Matsumoto *et al.*, *Gene 84*: 247, 1989), phage PhiPRD1 DNA polymerase (Jung *et al.*, *Proc. Natl. Acad. Sci. USA 84*: 8287, 1987), T5 DNA polymerase (Chatterjee *et al.*, *Gene 97*: 13-19, 1991), Sequenase (U.S. Biochemicals), PRD1 DNA polymerase (Zhu and Ito, *Biochim. Biophys. Acta. 1219*: 267-76, 1994), 9°N$_m$™ DNA polymerase (New England Biolabs) (Southworth *et al.*, *Proc. Natl. Acad. Sci. 93*: 5281-5, 1996; Rodriquez *et al.*, *J. Mol. Biol. 302*: 447-62, 2000), and T4 DNA polymerase holoenzyme (Kaboord and Benkovic, *Curr. Biol. 5*: 149-57, 1995).

In certain embodiments where a nicking agent nicks in the DNA strand of a RNA-DNA duplex, a RNA-dependent DNA polymerase may be used. In other embodiments where a nicking agent nicks in the RNA strand of a RNA-DNA duplex, a DNA-dependent DNA polymerase that extends from a DNA primer, such as Avian Myeloblastosis virus reverse transcriptase (Promega) may be used. In both instances, a target mRNA need not be reverse transcribed into cDNA and may be directly mixed

with a template nucleic acid molecule that is at least substantially complementary to the target mRNA.

iii. Amplification of Single-Stranded Nucleic Acids

The extension of a nicked nucleic acid from the 3′ terminus of the nicking site by a DNA polymerase produces a new nicking site. This new nicking site may be re-nicked by a nicking agent and re-extended in the presence of the DNA polymerase. Such a nicking-extension process may be repeated many times, resulting in the amplification of the single-stranded nucleic acid fragment downstream of the nicking site.

Preferably, both the nicking agent (*e.g.*, a NE or a RE) and the DNA polymerase are present in a nicking/extension reaction mixture and the mixture is optimized to be suitable for both the nicking agent and the DNA polymerase. For instance, if N.BstNB I is the nicking agent and exo⁻ Vent is the DNA polymerase, the nicking-extension buffer can be 0.5X N.BstNB I buffer and 1X DNA polymerase Buffer. Exemplary 1X N.BstNB I buffer may be 10 mM Tris-HCl, 10 mM $MgCl_2$, 150 mM KCl, and 1 mM dithiothreitol (pH 7.5 at 25°C). Exemplary 1X DNA polymerase buffer may be 10 mM KCl, 20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, and 0.1% Triton X-100.

Also preferably, nicking and extension reactions of the present invention are performed under isothermal conditions. As used herein, "isothermally" and "isothermal conditions" refer to a set of reaction conditions where the temperature of the reaction is kept essentially constant (*i.e.*, at the same temperature or within the same narrow temperature range wherein the difference between an upper temperature and a lower temperature is no more than 20°C) during the course of the amplification. An advantage of the amplification method of the present invention is that there is no need to cycle the temperature between an upper temperature and a lower temperature. Both the nicking and the extension reaction will work at the same temperature or within the same narrow temperature range. If the equipment used to maintain a temperature allows the temperature of the reaction mixture to vary by a few degrees, such a fluctuation is not detrimental to the amplification reaction. Exemplary temperatures for isothermal amplification include, but are not limited to, any temperature between 50°C to 70°C (*i.e.*, 51, 52, 53...69) or the temperature range between 50°C to 70°C, 55°C to 70°C, 60°C to 70°C, 65°C to 70°C, 50°C to 55°C, 50°C to 60°C, 50°C to 65°C. Many NAs and DNA polymerases are active at the above exemplary temperatures or within the above exemplary temperature ranges. For instance, both the nicking reaction using

N.BstNB I (New England Biolabs) and the extension reaction using exo⁻ Bst polymerases (BioRad) may be carried out at about 55°C. Other polymerases that are active between about 50°C and 70°C include, but are not limited to, exo⁻ Vent (New England Biolabs), exo⁻ Deep Vent (New England Biolabs), exo⁻ Pfu (Strategene), exo⁻ Bca (Panvera) and Sequencing Grade Taq (Promega).

As described above, in certain preferred embodiments, an identical sequence of one strand of a NARS is present in both members of an ODNP pair for synthesizing template nucleic acids. In these embodiments, two complementary single-stranded nucleic acid fragments are amplified in the presence of a NA that recognizes the NARS and a DNA polymerase. The detection of both of the amplified nucleic acid fragments by methods described in detail below increases the reliability and/or accuracy of the present method.

As described above, in certain embodiments, the amplified single-stranded nucleic acid may be relatively short and readily dissociate from the other portion of the template nucleic acid in the abosence of any strand displacement activity of a DNA polymerase or a strand displacement facilitator. The short length of the amplified single-stranded molecule may be advantageous because it increases amplification efficiencies and rates. In addition, it allows the use of a DNA polymerase that does not have a stand displacement activity or is not 5'→3' exonuclease deficient. It also facilitates the detection of A1 molecules in which A1 is used as an initial amplification primer via certain technologies such as mass spectrometric analysis.

The dissociation of an amplified single-stranded nucleic acid in the absence of a strand displacement activity or a strand displacement facilitator may occur under the circumstances where the melting temperature (Tm) of the duplex formed between the amplified single-stranded nucleic acid and its complementary strand is lower than that of the duplex formed between the nicked product that comprises the 3' terminus at the nicking site and its complementary strand, and where the amplification is performed at a temperature that is between the above two melting temperatures. Calculation of melting temperatures is known in the art. For this invention, a melting temperature of a nucleic acid duplex is calculated by the method of Baldino *et al.*, *Methods Enzymol. 168*: 761-777, 1989.

The present invention also provides methods for multiplex determination of genetic variations in target nucleic acids at defined positions. In such an application, multiple template nucleic acids may be combined together in a nicking-extension reaction. This combination reduces the expenses associated with reagents, nicking agents and DNA polymerases for the reaction. It also increases the throughput of

subsequent detection/characterization assays for the amplified single-stranded nucleic acid molecules.

d. Methodologies for Characterizing Amplified Single-Stranded Nucleic Acid Fragments

As discussed above, the present invention provides methodology whereby a single-stranded nucleic acid fragment that contains a genetic variation originating from a target nucleic acid is amplified. The subsequent characterization of the amplified single-stranded nucleic acid fragment allows one to determine the identity of the genetic variation in the target nucleic acid. Thus, the present invention transfers information about the genetic variation from a relatively large target nucleic acid into a relatively small nucleic acid fragment. This transfer facilitates the characterization of the genetic variation because a direct and complete characterization of the small nucleic acid fragment is often more feasible than adequate characterization of a large nucleic acid target.

According to the present invention, the characterization of a nucleic acid fragment (*i.e.*, an amplified single-stranded nucleic acid) can be done directly, that is, without the need to incorporate a tag or label into the fragment. Alternatively, in some embodiments, it may be advantageous to utilize one or more detectable labels. Both types of characterization methods are discussed in more detail below. In addition, the techniques for characterizing a nucleic acid fragment as described in detail below with respect to methods for characterizing target nucleic acids may also be used in methods for detecting a genetic variation.

i. Direct Characterization

The present invention transfers information about nucleotide(s) of interest (*i.e.*, genetic variations) from a relatively large target nucleic acid into a relatively small single-stranded nucleic acid fragment. Direct characterization of the small fragment may, in many instances, provide information about the nucleotide(s) of interest, and particularly the identity of the nucleotide of interest. For example, small nucleic acid fragments are amenable to direct detection by a variety of mass spectrometric methodologies (as discussed herein below) as well as by ultraviolet (UV) absorption.

In many instances according to the present invention, the complete nucleotide sequence, with the exception of a single nucleotide, will be known for the short nucleic acid fragment even before it is formed. The issue then becomes detecting

the nucleotide of interest over the "noise" created while concurrently detecting the other bases in the short nucleic acid fragment. However, if the identity of the other nucleotides is known and their signal in the detection method is known, then this signal can be subtracted from the overall signal for the fragment, to leave information about the nucleotide of interest. This approach is essentially adopted in using mass spectrometry to characterize the small nucleic acid fragment. Other suitable methods, as discussed in detail herein, include determining the mass-to-charge ratio of the small nucleic acid fragment(s), measuring fluorescence polarization and/or by quantifying ultraviolet (UV) absorption of the fragment(s).

In some instances, characterizing a small nucleic acid fragment may entail simply determining the size of the single-stranded fragment, and from this information the skilled artisan can deduce whether a target nucleic acid contains one or more mutations at a defined nucleotide locus. It will be apparent that the size of a single-strand fragment may be determined by numerous methods that are readily available in the art. Exemplary methods disclosed herein for measuring the size and/or molecular weight of a single-strand nucleic acid fragment, include, but are not limited to, fluorescence including fluorescence polarization (FP), mass spectrometry (MS), ultraviolet (UV) absorption, cleavable mass tags, TaqMan (homogeneous), fluorescence resonance energy transfer (FRET), colorimetric, luminescence and/or fluorescence methodologies employing substrates for horseradish peroxidase (HRP) and/or alkaline phosphatase (AP), as well as methods employing radioactivity.

In certain embodiments of the present invention, Mass Spectrometry (MS) is employed for characterizing a nucleic acid fragment, preferably a short nucleic acid fragment, comprising the nucleotide locus of interest from the target nucleic acid. MS may be particularly advantageous in those applications in which it is desirable to eliminate a fractionation step prior to detection. Alternatively, MS may also be employed in conjunction with a fractionation methodology, as discussed herein below, such as, for example, one of the liquid chromatography methodologies including HPLC and DHPLC. Typically, MS detection does not require the addition of a tag or label to the small nucleic acid fragment. Instead, the nucleic acid fragment can be adequately identified and/or characterized directly in the mass spectrometer.

As disclosed herein, MS may be particularly suitable for the detection of small nucleic acid fragments from as small as 1 nucleotide to as large as several hundred nucleotides. More preferably, MS is used to identify and/or characterize fragments of 1 to 50 nucleotides, still more preferably fragments of from 1 to 17 nucleotides.

Sensitivities may be achieved to at least to 1 Dalton. The smallest mass difference in nucleic acid bases is between adenine and thymine, which is 9 Daltons.

Particularly preferred MS methodologies employ Liquid Chromatography-Time-of-Flight Mass Spectrometry (LC-TOF-MS). LC-TOF-MS is composed of an orthogonal acceleration Time-of-Flight (TOF) MS detector for atmospheric pressure ionization (API) analysis using electrospray (ES) or atmospheric pressure chemical ionization (APCI). LC-TOF-MS provides high mass resolution (5000 FWHM), high mass measurement accuracy (to within 5 ppm) and very good sensitivity (ability to detect picomolar amount of DNA polymer) compared to scanning quadrupole instruments. TOF instruments are generally more sensitive than quadrupoles, but also are more expensive. In one embodiment, solid phase extraction (SPE) may be used in lieu of LC to purify nucleic acid fragments before subjecting the fragments to MS analysis.

LC-TOF-MS has a very efficient duty cycle since the current instruments can sequentially analyze one mass at a time while rejecting all others (this is referred to as single ion monitoring (SIM)). LC-TOF-MS samples and examines all of the ions passing into the TOF analyzer at the same time. This results in higher sensitivity and provides quantitative data which improves the sensitivity between 10 and 100 fold. Enhanced resolution (5000 FWHM) and mass measurement accuracy of better than 5 ppm imply that differences between nucleosides as small as 9 amu (Daltons) can be accurately measured. The TOF mass analyzer performs very high frequency sampling (10 spectra/sec) of all ions simultaneously across the full mass range of interest. The duty cycle of the LC-TOF-MS allows high sensitivity spectra to be recorded in quick succession making the instrument compatible with highly efficient separation techniques such as narrow bore LC, capillary chromatography and capillary electrochromatography (CEC). The ions are pulsed into the analyzer, effectively taking a "snapshot" of the ions present at any time.

In the first stage of ES or APCI, aerosol spray is directed perpendicularly past the sampling cone, which is displaced from the central axis of the instrument. Ions are extracted orthogonally from the spray into the sampling cone aperture leaving large droplets, nonvolatile materials, particulates and other unwanted components to collect in the vent port that is protected with an exchangeable liner. The second orthogonal step enables the volume of gas (and ions) sampled from the atmosphere to be increased compared with conventional API sources. Gas at atmospheric pressure is sampled through an aperture into a partial vacuum to form a freely expanding jet, which represents a region of high performance compared to the surrounding vacuum. When

this jet is directed into the second aperture of a conventional API interface it increases the flow of gas through the second aperture. Maintaining a suitable vacuum in the MS-TOF therefore places a restriction on the maximum diameter of the apertures in such an LC interface. Ions in the partial vacuum of the ion block are extracted electrostatically into the hexapole ion bridge which efficiently transports ions to the analyzer.

In one aspect, TOF mass analyzers are coupled with MUX-technology to allow the connection of up to 8 HPLC columns in parallel to a single LC-TOF-MS. (Micromass, Manchester UK). A multiplexed electrospray (ESI) interface may be used for on-line LC-MS utilizing an indexed stepper motor to sequentially sample from up to 8 HPLC columns or liquid inlets operated in parallel.

Use of LC-TOF-MS is generally preferred over use of MALDI-TOF because LC-TOF-MS is a quantitative method for analysis of the molecular weight of polymers. LC-TOF-MS does not fragment the polymers and it employs a very gentle ionization process compared to matrix-assisted-laser-desorption-ionization (MALDI). Because every MALDI blast is different, the ionization is not quantitative. LC-TOF-MS does, however, produce different m/z values for polymers, but, as disclosed in Example 1 and Figures 7-17, this property provides the additional advantage of reducing background and providing complementary information.

In one aspect, tandem MS or MS/MS is used for structure determination of molecular ions or fragments. In tandem MS, the ion of interest is selected with the first analyzer (MS-1), collided with inert gas atoms in a collision cell, and the fragments generated by the collision are separated by a second analyzer (MS-2). In Ion Trap and Fourier transform experiments, the analyses are carried out in one analyzer, and the various events are separated in time, not in space. The information can be used to sequence peptides and small DNA and RNA oligomers.

In one aspect, exact mass measurements, sometimes referred to as "high-resolution measurements," are used for elemental-composition determination of the sample molecular ion or an ionic fragment. The basis of the method is that each element has a unique mass defect (deviation from the integer mass). The measurement is carried out by scanning with an internal calibrant (in EI or CI mode) or by peak matching (in FAB mode). The elemental composition is determined by comparing the masses of many possible compositions to the measured one. The method is very reliable for fragments having masses up to 800 Da. At higher masses, higher precision and/or knowledge of expected composition is required to determine the elemental composition unambiguously.

Electron ionization (EI) is widely used in mass spectrometry for relatively volatile samples that are insensitive to heat and have relatively low molecular weight. The spectra, usually containing many fragment-ion peaks, are useful for structural characterization and identification. Small impurities in the sample are easy to detect. Chemical ionization (CI) is applied to similar samples; it is used to enhance the abundance of the molecular ion. For both ionization methods, the molecular weight range is 50 to 800 Da, and thus either of these ionization methods may be used according to the present invention, when the fragment is within the stated range. In rare cases it is possible to analyze samples of higher molecular weight. Accuracy of the mass measurement at low resolving power is ±0.1 Dalton and in the high resolution mode, ±5 ppm.

Fast atom bombardment ionization (FAB, or sometimes called liquid secondary ionization MS, LSIMS) is a softer ionization method than EI. The spectrum often contains peaks from the matrix, which is necessary for ionization, a few fragments and a peak for a protonated or deprotonated sample molecule. FAB is used to obtain the molecular weight of sensitive, nonvolatile compounds. The method is prone to suppression effects by small impurities. The molecular weight range is 100 to 4000 Da. Exact mass measurement is usually done by peak matching. The accuracy of the mass is the same as obtained in EI, CI, and thus this is another suitable ionization method to be used in the characterization of fragments prepared by methods described herein.

Matrix-assisted laser desorption (MALDI) has been used to determine the molecular weight of peptides, proteins, oligonucleotides, and other compounds of biological origin as well as of small synthetic polymers. The amount of sample needed is very low (pmoles or less). The analysis can be performed in the linear mode (high mass, low resolution) up to a molecular weight of *m/z* 300,000 (in rare cases) or reflectron mode (lower mass, higher resolution) up to a molecular weight of 10,000. The analysis is relatively insensitive to contaminants, and accordingly a purification step is not generally necessary as part of the characterization process when characterization includes MS. However, in certain embodiments, LC or SPE may be used to separate or purify single-stranded nucleic acid fragments before subjecting these fragments to MALDI. Mass accuracy (0.1 to 0.01%) of traditional MADLI is not as high as for other mass spectrometry methods. However, recent development in Delayed Extraction TOF allows higher resolving power and mass accuracy. Thus, MALDI is used in one aspect of the present invention. In addition, in the embodiments where amplified single-stranded nucleic acid molecules are first purified or separated via

liquid chromatography, MALDI allows for the pooling of multiple fractions of liquid chromatography before mass spectrometric analysis, thus enables higher throughput.

Electrospray ionization (ESI) allows production of molecular ions directly from samples in solution. It can be used for small and large molecular-weight biopolymers (peptides, proteins, carbohydrates, and DNA fragments), and lipids. Unlike MALDI, which is pulsed, ESI is a continuous ionization method that may be interfaced with HPLC or capillary electrophoresis. Multiply charged ions are usually produced. ESI may be considered a complement to MALDI, and may be used according to the present invention. The sample must be soluble, stable in solution, polar, and relatively clean (free of nonvolatile buffers, detergents, salts, etc.).

Electron-capture (sometimes called negative ion chemical ionization or NICI) is used for molecules containing halogens, NO_2, CN, etc, and it usually requires that the analyte be derivatized to contain highly electron-capturing moieties (*e.g.*, fluorine atoms or nitrobenzyl groups). Such moieties are generally inserted into the target analyte after isolation and before mass spectrometric analysis. The sensitivity of NICI analyses is generally two to three orders of magnitude greater than that of PCI or EI analyses. Little fragmentation occurs during NICI. Electron-capture can therefore be used when the nucleic-acid fragment has the required chemical or atomic groups, or when such groups are added to the fragment as discussed below.

Various mass analyzers may be used to analyze samples that have been ionized. Such mass analyzers include time of flight, quadrupoles, magnetic sector and ion trap mass analyzers. For general discussion of various types of ionization methods and mass analyzers, see Siuzdak, *Mass Spectrometry for Biotechnology*, Academic Press, 1996.

ii. Indirect characterization

In some embodiments of the present invention, it may be advantageous to add one or more detectable labels to a short single-stranded nucleic acid fragment or a reaction product thereof (*e.g.*, a portion or the whole complementary strand of the short nucleic acid fragment). Such labels may facilitate the characterization of the fragment and thereby the identification of nucleotide(s) of interest and/or genetic variations within the fragment.

Tables 1 and 2 summarize exemplary labels and detectors, respectively, that are generally suitable for use in methodologies for detecting small nucleic acid fragments, and which may be used to detect the nucleic acid fragments that are produced by the methods described herein.

TABLE 1

LABELS SUITABLE FOR USE IN METHODOLOGIES FOR DETECTING SMALL NUCLEIC ACID FRAGMENTS

Tagging Technologies	Attributes
Fluorophores	Multi-color, overlapping emission spectra, inexpensive detectors
FRET	High sensitivity
Fluorescent quenching	Homogenous assay formats
Time-resolved fluorescence	Low background
Colloidal gold	Good sensitivity
Mass Tags (CMSTs)	High level of multiplexing
Mass Tags (Electrophore)	High level of multiplexing
Radiolabels	Excellent sensitivity
Chemiluminescence	Excellent sensitivity
Colorimetric	Inexpensive
Assay product = "Tag"	Accurate, inexpensive, direct

TABLE 2

DETECTORS SUITABLE FOR USE IN METHODOLOGIES FOR DETECTING SMALL NUCLEIC ACID FRAGMENTS

Detector	Attributes
Film	Inexpensive
Scintillation Counter	Reliable, sensitive
Fluorescent plate reader	Reliable, inexpensive, sensitive, multicolor
Fluorescence Polarization	Permits homogeneous assay formats, some instruments very sensitive.
Time-resolved fluorescence	Low background, sensitive
Fluorescent-monitoring of PCR	Useful information on the process of PCR
ABI-377	Reliable
Capillary Instrument	High throughput, expensive
Chemiluminescence plate reader	Reliable, sensitive
CCD	Versatile, sensitive
Quadrupole MS	Wide spectral range, quantitative
GC/MS	
Maldi-TOF	Wide spectral range, not quantitative
Plate Reader (colorimetric assays)	Reliable, inexpensive, sensitive
Cell Sorter	High throughput
Light Microscopy (Confocal)	Excellent sensitivity
Electron microscopy	Sensitivity
Amphoteric device	Ability to multiplex
DHPLC (HPLC/UV)	Reliable, relatively inexpensive
HPLC/Fluorescence	Reliable, sensitive, relatively inexpensive
Text scanner	Very inexpensive, make your own assay
UV box (for stains)	Very inexpensive

Detectors for these tags and labels are available in generic and non-generic instruments. The generic instruments are the plate readers that usually read micro-plates in 96-well or 384-well formats, and are capable of reading multiple colors (4-6 fluorescent tags). These instruments can be found in customized versions to perform more specialized measurements like time-resolved-fluorescence (TFR) or fluorescence polarization. The detectors for PAGE sequencing and bundled capillary instruments are highly dedicated and non-generic. The generic mass spectrometers

MALDI-TOF, electrospray-TOF and APCI-quadrupole (and combinations thereof including ion-trap instruments) are opened-ended instruments with versatility. Suitable software packages have been developed for combinatorial chemistry applications and may be readily adapted for use in the detection of fragments as described herein. Scintillation counters are dedicated in that they need to be used with radioisotopes, but can accommodate a wide range of assays formats.

The following describes exemplary indirect characterization methodologies. However, the present invention is not limited to these examples. Any technique known in the art suitable for characterizing small nucleic acid fragments and thereby determining the identity of nucleotide(s) at a defined location may be used in the present invention.

(a) Sequencing

In one aspect of the invention, a single-stranded nucleic acid fragment is characterized by performing a complete nucleotide sequence analysis. Many techniques are known in the art for identifying each of the bases in a nucleic acid fragment, so as to obtain base sequence information. For instance, two different DNA sequencing methodologies that were developed in 1977 and are commonly known as "Sanger sequencing" and "Maxam-Gilbert sequencing," among other names, are still in wide use today and are well known to those of ordinary skill in the art. *See, e.g.,* Sanger, *Proc. Natl. Acad. Sci. (USA) 74*:5463, 1977) and Maxam and Gilbert, *Proc. Natl. Acad. Sci. (USA) 74*:560, 1977). Both methods produce populations of shorter fragments that begin from a particular point and terminate in every base that is found in the nucleic acid fragment that is to be sequenced. The shorter nucleic acid fragments are separated by polyacrylamide gel electrophoresis and the order of the DNA bases (adenine, cytosine, thymine, guanine; also known as A,C,T,G, respectively) is read from a autoradiograph of the gel. These methods are generally more successful with relatively long single-stranded nucleic acid fragments.

Automated DNA sequencing methods may also be used. Such methods are in wide-spread commercial use to sequence both long and short nucleic acid molecules. In one approach, these methods use fluorescent-labeled primers or ddNTP-terminators instead of radiolabeled components. Robotic components can utilize polymerase chain reaction (PCR) technology which has lead to the development of linear amplification strategies. Current commercial sequencing allows all 4 dideoxy-terminator reactions to be run on a single lane. Each dideoxy-terminator reaction is represented by a unique fluorescent primer (one fluorophore for each base type: A, T, C,

G). Only one template DNA (*i.e.*, DNA sample) is represented per lane. Current gels permit the simultaneous electrophoresis of up to 64 samples in 64 different lanes. Different ddNTP-terminated fragments are detected by the irradiation of the gel lane by light followed by detection of emitted light from the fluorophore. Each electrophoresis step is about 4-6 hours long. Each electrophoresis separation resolves up to about 400-600 nucleotides (nt), therefore, about 6000 nt can be sequenced per hour per sequencer.

Gilbert has described an automated DNA sequencer (EPA, 92108678.2) that consists of an oligomer synthesizer, an array on a membrane, a detector which detects hybridization and a central computer. The synthesizer synthesizes and labels multiple oligomers of arbitrary predicted sequence. The oligomers are used to probe immobilized DNA on membranes. The detector identifies hybridization patterns and then sends those patterns to a central computer which constructs a sequence and then predicts the sequence of the next round of synthesis of oligomers. Through an iterative process, a DNA sequence can be obtained in an automated fashion. This approach may be used to characterize a short nucleic acid fragment (either double or, more commonly single stranded) according to the present invention.

The use of mass spectrometry for the study of monomeric constituents of nucleic acids has also been described (Hignite, "Biochemical Applications of Mass Spectrometry", Waller and Dermer (eds.), Wiley-Interscience, Chapter 16, p. 527, 1972). Briefly, for larger oligomers, significant early success was obtained by plasma desorption for protected synthetic oligonucleotides up to 14 bases long, and for unprotected oligonucleotides up to 4 bases in length. As with proteins, the applicability of ESI-MS to oligonucleotides has been demonstrated (Covey et al., *Rapid Comm. in Mass Spec. 2*:249-256, 1988). These species are ionized in solution, with the charge residing at the acidic bridging phosphodiester and/ or terminal phosphate moieties, and yield in the gas phase multiple charged molecular anions, in addition to sodium adducts. These approaches to nucleic acid characterization may be used according to the present invention.

Sequencing nucleic acids with less than 100 bases by the common enzymatic ddNTP technique is more complicated than it is for larger nucleic acid templates, so that chemical degradation is sometimes employed. However, the chemical decomposition method requires about 50 pmol of radioactive ^{32}P end-labeled material, 6 chemical steps, electrophoretic separation, and film exposure. For small oligonucleotides (<14 nts), as may need to be characterized according to the present invention, the combination of electrospray ionization (ESI) and Fourier transform (FT) mass spectrometry (MS) is far faster and more sensitive, and is a preferred method for

the present invention. Dissociation products of multiply-charged ions measured at high (105) resolving power represent consecutive backbone cleavages providing the full sequence in less than one minute on sub-picomole quantity of sample (Little et al., *J. Am. Chem. Soc. 116*:4893, 1994). For molecular weight measurements, ESI/MS has been extended to larger fragments (Potier et al., *Nuc. Acids Res. 22*:3895, 1994). ESI/FTMS appears to be a valuable complement to classical methods for sequencing and pinpoint mutations in nucleotides as large as 100-mers. Spectral data have recently been obtained loading 3 x 10-13 mol of a 50-mer using a more sensitive ESI source (Valaskovic, *Anal. Chem. 68*:259, 1995).

Other methods for obtaining complete, or near complete base sequence information for a nucleic acid molecule are described in the following references: Brennen et al. (Biol. Mass Spec., New York, Elsevier, p. 219, 1990); U.S. Patent No. 5,403,708; PCT Patent Application No. PCT/US94/02938; and PCT Patent Application No. PCT/US94/11918. Each of these identified characterization techniques may be used to characterize single-stranded nucleic acid fragments produced as described herein.

(b) Fluid Handling

As used herein, the term "fluid handling" refers to those assays that are microtiter-plate based and use fluorescence, fluorescence-polarization, luminescence, radioactivity (scintillation counters), or colorimetric readouts. Fluid handling may be useful when the characterization method employs modification of the short nucleic acid fragment, *e.g.*, when a tag or label is incorporated into the short nucleic acid fragment. These assays can be amplified by the use of enzymes such as horseradish peroxidase or alkaline phosphatase that can generate soluble or insoluble colorimetric products from soluble substrates or sensitive luminescent products. These assays have large dynamic ranges (6-8 logs) and can be made robust. Fluid handling using microplates scales well and has been partially miniaturized by the use of 384-well plates. Fluid handling is especially compatible with commercial robotics and readout systems such as fluorometers, and plate readers. The data is easy to archive and manipulate.

iii. Fractionation Methodologies

According to the present invention, the small nucleic acid fragment(s) may, optionally, undergo a step of fractionation prior to a step of detection. The fractionation step may simply remove undesired impurities from the small fragment of interest, to allow more convenient and/or more accurate characterization of the

fragment. This type of fractionation step may be referred to as purification. Alternatively, or in addition, the fractionation may separate nucleic acids from one another (such as in chromatography) and the detection technique is simply determining whether the nucleic acid is, or is not, present at a particular time and space (*e.g.*, using ultraviolet detection to determine whether a nucleic acid is eluting from a chromatography column).

Thus, depending on the particular detection methodology employed, it may be advantageous to couple a detection methodology with one or more methodologies for the fractionation of small nucleic acid fragments. As discussed below, such fractionation methodologies include, but are not limited to, electrophoresis including polyacrylamide or agarose gel electrophoresis and capillary electrophoresis, and liquid chromatography (LC) including high pressure liquid chromatography (HPLC) and denaturing high pressure liquid chromatography (DHPLC).

(a) Gel Electrophoresis

As used herein, the term "electrophoresis" refers generally to those separation techniques based on the mobility of nucleic acid in an electric field. Negatively charged nucleic acid migrates towards a positive electrode and positively charged nucleic acid migrates toward a negative electrode. Charged species have different migration rates depending on their total charge, size, and shape, and can therefore be separated.

An electrophoresis apparatus consists of a high-voltage power supply, electrodes, buffer, and a support for the buffer such as a polyacrylamide gel, or a capillary tube. Open capillary tubes are used for many types of samples and the other gel supports are usually used for biological samples such as protein mixtures or nucleic acid fragments.

The most powerful separation method for nucleic acid fragments is PAGE, generally in a slab gel format. The major limitation of the current technology is the relatively long time required in performing the gel electrophoresis of nucleic acid fragments produced in sequence reactions. An increased magnitude (10-fold) can be achieved with the use of capillary electrophoresis which utilize ultrathin gels.

Capillary electrophoresis (CE) in its various forms, including free solution, isotachophoresis, isoelectric focusing, PAGE, and micellar electrokinetic "chromatography," is a suitable technology for the rapid, high resolution separation of very small sample volumes of complex mixtures. In combination with the inherent sensitivity and selectivity of mass spectrometry (CE-MS; see below), CE is a powerful

technique for bioanalysis. In the methodology disclosed herein, the interfacing of these two methods provides superior DNA sequencing methods that are superior to the current rate methods of sequencing.

By alternate embodiments, CE may be employed in conjunction with electrospray ionization (ESI) flow rates. The combination of both capillary zone electrophoresis (CZE) and capillary isotachophoresis with quadrupole mass spectrometers based upon ESI have been described. (Olivares et al., *Anal. Chem. 59*:1230 (1987); Smith et al., *Anal. Chem. 60*:436 (1988); Loo et al., *Anal. Chem. 179*:404 (1989); Edmonds et al., *J. Chroma. 474*:21 (1989); Loo et al., *J. Microcolumn Sep. 1*:223 (1989); Lee et al., *J. Chromatog. 458*:313 (1988); Smith et al., *J. Chromatog. 480*:211 (1989); Grese et al., *J. Am. Chem. Soc. 111*:2835 (1989) each of which is incorporated herein by reference in its entirety). Small nucleic acids are easily amenable to CZE analysis with good (femtomole) sensitivity.

Polyacrylamide gels, such as those discussed above, may be applied to CE methodologies. Remarkable plate numbers per meter have been achieved with cross-linked polyacrylamide. (*See, e.g.*, Cohen et al., *Proc. Natl. Acad. Sci., USA 85*:9660 (1988) reporting 10^{+7} plates per meter). Such CE columns as described can be employed for nucleic acid (particularly DNA) sequencing. The CE methodology is in principle 25 times faster than slab gel electrophoresis in a standard sequencer. For example, about 300 bases can be read per hour. The separation speed is limited in slab gel electrophoresis by the magnitude of the electric field that can be applied to the gel without excessive heat production. Therefore, the greater speed of CE is achieved through the use of higher field strengths (300 V/cm in CE versus 10 V/cm in slab gel electrophoresis). The capillary format reduces the amperage and thus power and the resultant heat generation.

In alternative embodiments of the present invention, multiple capillaries may be used in parallel to increase throughput and may be used in conjunction with high throughput sequencing. (Smith et al., *Nuc. Acids. Res. 18*:4417 (1990); Mathies et al., *Nature 359*:167 (1992); Huang et al., *Anal. Chem. 64*:967 (1992); Huang et al., *Anal. Chem. 64*:2149 (1992)). The major disadvantage of capillary electrophoresis is the limited volume of sample that can be loaded onto the capillary. This limitation may be circumvented by concentrating large sample volumes prior to loading the capillary with the accompanying benefit of >10-fold enhancement in detection. In one aspect of the invention, the small single-stranded nucleic acid fragments are preconcentrated prior to characterization.

A suitable method for preconcentration in CE is sample stacking. (Chien et al., *Anal. Chem. 64*:489A (1992)). Sample stacking depends on the matrix difference (*i.e.*, pH and ionic strength) between the sample buffer and the capillary buffer, so that the electric field across the sample zone is more than in the capillary region. In sample stacking, a large volume of sample in a low concentration buffer is introduced for preconcentration at the head of the capillary column. The capillary is filled with a buffer of the same composition, but at higher concentration. When the sample ions reach the capillary buffer and the lower electric field, they stack into a concentrated zone. Sample stacking has increased detectability by 1-3 orders of magnitude.

Alternatively, preconcentration may be achieved by applying isotachophoresis (ITP) prior to the free zone CE separation of analytes. ITP is an electrophoretic technique that allows microliter volumes of sample to be loaded onto the capillary, in contrast to the low nL injection volumes typically associated with CE. This technique relies on inserting the sample between two buffers (leading and trailing electrolytes) of higher and lower mobility followed by the analyte. The technique is inherently a concentration technique, where the analytes concentrate into pure zones migrating with the same speed. The technique is typically less preferred than the stacking methods described above because of the need for several choices of leading and trailing electrolytes, and the ability to separate only cationic or anionic species during a separation process.

Central to the nucleic acid sequencing process is the remarkably selective electrophoretic separation that may be achieved with nucleic acid fragments. Separations are routinely achieved with fragments differing in sequence by only a single nucleotide. This methodology is suitable for separations of fragments up to 1000 bp in length.

A further advantage may be achieved by sequencing with cleavable tags. *See, e.g.,* European Patent Nos. 0 868 535; 0 840 804; and 0 850 320, as well as U.S. Patent No. 6,027,890. When using cleavable tags, there is no requirement to use a slab gel format when nucleic acid fragments are separated by PAGE. Since numerous samples are combined (4 to 2000) there is no need to run samples in parallel as is the case with current dye-primer or dye-terminator methods (*i.e.*, ABI 373 sequencer). Since there is no reason to run parallel lanes, there is no reason to use a slab gel. Therefore, one can employ a tube gel format for the electrophoretic separation method. It has been shown that considerable advantage is gained when a tube gel format is used in place of a slab gel format. (Grossman et al., *Genet. Anal. Tech. Appl. 9*:9 (1992)). This is due to the greater ability to dissipate Joule heat in a tube format compared to a

slab gel which results in faster run times (by 50%), and much higher resolution of high molecular weight nucleic acid fragments (greater than 1000 nt). Long reads are critical in genomic sequencing. Therefore, the use of cleavable tags in sequencing has the additional advantage of allowing the user to employ the most efficient and sensitive nucleic acid separation method that also possesses the highest resolution.

As discussed above, CE is a powerful method for nucleic acid sequencing, particularly DNA sequencing, forensic analysis, PCR product analysis and restriction fragment sizing. CE is faster than traditional slab PAGE since with capillary gels a higher potential field can be applied, but has the drawback of allowing only one sample to be processed per gel. Thus, by alternative embodiments according to the present invention, micro-fabricated devices (MFDs) are employed to combine the faster separations times of CE with the ability to analyze multiple samples in parallel.

MFDs permit an increase in information density in electrophoresis by miniaturizing the lane dimension to about 100 micrometers. The current density of capillary arrays is limited to the outside diameter of the capillary tube. Microfabrication of channels produces a higher density of arrays. Microfabrication also permits physical assemblies not possible with glass fibers, and links the channels directly to other devices on a chip. A gas chromatograph and a liquid chromatograph have been fabricated on silicon chips, but these devices have not been widely used. (Terry et al., *IEEE Trans. Electron Device ED-26*:1880 (1979) and Manz et al., *Sens. Actuators B1*:249 (1990)). Several groups have reported separating fluorescent dyes and amino acids on MFDs. (Manz et al., *J. Chromatography 593*:253 (1992); Effenhauser et al., *Anal. Chem. 65*:2637 (1993)).

Photolithography and chemical etching can be used to make large numbers of separation channels on glass substrates. The channels are filled with hydroxyethyl cellulose (HEC) separation matrices. DNA restriction fragments could be separated in as little as two minutes. (Woolley et al., *Proc. Natl. Acad. Sci. 91*:11348 (1994)).

(b) Liquid Chromatography (LC)

Liquid chromatography, including HPLC and DHPLC, may be used in conjunction with one of the detection methodologies discussed above such as, for example, fluorescence polarization, mass spectrometry and/or electron ionization. Alternatively LC, HPLC and/or DHPLC may be utilized in conjunction with a UV detection methodology. Regardless of the detection methodology employed, liquid chromatography provides the separation of complex mixtures of non-volatile

compounds prior to detection. In certain embodiments of the present invention, LC alone can be used to identify and/or distinguish small nucleic acid molecules from each other, and thus allow an identification of the nucleotide(s) of interest in the target nucleic acid (based on characterizing the short nucleic acid fragments prepared from the target nucleic acid as described herein).

Remarkably, two nucleic acid molecules may be of the same length, *i.e.*, contain the same number of nucleotides, and differ from each other only in their sequences at one nucleotide position, and yet these two molecules can be distinguished from one another based on retention time obtained from liquid chromatographic analysis performed according to the present invention. Also remarkable is that two nucleic acid molecules that are of the same length and of the same nucleotide composition, and differ from each other only in the order in which the nucleotides are arranged, may be distinguish from one another based on retention time obtained from liquid chromatographic analysis performed according to the present invention. As one example, the order of two nucleotides in a nucleic acid molecule can be switched to provide a variant nucleic acid molecule, and the original and variant nucleic acid molecules may be distinguished by their different retention times as observed by liquid chromatographic analysis (*see, e.g.,* 8merB and 8merC in Example 1). An important factor in being able to use retention time to distinguish two similar nucleic acid molecules is the selection of the mobile phase for the chromatographic analysis, and thus this subject is described in detail below.

The mobile phase is a gradient formed from two buffer solutions (Buffer A and Buffer B). That is, after the nucleic acid molecules have been applied to the column, Buffer A, or an elution buffer that is largely Buffer A, is used to elute the molecules. During the course of the chromatography, Buffer B is gradually (incrementally) added to Buffer A so that the eluting buffer gradually becomes enriched in the components specific to Buffer B. Buffer A is an aqueous solution of an (one or more) ammonium salt, while Buffer B is an organic solution of an (one or more) ammonium salt. As referred to herein, an aqueous solution necessarily contains water, but may also contain non-aqueous components including organic components, *e.g.*, organic solvent(s). As referred to herein, an organic solution necessarily contains organic material, but may and typically will also contain non-organic components including inorganic components, *e.g.*, water as a solvent. The water used in Buffers A and/or B is preferably HPLC grade water, where this quality of water is well known in the art.

The ammonium salt(s) present in Buffer A need not be identical to the ammonium salt(s) present in Buffer A, however, in one aspect of the invention Buffer A and Buffer B contain the same ammonium salts. Either of Buffer A or Buffer B may, independently of the other, contain more than one ammonium salt, however, in one aspect of the invention Buffer A and Buffer B each contain a single ammonium salt structure. In one aspect of the invention, each of Buffer A and Buffer B contain a single ammonium salt and furthermore these two buffers contain the same ammonium salt.

Referring to Buffer A, in one aspect of the invention water is the only solvent present in this buffer. In other aspects of the invention, water constitutes at least 99%, at least 97%, at least 95%, at least 90%, at least 85%, or at least 80% of the total volume of solvent present in Buffer A. Referring to Buffer B, in one aspect of the invention this buffer contains both organic solvent and water. Typically, water is necessarily present in Buffer B in order to solubilize the ammonium salt in the buffer. In various aspects of the invention, the organic solvent constitutes up to 75%, up to 60%, up to 45%, up to 30%, or up to 15% of the total volume of solvents present in Buffer B, with water contributing the residual volume of solvent. In various aspects of the invention, the volume ratio of organic solvent:water in Buffer B is 5-75:95-25, or 10-50:90-50; or 15-40:85-60; or 20-35:80-65 with the total of the organic solvent and the water equaling 100.

As to the organic solvent, this is preferably a solvent that is very miscible with water. Preferably, the organic solvent is miscible with water to an extent of at least about 5 vol% based on the total volume of water and organic solvent, and in various aspects is miscible to an extent of 10 vol%, 20 vol%, 30 vol%, 40 vol%, or 50 vol% (*i.e.*, an equal volume of water and organic solvent forms a homogeneous solution). As referred to herein, a solvent is "organic" so long as it contains at least one carbon. The organic solvent is preferably a liquid at room (ambient, standard) temperature and pressure. Such solvents are very well known in the art. In one aspect, the organic solvent is an alcohol, *i.e.*, an organic liquid at room temperature than contains at least one hydroxyl (OH) group. Exemplary alcohols include, without limitation, methanol, ethanol, ethylene glycol, *iso*-propanol, *n*-propanol, propylene glycol, and glycerol. Typically, as the number of carbon atoms present in the alcohol increases, the alcohol becomes less soluble in water, and the alcohol must contain more than one hydroxyl group in order to embody effective miscibility with water. In another aspect the organic solvent is acetonitrile. In yet another aspect the organic solvent is dimethylsulfoxide (DMSO).

As to the ammonium salt(s), these are selected from primary (RNH_2), secondary (R_2NH) and tertiary (R_3N) amines that are complexed with protic acid (H-A). In one aspect, the ammonium salt incorporates a primary amine, in another aspect the ammonium salt incorporates a secondary amine, in yet another aspect the ammonium salt incorporates a tertiary amine. In another aspect the ammonium salt selected from secondary and tertiary amine salts. The ammonium salt should be soluble in the solvent(s) that form the buffer into which the ammonium salt is to be added. In various aspects the ammonium salt is soluble in water at a concentration up to 1,000 mM, or up to 800 mM, or up to 600 mM, or up to 400 mM, or up to 200 mM, or up to 100 mM. Particularly if the liquid chromatographic analysis will be followed by mass spectroscopic analysis, it is desirable to minimize the amount of ammonium salt in the fractions that are obtained from liquid chromatography. High concentrations of ammonium salt can impair the ionization efficiency of the mass spectrometer, particularly when the mass spectrometer operates by electrospray ionization. Ammonium salts typically display water solubility so long as the R groups that form the ammonium salts are not, as a whole, too hydrophobic. In various aspect the ammonium salt is selected so as to provide an aqueous solution with a pH in the range of from 5.0 to 9.0, or 5.5 to 8.5, or 6.0 to 8.0, or 6.5 to 7.5.

In various aspects an R group is, independently at each occurrence within an ammonium molecule, and for each of primary, secondary and tertiary ammonium salts, hydrocarbon, *i.e.*, the R group is formed entirely of hydrogen and carbon; hydrocarbons with 1-10 carbons; hydrocarbons with 1-6 carbons; hydrocarbons with 5-8 carbons; selected from alkyl and cycloalkyl groups; selected from methyl, ethyl, propyl (including geometric isomers thereof), butyl (including geometric isomers thereof), pentyl (including geometric isomers thereof), hexyl (including geometric isomers thereof), cyclohexyl, and cycopentyl (including methyl-substituted cyclopentyl); selected from organic groups having an atomic mass of 15-250, which may or may not include atoms other than carbon and hydrogen; halide substituted in one occurrence; and/or halide substituted in two occurrences; where an R group may according to the present invention embody any two or more of these features in every combination so long as two features are not contradictory. As referred to herein, alkyl and cylcoalkyl groups may contain unsaturation, *e.g.*, a double bond, however cycloalkyl does not include aromatic rings. Nevertheless, in one aspect the R group is an aromatic ring selected from phenyl and C_1-C_6alkyl substituted phenyl. In another aspect, the R group is an alkyl group having aryl substitution, *e.g.*, benzyl. In one aspect, the ammonium salt is a secondary or tertiary ammonium salt having alkyl or

cycloalkyl groups with 5-8 carbons. In one aspect, the ammonium salt contains a cationic, *i.e.*, protonated form of triethylamine, diallylamine, diisopropylamine, N,N-dimethyl-N-cyclohexylamine, or N,N-dimethyl-N-isopropylamine. In one aspect, the ammonium salt is the salt form of N-N-dimethylaminobutane. In another aspect, the ammonium salt is the salt form of N,N-dimethylcyclohexylamine. In another aspect the ammonium salt is the salt form of triethylamine.

The selection of the amino component of the ammonium salt can have an effect on the signal generated by the nucleic acid molecule using liquid chromatography followed by mass spectroscopy. For instance, while dimethylaminobutane and dimethylcyclohexylamine provide similar responses, triethylamine decreases response as observed by mass spectrometry following liquid chromatography, by about 25% compared to dimethylaminobutane.

In one aspect, the protic acid of the ammonium salt is of the formula Ra-COOH so that the anionic counterion to the ammonium group has the formula Ra-COO^-. In various aspects, Ra is hydrogen (*i.e.*, the counterion is derived from formic acid) methyl, ethyl, propyl, selected from methyl and ethyl, or selected from methyl, ethyl and propyl, and mono- and poly-halogenated versions thereof. In another aspect, the protic acid used to form the ammonium salt is a protonated form of the carbonate family of anions, *i.e.*, bicarbonate (HCO_3) and/or carbonate (CO_3) is the anionic counterion to the ammonium group. In other aspects of the invention, the protic acid is an inorganic acid, *e.g.*, HCl and HBr. Generally, in order of preference, the protic acid of the ammonium salt is acetic acid, formic acid, carbonic acid (so as to provide bicarbonate anion), or hydrogen chloride.

In one aspect of the invention, the ammonium group is a tertiary ammonium group having R groups as defined above, and the counterion is acetate (not including halogenated acetate), carbonate/ bicarbonate, or is selected from acetate (not including halogenated acetate) and carbonate/ bicarbonate. When the counterion is acetate, exemplary ammonium salts include, without limitation, triethylamine acetate, dimethylbutamine acetate, dimethylisopropylamine acetate, dimethylhexylamine acetate, dimethylcyclohexylamine acetate, and disopropylamine acetate.

Beside the selection of Buffers A and B, the conditions under which the liquid chromatography is operated are important in being able to resolve, or distinguish, two nucleic acid structures of identical length but of somewhat different nucleotide base sequence. For example, generally the liquid chromatographic analysis of the present invention is carried out at either room temperature, *i.e.*, about 25°C, or at elevated temperature. Typically, elevated temperatures are less than 75°C. In various aspects,

the chromatography column is maintained within the following temperature ranges during the chromatography: 20°C-80°C, 25°C-70°C, 25°C-65°C, 25°C-60°C, 30°C-80°C, 30°C-70°C, 30°C-65°C, 30°C-60°C, 30°C-55°C, 30°C-50°C, or 30°C-45°C. Elevated temperature is typically desirable because it may provide a chromatogram wherein the peaks are higher, narrower, and/or display greater resolution. However, as the temperature exceeds about 70°C, bubble formation in the eluent is sometimes observed and this may cause a loss in resolution.

As mentioned above, the pH of the elution buffer is also an important factor in the successful chromatographic analysis. Generally, the elution buffer maintains a pH ranging from about 5.0 to 9.0, as discussed above. Typically, retention time is unaffected so long as the pH of the elution buffer is maintained within this pH range.

The chromatography is typically run under pressure, *i.e.,* pressure is used to push the elution buffer through the column. As the pressure is increased there is typically an increase in the flow rate of the buffer through the column. A pressure ranging from about 200-1600 bars, and a flow rate ranging from about 10 µL to 2000 µL per minute are typically suitable conditions for operating the column chromatography.

The length and stationary phase of the liquid chromatography column are two other parameters that must be selected. Typically, the column is 10-500 mm in length. In various aspects, the column is 10, 18, 25, 50, 100, 250, or 500 mm in length and can be at a micro-, or macro-scale. The stationary phase is selected so as to provide a reverse phase column, i.e., a column having a hydrophobic phase surrounding the solid phase. A suitable reverse phase column is the MICROSORB™ C18 column from Varian Inc. (Palo Alto, CA; www.varianinc.com). This column is a monomeric silica column, with 5 micron spherical particles, 300Å pores, C18 stationary phase, 12% carbon load and is endcapped. A substantially equivalent column is the JUPITER™ C18 column made by Phenomenex U.S.A. (Torrance, CA; www.phenomenex.com). A similar column is the XTERRA™ column from Waters (Milford, MA: www.waters.com) which contains a hybrid particle made of silica and polymer to extend pH stability with a pore size of 120Å and particle size from 2.5 micron to 5 micron. Columns having a completely polymeric solid support may also be used. In one preferred aspect of the invention, the reverse phase column has a C18 stationary phase and a pore size of at least 120Å. Endcapping of silica columns is desirable in order to minimize tailing and improve peak shape, and accordingly a preferred column has this feature. The carbon load of the column is important to ensure sufficient

retention (carbon load does not apply to polymeric columns), and is preferably in the range of 5-20%. The particle size typically varies from about 2 microns up to about 10 microns, where these are typical sizes for columns that are currently commercially available. Smaller particle size generally result in improved chromatography, so particle sizes smaller than 5 μM are preferred. Column dimensions are not critical, but may be chosen based on scale and type of analysis. Generally for fast, higher throughput analysis of small samples (<25 μL or 250 μg of analyte) a small column is preferred, such as the 2.1 x 15 mm (diameter x length) XTERRA™ column with 2.5 μM particles allowing a complete run in four minutes at a flow rate of 250 μL/min. When injecting larger samples sizes it is typically preferred to use a larger column, where larger columns are about 4.6 mm in diameter and about 250 mm in length. Columns having dimensions from 0.3 mm to 4.6 mm in internal diameter and from 10 mm to 250 mm long are available from many commercial suppliers (*e.g.*, Water) and are suitably used in the present invention. A typical LCMS column will be 1 mm x 50 mm. The flow rate will be dependent on the column dimensions and will vary from a few microliters per minute for a 0.3 mm ID column up to about 1500-2000 microliters per minute for a 4.6 mm ID column.

As explained above, the mobile phase is preferably formed by incrementally combining two different solutions (Buffers A and B). The salt is preferably present in the combined solution at a concentration of 1 mM to 200 mM, more preferably at a concentration of 1 mM to 100 mM, still more preferably at a concentration of 5 mM to 50 mM. Particularly when the liquid chromatographic analysis is followed by mass spectroscopic analysis, lower salt concentration is preferred, and in such a situation a salt concentration of about 5 mM is preferred. Buffer B preferably contains 10-90% (volume/volume) polar organic molecules (*e.g.*, acetonitrile, methanol, or isopropanol) in Buffer A. Buffers A and/or B may contain optional components, and in one aspect an optional component present in Buffer A is also present in Buffer B. A suitable optional component is EDTA, where EDTA may be used at a concentration of about 0.1 mM in the buffers.

For instance, liquid chromatography according to the present invention may be performed as follows. First, a shallow gradient of acetonitrile or other suitable solvent may be used to elute the nucleic acid molecules and provide for sample clean up. For instance, when using acetonitrile, the gradient can start at, or about at, 5% and increase to, or about to, 20%, where these percent values refer to the volume percent of organic solvent in water as the elution buffer. Methanol can be used in place of acetonitrile, but the use of methanol may require increasing the final gradient

composition to 50% or even 75% methanol. The fraction of solvent required depends partly on the column used, and also on the length range of nucleic acid molecules being analyzed. Generally a strong solvent wash is applied to the column at the end of the run to elute any large, hydrophobic components. In a preferred embodiment, the analysis portion of the gradient starts at 5% acetonitrile and increases to 15% over about 90 seconds, where this is followed by a wash which quickly pushes a "plug" of 45% acetonitrile onto the column for just a few seconds followed by a return to starting conditions of 5% acetonitrile. A preferred buffer system incorporates 5 mM N,N-dimethylaminobutyl acetate, and operates at pH 7. Concentrations of ammonium salts from 1 mM up to 50 mM are observed to have equivalent response in a mass spectrometer (*e.g.,* there is no evidence of ion suppression within this range) but variation of ammonium salt concentration within this range may have slight effects on the retention times of the nucleic acid molecules. Variation in retention times may be compensated for by adjusting the solvent composition during the LC run. As mentioned above, the pH range is flexible, where a range from pH 6 to pH 8 can be used with little or no noticeable change in outcome, *i.e.,* little or no effect on nucleic acid retention times.

In certain other embodiments, High-Performance Liquid Chromatography (HPLC) may be used. HPLC is a chromatographic technique for separation of compounds dissolved in solution. HPLC instruments consist of a reservoir of mobile phase, a pump, an injector, a separation column, and a detector. Compounds are separated by injecting an aliquot of the sample mixture onto the column. The different components in the mixture pass through the column at different rates due to differences in their partitioning behavior between the mobile liquid phase and the stationary phase. The pumps provide a steady high pressure with no pulsating, and can be programmed to vary the composition of the solvent during the course of the separation.

Exemplary detectors useful within the methods of present invention include UV-VIS absorption, or fluorescence after excitation with a suitable wavelength, mass spectrometers and IR spectrometers. Oligonucleotides labeled with fluorochromes may replace radio-labeled oligonucleotides in semi-automated sequence analysis, minisequencing and genotyping. (Smith et al., *Nature 321*:674 (1986)).

In one aspect of the invention, the (short) single-stranded nucleic acid fragment is characterized solely by chromatography as described herein in conjunction with a suitable detector for nucleic acid-containing element. In another aspect the characterization step entails chromatography to separate nucleic acid fragments as

described above, in conjunction with mass spectrometric analysis of the separated fragments.

IP-RO-HPLC on non-porous PS/DVB particles with chemically bonded alkyl chains may be employed in the analysis of both single and double-strand nucleic acids. (Huber et al., *Anal. Biochem. 212*:351 (1993); Huber et al., *Nuc. Acids Res. 21*:1061 (1993); Huber et al., *Biotechniques 16*:898 (1993)). In contrast to ion-exchange chromatography, which does not always retain double-strand DNA as a function of strand length (since AT base pairs interact with the positively charged stationary phase, more strongly than GC base-pairs), IP-RP-HPLC enables a strictly size-dependent separation.

A method has been developed using 100 mM triethylammonium acetate as ion-pairing reagent, so that phosphodiester oligonucleotides could be successfully separated on alkylated non-porous 2.3 μM poly(styrene-divinylbenzene) particles by means of high performance liquid chromatography. (Oefner et al., *Anal. Biochem. 223*:39 (1994)). The technique described allows the separation of PCR products differing by only 4 to 8 base pairs in length within a size range of 50 to 200 nucleotides.

Denaturing HPLC (DHPLC) is an ion-pair reversed-phase high performance liquid chromatography methodology (IP-RP-HPLC) that uses a non-porous C-18 column as the stationary phase that may be utilized in characterizing (short) single-stranded nucleic acid fragments produced as described herein.. The column is comprised of a polystyrene-divinylbenzene copolymer. The mobile phase is comprised of an ion-pairing agent of triethylammonium acetate (TEAA), which mediates binding of DNA to the stationary phase, and acetonitrile (ACN) as an organic agent to achieve subsequent separation of the DNA from the column. A linear gradient of acetonitrile allows separation. DHPLC identifies mutations and polymorphisms based on detection of heteroduplex formation between mismatched nucleotides in double stranded PCR amplified DNA. Sequence variation creates a mixed population of heteroduplexes and homoduplexes during reannealling of wild type and mutant DNA of fragments based on size and/or presence of heteroduplexes (this is the traditional use of the DHPLC technology). When this mixed population is analyzed by HPLC under partially denaturing temperatures, the heteroduplexes elute from the column earlier than the homoduplexes because of their reduced melting temperature. Analysis can be performed on individual samples to determine heterozygosity, or on mixed samples to identify sequence variation between individuals.

In certain applications, it may be preferred to use the DHPLC column in a non-denaturing mode in order to separate identically sized DNA fragments which

possess a different nucleotide composition. For example, the non-denaturing mode may be applicable where, for example, a 6-mer contains a C → T single nucleotide polymorphism (SNP) such as where the wild-type single strand DNA fragment has the nucleotide sequence 5′-AACCCC-3′ and where the mutant single strand DNA fragment has the nucleotide sequence 5′-AATCCC-3′. Fragments as short as 1-mers, 2-, 3-, 4-, 5-, 6-, 7-, 8-, to 16-mers show different mobilities (retention times) on the DHPLC instrument. Alternative to applications employing non-porous materials for performing the chromatography of the small nucleic acid fragments, HPLC as both sizing and DHPLC applications work on a wide pore silica based material. Porous materials have the advantage of high sample capacity for semipreparative work. This is marketed by Hewlett-Packard (Palo Alto, CA) as their Eclipse dsDNA columns.

e. Allelic Frequency Determination

Many of the methodologies for detecting and/or characterizing amplified single-stranded nucleic acid fragments may also be used to measure the amount of a particular amplified single-stranded nucleic acid fragment in the amplification reaction mixture. For instance, in the embodiments where an amplified single stranded nucleic acid molecule is first separated from the other molecules in the amplification reaction mixture by liquid chromatography and then subject to mass spectrometry analysis, the amount of the amplified single-stranded nucleic acid molecule may be quantified either by liquid chromatography of the fraction that contains the nucleic acid molecule, or by ion current measurement of the mass spectrometry peak corresponding to the nucleic acid molecule.

Such methodologies may be used to determine the allelic frequency of a target nucleic acid in a population of nucleic acids where the allelic variant(s) of the target nucleic acid may also be present. "Allelic variants," as used herein, refer to nucleic acid molecules that have an identical sequence to the target nucleic acid except at a defined location of the target nucleic acid. Both the sequence of the target nucleic acid at the defined location and the sequence of an allelic variant of the target nucleic acid at the same defined position may be referred to herein as a "genetic variation." "Allelic frequency of a target nucleic acid in a population of nucleic acids," as used herein, refers to the percentage of the total amount of the target nucleic acid and its allelic variant(s) in a nucleic acid population that is the target nucleic acid. Because the ODNP pairs of the present invention as described above are designed to anneal to portions of a target nucleic acid located at both sides of a genetic variation of interest present in the target, the amplification using the ODNP pairs as primers and a nucleic

acid population containing the target nucleic acid as templates produces the nucleic acid fragment that contains the genetic variation of the target nucleic acid (referred to as "template nucleic acid fragment derived from the target nucleic acid") as well as the nucleic acid fragment(s) that contains genetic variation(s) of allelic variant(s) of the target nucleic acid (referred to as "template nucleic acid fragment derived from an allelic variant of the target nucleic acid") if the variant(s) is present in the nucleic acid population. Because the sequences of the target nucleic acid and its allelic variant(s) differ only at the defined location, the template nucleic acid fragment derived from the target nucleic acid and those derived from the variant(s) are amplified at an identical, or a similar, efficiency. Likewise, the amplification of single-stranded nucleic acid molecules that contain the genetic variation of the target nucleic acid (referred to as "single-stranded target") and that of single-stranded nucleic acid molecules that contain the genetic variation(s) of the allelic variants of the target nucleic acid (referred to as "single-stranded variant(s)") using the template nucleic acid fragment derived from the target nucleic acid and that derived from the allelic variant(s) of the target nucleic acid as templates, respectively, are conducted at an identical, or a similar, efficiency. Consequently, the percentage of the single-stranded target within the total of the amplified single-stranded nucleic acids (*i.e.*, the sum of single-stranded target and single-stranded variant(s)), reflects the allelic frequency of the target nucleic acid in the nucleic acid population.

The present invention also provides methods for multiplex determination of allelic frequencies of a plurality of target nucleic acids in a nucleic acid population. In such a method, an ODNP pair is designed for each individual target nucleic acid to prepare the template nucleic acid that contains a genetic variation at a defined location in the target nucleic acid as well as the template nucleic acid(s) that contain the sequence(s) at the (corresponding) defined location in the allelic variant(s) of the target nucleic acid. The amplified template nucleic acids for every target nucleic acid of interest or digestion products thereof (*e.g.*, wherein one ODNP of the ODNP pairs comprises a RERS) are then used as templates for amplifying single-stranded nucleic acid fragments that contain the genetic variations of target nucleic acids or of the allelic variants of the target nucleic acids. Because the detection/characterization methodologies of the present invention (*e.g.*, liquid chromatography and mass spectrometry) can distinguish between amplified single stranded nucleic acid from one target nucleic acid or an allelic variant thereof and that from another target nucleic acid or an allelic variant thereof, the allelic frequencies of multiple target nucleic acid molecules may be determined simultaneously.

2. Kits for Identifying Genetic Variations

The present invention also provides kits for identifying genetic variations. Such kits generally comprise the ODNP pairs described above that are useful for preparing a template nucleic acid that contains a nicking site and a genetic variation located 3′ to the NS. They may also further comprise at least one, two, several, or each of the following components: (1) a nicking agent (*e.g.*, a NE or a RE) that recognizes the nicking agent recognition sequence (NARS) in at least one primer of the ODNP pair; (2) a suitable buffer for the nicking agent (1); (3) a RE that recognizes the RERS that is present in at least one primer of the ODNPs; (4) a suitable buffer for the RE (3); (5) a DNA polymerase useful for making the template nucleic acid (*i.e.*, extending from the 3′ termini of the above ODNP pair); (6) a suitable buffer for the DNA polymerase (5); (7) a DNA polymerase for amplifying single-stranded nucleic acid fragments; (8) a suitable buffer for the DNA polymerase (7); (9) dNTPs; (10) a modified dNTP; (11) a control template and/or control oligonucleotide primers for amplifying a template nucleic acid; (12) a chromatography column; (13) Buffer A for the chromatography column; (14) Buffer B for the chromatography column; (15) water; (16) a strand displacement facilitator (*e.g.*, 1M trehalose); (17) microtiter plates or microwell plates; (18) oligonucleotide standards (*e.g.*, 6mer, 7mer, 8mer, 12mer and 16mer) for liquid chromatography and/or mass spectrometry; (19) an instruction booklet for using the kit; and (20) an access code for a software used in designing and/or ordering ODNP pairs. Detailed descriptions of many of the above components have been provided above.

An exemplary nicking agent is a nicking endonuclease N.BstNB I. The buffer for this nicking endonuclease may be 10 mM Tris-HCl, 10 mM $MgCl_2$, 150 mM KCl, 1 mM dithiothreitol (pH 7.5 at 25°C). N.BstNB I may be stored in the following storage buffer: 50 mM KCl, 10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 1 mM DTT, 200 μg/ml BSA, and 50% glycerol.

An exemplary 5′→3′ exonuclease deficienty DNA polymerase is exo$^-$ Vent. The reaction buffer for this polymerase may be 10 mM KCl, 20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100.

As discussed above, Buffer A comprises an ammonium salt(s) dissolved in water, where the water should preferably be HPLC grade water. The ammonium salt may be any of the one or more ammonium salts identified above. For example, in one aspect, the ammonium salt is N,N-dimethyl-N-butylammonium with a counterion, where, in a further aspect, the counterion is acetate. In a preferred embodiment the ammonium salt is present in the water at a concentration of about 1-200 mM, preferably

about 1-100 mM, preferably about 5-50 mM, and more preferably about 5 mM. The container may hold ammonium salt at a concentration greater than 100 mM, but in such case the components in this container will probably need to be diluted with water in order to form an effective buffer. This solution of ammonium salt and water preferably has a pH of about 7.0-7.5, and is typically about 7.2, and is referred to herein as Buffer A. In one aspect, this container holds only ammonium salt dissolved in water.

Also as discussed above, Buffer B comprises ammonium salt, water and organic solvent. In one aspect, the components comprise Buffer A and organic solvent, where the mixture is preferably homogeneous, *i.e.,* the organic solvent dissolves or is miscible in Buffer A. In one aspect the organic solvent is acetonitrile. In another aspect the organic solvent is methanol. In various aspects, the organic solvent constitutes, on a volume percent basis, based on the total volume of the components in the container, 15%-20%, 20%-25%, 25%-30%, 30%-35%, 35%-40%, 40%-45%, 45%-50% or 50%-55% or 55-60% or 65-70% or 75-80%. In a preferred embodiment the organic solvent constitutes 25% acetonitrile. In another preferred embodiment the organic solvent constitutes 50% acetonitrile. Thus, a preferred component is a solution of 75 vol% Buffer A and 25 vol% acetonitrile, while another preferred component is a solution of 50 vol% Buffer A and 50 vol% acetonitrile.

The chromatography column is preferably a reverse phase chromatography column, more preferably a C18 reverse phase chromatography column. A preferred C18 reverse phase column may have one, two or all of the following features: a pore size of at least 120Å, a particle size of 2 microns to 10 microns, and 0.3 mm to 4.6 mm in inner diameter and from 10 mm to 250 mm in length.

3. Applications for Genetic Variation Identification Methods of the Present Invention

As discussed in detail herein above, the present invention provides methods for detecting and/or identifying genetic variations in target nucleic acids. Also provided herein, are various "readout" technologies that may be employed with the methodologies of the present invention for detecting, for example, the size and/or molecular weight of one or more single-stranded fragment(s) comprising the mutations and/or genetic variations. Methods according to the present invention will find utility in a wide variety of applications wherein it is desirable or necessary to identify such a mutation at a defined nucleotide locus or measure genetic variations. Such applications include, but are not limited to, genetic analysis for hereditarily transferred diseases, tumor diagnosis, disease predisposition, forensics, paternity determination,

enhancements in crop cultivation or animal breeding, expression profiling of cell function and/or disease marker genes, and identification and/or characterization of infectious organisms that cause infectious diseases in plants or animals and/or that are related to food safety. Furthermore, the present methods, compositions and kits may be utilized to greatly increase the specificity, sensitivity and throughput of the assay while lowering costs in comparison to conventional methods currently available in the art. Described below are certain exemplary applications of the present invention.

a. Expression Profiling

Most mRNAs are transcribed from single copy sequences. Another property of cDNAs is that they represent a longer region of the genome because of the introns present in the chromosomal version of most genes. The representation varies from one gene to another but can be very significant as many genes cover more than 100 kb in genomic DNA, represented in a single cDNA. One possible use of molecular profiling is the use of probes from one species to find clones made from another species. Sequence divergence between the mRNAs of mouse and man permits specific cross-reassociation of long sequences, but except for the most highly conserved regions, prevents cross-hybridization of PCR primers.

Differential screening in complex biological samples such as developing nervous system using cDNA probes prepared from single cells is now possible due to the development of PCR-based and cRNA-based amplification techniques. Several groups reported previously the generation of cDNA libraries from small amounts of poly (A)+ RNA (1 ng or less) prepared from 10-50 cells (Belyav et al., *Nuc. Acids Res. 17*:2919, 1989). Although the libraries were sufficiently representative of mRNA complexity, the average cDNA insert size of these libraries was quite small (<2 kb).

More recently, methodologies have been combined to generate both PCR-based (Lambolez et al., *Neuron 9*:247, 1992) and cRNA-based (Van Gelder et al., *Proc. Natl. Acad. Sci. USA 87*:1663, 1990) probes from single cells. After electrical recordings, the cytoplasmic contents of a single cell were aspirated with patch-clamp microelectrodes for *in situ* cDNA synthesis and amplification. PCR was used to amplify cDNA of selective glutamate receptor mRNAs from single Purkinje cells and GFAP mRNA from single glia in organotypic cerebellar culture (Lambolez et al., *Neuron 9*:247, 1992). In the case of cRNA amplification, transcription promoter sequences were designed into primers for cDNA synthesis and complex antisense cRNAs were generated by in vitro transcription with bacteriophage RNA polymerases.

The methods of the present invention are useful for determining whether a particular cDNA molecule is present in cDNAs from a biological sample and further determine whether genetic variation(s) exist in the cDNA molecule.

b. Forensics

The identification of individuals at the level of DNA sequence variation offers a number of practical advantages over such conventional criteria as fingerprints, blood type, or physical characteristics. In contrast to most phenotypic markers, DNA analysis readily permits the deduction of relatedness between individuals such as is required in paternity testing. Genetic analysis has proven highly useful in bone marrow transplantation, where it is necessary to distinguish between closely related donor and recipient cells. Two types of probes are now in use for DNA fingerprinting by DNA blots. Polymorphic minisatellite DNA probes identify multiple DNA sequences, each present in variable forms in different individuals, thus generating patterns that are complex and highly variable between individuals. VNTR probes identify single sequences in the genome, but these sequences may be present in up to 30 different forms in the human population as distinguished by the size of the identified fragments. The probability that unrelated individuals will have identical hybridization patterns for multiple VNTR or minisatellite probes is very low. Much less tissue than that required for DNA blots, even single hairs, provides sufficient DNA for a PCR-based analysis of genetic markers. Also, partially degraded tissue may be used for analysis since only small DNA fragments are needed. The methods of the present invention are useful in characterizing polymorphism of sample DNAs, therefore useful in forensic DNA analyses. For example, the analysis of 22 separate gene sequences in a sample, each one present in two different forms in the population, could generate 1010 different outcomes, permitting the unique identification of human individuals.

c. Tumor Diagnostics

The detection of viral or cellular oncogenes is another important field of application of nucleic acid diagnostics. Viral oncogenes (v-oncogenes) are transmitted by retroviruses while their cellular counterparts (c-oncogenes) are already present in normal cells. The cellular oncogenes can, however, be activated by specific modifications such as point mutations (as in the c-K-ras oncogene in bladder carcinoma and in colorectal tumors), small deletions and small insertions. Each of the activation processes leads, in conjunction with additional degenerative processes, to an increased and uncontrolled cell growth. In addition, point mutations, small deletions or insertions

may also inactivate the so-called "recessive oncogenes" and thereby leads to the formation of a tumor (as in the retinoblastoma (Rb) gene and the osteosarcoma). Accordingly, the present invention is useful in detecting or identifying the point mutations, small deletions and small mutations that activate oncogenes or inactivate recessive oncogenes, which in turn, cause cancers.

d. Transplantation Analyses

The rejection reaction of transplanted tissue is decisively controlled by a specific class of histocompatibility antigens (HLA). They are expressed on the surface of antigen-presenting blood cells, *e.g.*, macrophages. The complex between the HLA and the foreign antigen is recognized by T-helper cells through corresponding T-cell receptors on the cell surface. The interaction between HLA, antigen and T-cell receptor triggers a complex defense reaction which leads to a cascade-like immune response on the body.

The recognition of different foreign antigens is mediated by variable, antigen-specific regions of the T-cell receptor—analogous to the antibody reaction. In a graft rejection, the T-cells expressing a specific T-cell receptor which fits to the foreign antigen, could therefore be eliminated from the T-cell pool. Such analyses are possible by the identification of antigen-specific variable DNA sequences which are amplified by PCR and hence selectively increased. The specific amplification reaction permits the single cell-specific identification of a specific T-cell receptor.

Similar analyses are presently performed for the identification of auto-immune disease like juvenile diabetes, arteriosclerosis, multiple sclerosis, rheumatoid arthritis, or encephalomyelitis.

Accordingly, the present invention is useful for determining gene variations in T-cell receptor genes encoding variable, antigen-specific regions that are involved in the recognition of various foreign antigens.

e. Genome Diagnostics

Four percent of all newborns are born with genetic defects; of the 3,500 hereditary diseases described which are caused by the modification of only a single gene, the primary molecular defects are only known for about 400 of them.

Hereditary diseases have long since been diagnosed by phenotypic analyses (anamneses, *e.g.*, deficiency of blood: thalassemias), chromosome analyses (karyotype, *e.g.*, mongolism: trisomy 21) or gene product analyses (modified proteins, *e.g.*, phenylketonuria: deficiency of the phenylalanine hydroxylase enzyme resulting in

enhanced levels of phenylpyruvic acid). The additional use of nucleic acid detection methods considerably increases the range of genome diagnostics.

In the case of certain genetic diseases, the modification of just one of the two alleles is sufficient for disease (dominantly transmitted monogenic defects); in many cases, both alleles must be modified (recessively transmitted monogenic defects). In a third type of genetic defect, the outbreak of the disease is not only determined by the gene modification but also by factors such as eating habits (in the case of diabetes or arteriosclerosis) or the lifestyle (in the case of cancer). Very frequently, these diseases occur in advanced age. Diseases such as schizophrenia, manic depression or epilepsy should also be mentioned in this context; it is under investigation if the outbreak of the disease in these cases is dependent upon environmental factors as well as on the modification of several genes in different chromosome locations.

Using direct and indirect DNA analysis, the diagnosis of a series of genetic diseases has become possible: bladder carcinoma, colorectal tumors, sickle-cell anemia, thalassemias, al-antitrypsin deficiency, Lesch-Nyhan syndrome, cystic fibrosis/mucoviscidosis, Duchenne/Becker muscular dystrophy, Alzheimer's disease, X-chromosome-dependent mental deficiency, and Huntington's chorea, phenylketonuria, galactosemia, Wilson's disease, hemochromatosis, severe combined immunodeficiency, alpha-1-antitrypsin deficiency, albinism, alkaptonuria, lysosomal storage diseases, Ehlers-Danlos syndrome, hemophilia, glucose-6-phosphate dehydrogenase disorder, agammaglobulimenia, diabetes insipidus, Wiskott-Aldrich syndrome, Fabry's disease, fragile X-syndrome, familial hypercholesterolemia, polycystic kidney disease, hereditary spherocytosis, Marfan's syndrome, von Willebrand's disease, neurofibromatosis, tuberous sclerosis, hereditary hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos syndrome, myotonic dystrophy, osteogenesis imperfecta, acute intermittent porphyria, and von Hippel-Lindau disease. The present invention is useful in diagnosis of any genetic diseases that are caused by point mutations, small deletions or small insertions at defined positions.

f. Infectious Disease

The application of recombinant DNA methods for diagnosis of infectious diseases has been most extensively explored for viral infections where current methods are cumbersome and results are delayed. *In situ* hybridization of tissues or cultured cells has made diagnosis of acute and chronic herpes infection possible. Fresh and fomalin-fixed tissues have been reported to be suitable for detection of papillomavirus in invasive cervical carcinoma and in the detection of HIV, while cultured cells have

been used for the detection of cytomegalovirus and Epstein-Barr virus. The application of recombinant DNA methods to the diagnosis of microbial diseases has the potential to replace current microbial growth methods if cost-effectiveness, speed, and precision requirements can be met. Clinical situations where recombinant DNA procedures have begun to be applied include the identification of penicillin-resistant Neisseria gonorrhoeae by the presence of a transposon, the fastidiously growing chlamydia, microbes in foods; and simple means of following the spread of an infection through a population. The worldwide epidemiological challenge of diseases involving such parasites as leishmania and plasmodia is already being met by recombinant methods.

The present invention is useful to detect and/or measure genetic variations that are involved in infectious diseases, especially those in drug resistance genes. Thus, the present invention facilitates the characterization and classification of organisms that cause infectious diseases and consequently the treatment of such diseases caused by these organisms.

g. Additional Uses

In addition to the above applications, the present invention may also be used in testing disease susceptibility. Certain gene variations, although they do not directly cause diseases, are associated to the diseases. In other words, the possession of the gene variations by a subject renders the subject susceptible to the diseases. The detection of such gene variations using the present methods enables the identification of the subjects that are susceptible to certain diseases and subsequent performance of preventive measures.

The present invention is also applicable to pharmocogenomics. For instance, it may be used to detect or identify genes that involve in drug tolerance, such as various alleles of cytochrome P450 gene.

In addition, the present invention provides methods useful for detecting or characterizing residual diseases. In other words, the present methods may be used for detecting or identifying remaining mutant genotypes as in cancer after certain treatments, such as surgery of chemotherapy. It may also useful in identifying emerging mutants, such as genetic variations in certain genes that render a pathogenic organism drug resistant.

C. Detection of Nucleic Acids in Biological Sample

In various aspects, the present invention provides methods, compounds, compositions, and kits for detecting the presence, or the absence, of a target nucleic acid in a biological sample.

1. Methods for Detecting Target Nucleic Acids

The present invention provides methods for detecting the presence, or the absence, of a target nucleic acid in a biological sample. In certain embodiments, the methods utilize a partially double-stranded nucleic acid probe (described in more detail below) that comprises (1) an overhang that is at least substantially complementary to a target nucleic acid, (2) a NARS, and (3) a sequence that uniquely correlates to the target nucleic acid. A sequence that "uniquely correlates to" a target nucleic acid refers to a sequence that is present only in a partially double-stranded nucleic acid probe specific for a particular target nucleic acid of interest, but is absent in a partially double-stranded nucleic acid probe specific for another target nucleic acid of interest. A partially double-stranded nucleic acid probe is "specific" for a target nucleic acid of interest if its overhang is at least substantially complementary to the target nucleic acid. In certain embodiments, the overhang of the probe is exactly complementary to a portion of the target. A target nucleic acid of interest, in certain embodiments (*e.g.*, where the detection of closely related nucleic acid molecules are intended as described below), may be a nucleic acid with a degenerate sequence that represents various nucleic acid molecules that are substantially identical to each other.

According to the present invention, the probe described above is combined with a biological sample that may contain the target nucleic acid under conditions that allow for hybridization of the probe to the target nucleic acid, if present, in the biological sample. The unhybridized probe is removed from a reaction mixture, while the hybridized probe is used as a template to amplify a single-stranded nucleic acid in the presence of a nicking agent that recognizes the NARS. Because the amplified single-stranded nucleic acid fragment has a sequence that is complementary to the sequence in the probe that uniquely correlates to the target nucleic acid, the detection and/or characterization of the single-stranded nucleic acid fragment indicates the presence of the target nucleic acid in the sample.

In other embodiments, the present invention uses a single-stranded nucleic acid probe or various ODNP pairs to prepare template nucleic acids that comprise a NARS. The single-stranded nucleic acid probe and ODNP pairs are specific to a target nucleic acid and can be used to prepare the template nucleic acids only when

the target nucleic acid is present in a biological sample. The template nucleic acid may be then used to amplify a single-stranded nucleic acid in the presence of a nicking agent that recognizes the NARS. The detection and/or characterization of the amplified single-stranded nucleic acid indicates the presence of the target nucleic acid in the sample.

The above methods may be multiplexed to detect the presence, or the absence, of multiple target nucleic acids in a sample. For instance, multiple partially double-stranded nucleic acid probes may be combined with the sample. Each of the multiple probes has a nucleotide sequence within the strand that does not contain the NS, where the nucleotide sequence is located 5′ to the position corresponding to the NS and uniquely correlates to a particular target nucleic acid. Such a sequence is used as a template for the amplification of a single-stranded nucleic acid fragment in the presence of (i) a NA that recognizes the NARS and (ii) a DNA polymerase. The detection/characterization of the single-stranded nucleic acid fragment indicates the presence of the particular target nucleic acid to which a portion of the probe uniquely correlates.

a. Biological Samples and Target Nucleic Acids

Biological samples of the present invention include any sample that originates from an organism and that may contain a nucleic acid of interest (*i.e.*, target nucleic acid). They may be provided by obtaining a blood sample, biopsy specimen, tissue explant, organ culture or any other tissue or cell preparation from a subject or a biological source. Biological samples also include tissue prints made onto a solid support. The subject or biological source may be a human or non-human animal, a plant, a primary cell culture or culture adapted cell line. In certain preferred embodiments of the invention, the subject or biological source may be suspected of having, or being at risk for having, a genetic disease or a pathogen infection. In other preferred embodiments, the subject or biological source may be a patient that has a genetic disease or a pathogen infection. In certain other embodiments, the subject or biological source may be a control subject that does not have a genetic disease or a pathogen infection.

Target nucleic acids may be any nucleic acid present in a biological sample. They can be DNA or RNA, single-stranded or double-stranded. In certain preferred embodiments, the target nucleic acid originates from a pathogenic organism such as a bacterium, a virus, a fungus or a parasite. Exemplary pathogenic organisms are provided below in the descriptions for single-stranded nucleic acid probe

preparation. The target nucleic acid may have a sequence common to a particular species or subspecies, but not present in other species or subspecies. Detection of such a target nucleic acid indicates the presence of the species or subspecies in a biological sample and can be used as an initial screen for a particular organism or particular organisms within the species or subspecies. The biological sample may be further characterized to determine the presence or the absence of the particular organisms in the sample. In other preferred embodiments, the target nucleic acid is an mRNA of which expression is associated with a disease.

In certain embodiments, nucleic acids of a biological sample are immobilized before being hybridized with partial double-stranded or single-stranded oligonucleotide probes of the present invention. The immobilization of target nucleic acid allows easy removal of non-hybridized probes from hybridization mixture before the amplification of single-stranded nucleic acid molecules. Techniques for the immobilization of target nucleic acids are well known, and the immobilization may be accomplished by any method known in the art. For instance, a biological sample may be fixed with a fixative that maintains the morphological integrity of the cells in the sample but does not cross-link or precipitate cellular proteins so extensively that penetration of nucleic acid probes and other reagents is prevented. Such fixation allows *in situ* detection of the target nucleic acid. Alternatively, nucleic acids of a biological sample may be transferred to a solid support (*e.g.*, nitrocellulose membranes) using techniques known in the art, such as tissue printing. In certain embodiments, the target nucleic acid may be first isolated from the biological sample and subsequently immobilized to a solid support. The solid support may be made of various materials, including but not limited to, silicon, glass, paper, ceramic, metal (*e.g.*, stainless steel), metalloid and plastic. The solid support may be further coated with a chemical layer that is bindable to nucleic acid molecules so that it is further coated with the nucleic acid molecules isolated from the biological sample via the chemical layer when the coated solid support is in contact with those nucleic acid molecules. Exemplary chemicals that may form a suitable chemical layer for binding with nucleic acid molecules include, but are not limited to, PADMAC and poly(ethyleneimine) (PEI). Methods for solid support coated with the suitable chemical layer are known (*see, e.g.*, U.S. Pat. No. 6,150,103; U.S. Pat. Appl. No. 09/120,689, now allowed, both incorporated herein by reference in their entireties).

In certain preferred embodiments, the solid support coated with a chemical layer for binding with nucleic acids is in a pin shape. Such solid support may be prepared according to U.S. Patent Appl. No. 09/323,695 (now pending), incorporated

herein by reference in its entirety. This type of solid support allows the automation of the diagnostic methods described herein. For instance, multiple pins bearing nucleic acid molecules isolated from a single biological sample may be simultaneously put into individual wells of a multi-well plate where each well contains a partially double-stranded or single-stranded nucleic acid probe specific for detecting a particular target nucleic acid molecule as well as other components necessary for single-stranded nucleic acid amplification (*e.g.*, a nicking agent, a DNA polymerase and dNTPs). Upon the completion of the amplification reaction, the pins may be removed from the wells and the amplified single-stranded nucleic acid molecules in the well be characterized in parallel. Such characterization of the amplified nucleic acid molecule allows the determination of the presence, or the absence, of particular target nucleic acids in the biological sample.

Likewise, multiple pins bearing nucleic acid molecules isolated from different biological samples may be simultaneously put into individual wells of a multi-well plate where each well contains a same partially double-stranded or single-stranded nucleic acid probe specific for detecting a particular target nucleic acid molecule as well as other components necessary for single-stranded nucleic acid amplification (*e.g.*, a nicking agent, a DNA polymerase and dNTPs). Upon the completion of the amplification reaction, the pins may be removed from the wells and the amplified single-stranded nucleic acid molecules in the well be characterized in parallel. Such characterization of the amplified nucleic acid molecule allows the determination of the presence, or the absence, of the particular target nucleic acid in various biological samples.

b. Mehods for Preparing Template Nucleic Acids

Template nucleic acid molecules useful for diagnostic applications may be provided by various approaches. For instance, templates may be obtained by annealing of a trigger nucleic acid to a single-stranded nucleic acid probe where the trigger nucleic acid is derived from a nucleic acid molecule originated from a pathogenic organism. Alternatively, a template may be directly derived from a double-stranded nucleic acid molecule originated from a pathogenic organism. A template may also be a partially double-stranded nucleic acid molecule having an overhang derived from a target nucleic acid and functioning as a template for single-stranded nucleic acid amplification, or an overhang capable of hybridizing with a target nucleic acid but not functioning as a template for single-stranded nucleic acid amplification. These and

other means for providing templates relevant to diagnostic applications are described below.

i. First Type of Exemplary Methods for Providing Template Molecules

In certain embodiments of the present invention where a template nucleic acid molecule is provided by annealing a trigger nucleic acid to a single-stranded nucleic acid probe, the trigger nucleic acid may be derived from either a DNA molecule (*e.g.*, a genomic DNA molecule) or a RNA molecule (*e.g.*, a mRNA molecule) originated from a pathogenic organism. If the nucleic acid molecule originated from a pathogenic organism is single-stranded, it may be directly used as a trigger nucleic acid. Alternatively, the single-stranded nucleic acid may be cleaved to produce shorter fragments, where one or more of these fragments may to be used as a trigger nucleic acid. If the nucleic acid molecule originated from a pathogenic organism is double-stranded, it may be denatured and directly used as a trigger nucleic acid or the denatured product may be cleaved to provide multiple shorter single-stranded fragments where one or more of these fragments may function as a trigger nucleic acid. Alternatively, it may be first cleaved to obtain multiple shorter double-stranded fragments, and the shorter fragments are then denatured to provide one or more trigger nucleic acids.

A single-stranded nucleic acid probe must be at least substantially complementary to a trigger nucleic acid so that when a target nucleic acid from which the trigger nucleic acid is derived is present in a biological sample, it is capable of annealing to the target. In addition, the probe comprises a sequence of one strand of a NARS, so that the extension product of the duplex formed when the trigger nucleic acid anneals to the probe comprises a nicking agent recognition sequence.

An example of the first type of methods for preparing template molecules is shown in Figure 36. As indicated in this figure, a double-stranded genomic DNA may be first cleaved by a restriction endonuclease. The digestion products may be denatured and one strand of one of the digestion products may be used as a trigger nucleic aicd to initiate nucleic acid amplification reactions.

ii. Second Type of Exemplary Methods for Providing Template Molecules

In certain embodiments of the present invention where a template is provided by annealing a trigger nucleic acid to a single-stranded nucleic acid probe, the trigger nucleic acid may comprise the sequence of the sense strand of a NARS. The

trigger nucleic acid may be derived from a target nucleic acid (*e.g.*, a genomic nucleic acid) originated from a pathogenic organism. A specific embodiment where a template nucleic acid comprises a NERS recognizable by a nicking endonuclease that nicks outside its recognition sequence (*e.g.*, N.BstNB I) is illustrated in Figure 37. As illustrated by this figure, a genomic DNA or a fragment thereof comprising a NERS is denatured, and one strand of the genomic DNA or a fragment of that strand anneals to a single-stranded nucleic acid probe. The probe is a portion of the other strand of the genomic DNA that comprises a sequence of the antisense strand of the NERS. The annealing of the trigger nucleic acid to the probe provides the template nucleic acid for amplification reactions. The number of the probe molecules in an amplification reaction mixture is preferably greater than the number of strands of genomic DNA or fragments thereof that contain the sequence of the sense strand of the NERS.

In related embodiments where the trigger nucleic acid is derived from a target nucleic acid and comprises the sequence of the sense strand of a NARS, a single-stranded nucleic acid probe may be at least substantially complementary to the trigger nucleic acid at the 3' portion, but not at the 5' portion, of the probe. The 3' portion of the probe includes the sequence of the antisense strand of the NARS so that the template nucleic acid formed by annealing the probe to the trigger nucleic acid comprises a double-stranded NARS. In the presence of a NA that recognizes the NARS, the template molecule is nicked. The 3′ terminus at the nicking site is then extended using a region 5′ to the sequence of the antisense strand of the NARS in the template molecule as a template. The resulting amplification product is a single-stranded nucleic acid molecule that is complementary to a region of the probe located 5′ to the sequence of the antisense strand of the NARS rather than a portion of the trigger nucleic acid.

In certain embodiments, the single-stranded nucleic acid probe may be immobilized to a solid support. Alternatively, the single-stranded nucleic acid probe may not be attached to any solid support.

iii. Third Type of Exemplary Methods for Providing Template Molecules

In certain embodiments of the present invention, a template nucleic acid is a double-stranded nucleic acid derived directly from a genomic nucleic acid that contain both a NARS and a RERS, where the NS corresponding to the NARS lies within the NARS or between the NARS and the RERS, and the RERS is located near the NARS. A RERS is located near a NARS when the distance between the RERS and the NARS is at most 500, 400, 300, 200, 100, 50, 40, 30, 20, 15 or 10 base pairs. An

embodiment with a NERS recognizable by a nicking endonuclease that nicks outside its recognition sequence (*e.g.*, N.BstNB I) as an exemplary NARS is illustrated in Figure 38. As shown in this figure, genomic DNA may be digested by a restriction endonuclease that recognizes a RERS in the genomic DNA. The digestion product that contains the NERS may function as a template nucleic acid.

iv. Fourth Type of Exemplary Methods for Providing Template Molecules

In certain embodiments of the present invention, a template nucleic acid is a completely or partially double-stranded nucleic acid molecule produced using various ODNP pairs. The methods for using ODNP pairs to prepare template molecules are described below in connection with Figures 39-41, which are similar to those in Figures 1, 2 and 5.

In one embodiment, a precursor to a template nuleic acid contains a double-stranded NARS and a RERS. The NARS and RERS are incorporated into the precursor using an ODNP pair. An embodiment with a NERS recognizable by a NE that nicks outside its recognition sequence (*e.g.*, N.BstNB I) as an exemplary NARS, and a type IIs restriction endonuclease recognition sequence (TRERS) as an exemplary RERS is illustrated in Figure 39. As shown in this figure, a first ODNP comprises the sequence of one strand of a NERS while a second ODNP comprises the sequence of one strand of a TRERS. When these two ODNPs are used as primers to amplify a portion of a target nucleic acid, the resulting amplification product (*i.e.*, a precursor to a template nucleic acid), contains both a double-stranded NERS and a double-stranded TRERS. In the presence of a type IIs restriction endonuclease that recognizes the TRERS, the amplification product is digested to produce a template nucleic acid molecule that comprises a double-stranded NERS.

In another embodiment, a precursor to a template nucleic acid contains two double-stranded NARSs. The two NARSs are incorporated into the precursor to the template using two ODNPs. An embodiment with a NERS recognizable by a nicking endonuclease that nicks outside its recognition sequence as an exemplary NARS is illustrated in Figure 40. As shown in this figure, both ODNPs comprise a sequence of a sense strand of a NERS. When these two ODNPs are used as primers to amplify a portion of a target nucleic acid, the resulting amplification product contains two NERSs. These two NERSs may or may not be identical to each other, but preferably, they are identical. In the presence of a NE or NEs that recognize the NERSs, the amplification product is nicked twice (once on each strand) to produce two template nucleic acids that each comprises a double-stranded NERS.

In yet another embodiment, a precursor to a template nucleic acid contains two hemimodified RERS. The two hemimodified RERSs are incorporated into the precursor by the use of two ODNPs. This embodiment is illustrated in Figure 41. As shown in this figure, both the first and the second ODNPs comprise a sequence of one strand of a RERS. When these two ODNPs are used as primers to amplify a portion of a target nucleic acid in the presence of a modified deoxynucleoside triphosphate, the resulting amplification product contains two hemimodified RERSs. These two hemimodified RERS may or may not be identical to each other. In the presence of a RE or REs that recognize the hemimodified RERS, the above amplification product is nicked to produce two partially double-stranded template nucleic acid molecule that each comprises a sequence of at least one strand of the hemimodified RERS.

v. Fifth Type of Exemplary Methods for Providing Template Molecules

In other embodiments of the present invention, a partially double-stranded nucleic acid probe may function as a template nucleic acid molecule. The probe comprises (1) a nicking agent recognition sequence (NARS) cleavable by a nicking agent to produce a nicking site (NS); (2) either a 5′ overhang in the strand that either the strand itself or an extension product thereof contains the NS or a 3′ overhang in the strand that neither the strand itself nor an extension product thereof contains the NS, wherein the overhang comprises a nucleotide sequence at least partially complementary to the target nucleic acid; and (3) a nucleotide sequence within the strand that neither the strand or an extension product thereof contains the NS, the nucleotide sequence located 5′ to the position corresponding to the NS and uniquely correlating to the target nucleic acid. An exemplary embodiment where a template nucleic acid has a NERS recognizable by a nicking endonuclease that nicks outside its recognition sequence as an exemplary NARS is illustrated in Figure 6. The probe is combined with a biological sample that may contain the target nucleic acid under conditions that allow for hybridization of the probe to the target nucleic acid, if present, in the sample. The unhybridized probe is removed, while the hybridized probe is used as a template in amplifying a single-stranded nucleic acid in the presence of a nicking agent that recognizes the NARS. Because the amplified single-stranded nucleic acid fragment has a sequence that is complementary to the sequence in the probe that uniquely correlates to the target nucleic acid, the detection and/or characterization of the amplified single-stranded nucleic acid indicates the presence of the target nucleic acid in the sample.

The removal of unhybridized probes may be facilitated by immobilizing nucleic acid molecules of a biological sample. Such immobilization may be performed by any method known in the art as described above. When a partially double-stranded nucleic acid probe having an overhang that is substantially complementary to a particular target nucleic acid molecule is applied to the sample, the probe hybridizes to the target nucleic acid via its overhang if the target nucleic acid is present in the sample. The sample is subsequently washed to remove any unhybridized probe. In the presence of a DNA polymerase and nicking agent that recognizes the NARS in the probe, a single-stranded nucleic acid molecule is amplified. However, if the target nucleic acid is absent in the sample, the probe is unable to hybridize to any nucleic acid molecule in the sample and thus is washed off from the sample. Thus, when the washed biological sample is incubated with a nucleic acid amplification reaction mixture (*i.e.*, a mixture containing all the necessary components for single strand nucleic acid amplification using a portion of the probe as a template, such as a NA that recognizes the NARS in the probe and a DNA polymerase), no single-stranded nucleic acid molecule that is complementary to the above portion of the probe is amplified. Thus, template molecules are only formed when N1 hybridizes to a specific immobilized sequence.

Besides immobilizing a target nucleic acid molecule, probe molecules that have annealed to a target nucleic acid may be separated from those that have not annealed to the target nucleic acid by a functional group associated with the target nucleic acid. For instance, the target nucleic acid may be labeled with a biotin molecule, and probe molecules that have annealed to the target nucleic acid may be separated from those that have not annealed to the target nucleic acid via the biotin molecule associated with the target. For instance, the target nucleic acid, as well as the probe that has annealed to the target, may be precipitated with immobilized streptavidin.

The NARS of the probe can be a recognition sequence of any nicking agent, including but not limited to, a NERS and a hemimodified RERS. The NS may be within the NARS or may be outside the NARS. In certain embodiments, the NS may be absent in the probe, but may be generated by extending the probe in the presence of a DNA polymerase. In a preferred embodiment, the NARS is recognizable by nicking endonuclease N.BstNB I.

The double-stranded portion of the probe must be long enough to allow both strands to be able to anneal with each other under the condition for removing nonhybridized probe from the sample or for amplifying a single-stranded nucleic acid fragment in the presence of a NA and a DNA polymerase. This portion may be at least

8, 10, 12, 15, 20 nucleotides in length. It contains a NARS and may or may not contain a corresponding nicking site.

The overhang of the probe (either a 5′ overhang in the strand that contains the NS or the extension product of which contains the NS, or a 3′ overhang in the other strand) also must be long enough to allow specific hybridization between the overhang and the target nucleic acid. Such hybridization should not be dissociated under the condition for removing nonhybridized probe from the sample or for subsequently amplifying a single-stranded nucleic acid. The overhang may be at least 6, 8, 10, 12, 14, 16, 18 or 20 nucleotides in length.

The probe must also contain a nucleotide sequence within the strand that either that strand itself or an extension product thereof does not contain the NS, where this nucleotide sequence is located 5′ to the position corresponding to the NS. This sequence will be used as a template for the amplification of a single-stranded nucleic acid fragment in the presence of a nicking agent and a DNA polymerase. It can be within a double-stranded portion or a single-stranded portion of the probe. When multiple probes are used to detect the presence, or the absence, of multiple target nucleic acids of interest, the above sequence of each probe must be different from each other so that each can uniquely correlate to a particular target nucleic acid.

The probe of the present invention may be synthesized using any applicable method known in the art. Each strand may be individually synthesized and subsequently annealed to form a partially double-stranded nucleic acid probe.

In certain related embodiments, a template nucleic acid is formed by hybridizing an immobilized target nucleic acid from a biological sample with a single-stranded nucleic acid probe. An example of these embodiments using a NARS recognizable by a nicking agent that nicks outside the NARS is illustrated in Figure 42. As shown in this figure, nucleic acids of a biological sample are immobilized via their 5' termini. The resulting immobilized nucleic acids are then hybridized with a single-stranded probe. The single-stranded probe comprises, from 3' to 5', a sequence that is at least substantially complementary to a target nucleic acid suspected to be present in the biological sample and a sequence of the antisense strand of a NARS. If the target nucleic acid is present in the biological sample, the probe molecule hybridizes to the target nucleic acid to form a template molecule. The template molecule is separated from unhybridized probe molecule by washing the solid phase to which the target nucleic acid is attached. In the presence of a DNA polymerase and a nicking agent that recognizes the NARS, the template nucleic acid is used as a template to amplify a single-stranded nucleic acid molecule. However, if the target nucleic acid is absent in

the sample, the single-stranded probe is unable to hybridize to any nucleic acid molecule in the sample and thus is washed off from the solid support. Consequently, no template nucleic acid can be formed that attaches to the solid support, and no single-stranded nucleic acid molecule complementary to a portion of the template nucleic acid can be amplified.

In other related embodiments, a template nucleic acid is formed by hybridizing a target nucleic acid from a biological sample with an immobilized single-stranded nucleic acid probe. An example of these embodiments is where a target nucleic acid is not immobilized, but a single-stranded nucleic acid probe as described above is immobilized to a solid support via its 5' terminus. If a target nucleic acid is present in a sample, the hybridization of the nucleic acids of the sample to the probe allows the target to remain attached to the solid support when the solid support is washed. In the presence of a nicking agent that recognizes the nicking agent recognition sequence of which the antisense strand is present in the probe and a DNA polymerase, a single-stranded nucleic acid molecule is amplified using a sequence located 5' to the sequence of the antisense strand of the recognition sequence in the probe as a template. If the target is absent in the sample, the nucleic acids of the sample will be washed off the solid support to which the probe is attached. Thus, no single-stranded nucleic acid molecule is amplified using a portion of the probe as a template.

Another example of the above embodiments using a NARS recognizable by a nicking agent that nicks outside the NARS is illustrated in Figure 44. As shown in this figure, a single-stranded nucleic acid probe is immobilized to a solid support via its 5' terminus. The probe comprises, from 5' to 3', a sequence of the sense strand of the NARS and a sequence that is substantially complementary to the 3' portion of the target nucleic acid. The probe is mixed with the nucleic acids from a biological sample. If the target nucleic acid is present in the sample, the probe molecule is hybridized to the target to form a template molecule. When the solid support to which the probe is attached is washed, the target remains attached to the solid support via its hybridization with the probe. In the presence of a DNA polymerase, the target extends from its 3' terminus using the probe as a template. The duplex formed between the extension product of the target and that of the probe comprises a double-stranded NARS. In the presence of a nicking agent that recognizes the NARS as well as the DNA polymerase, a single-stranded nucleic acid molecule is amplified using a portion of the target nucleic acid as a template. However, if the target nucleic acid is absent in the sample, the probe will not be able to hybridize with the target. Thus, no single-stranded nucleic acid molecule will be amplified using the target as a template.

Another example of the above embodiments is illustrated in Figure 45. In this example, the immobilized single-stranded nucleic acid probe is substantially complementary to the target nucleic acid, but not necessarily complementary to the 3' portion of the target. The probe also comprises a sequence of the sense strand of a nicking agent recognition sequence. If the target is present in a biological sample, when the probe is mixed with the nucleic acids in the sample, it may hybridize with the target. When the solid support to which the probe is attached is washed, the target remains attached to the solid support via its hybridization with the probe. In the presence of a DNA polymerase, and a nicking agent that recognizes the NARS, even when one or more nucleotides in the sequence of the sense strand of the NARS may not form conventional base pairs with nucleotides in the target, in certain circumstances, a single-stranded nucleic acid may be amplified using a portion of the target as a template. The detailed descriptions for the circumstances where a single-stranded nucleic acid is amplified when a template nucleic acid does not comprise a double-stranded NARS are provided below in the sections related to preparing single-stranded nucleic acid molecules. However, if the target nucleic acid is absent in the sample, the probe will not be able to hybridize with the target. Thus, no single-stranded nucleic acid molecule will be amplified using the target as a template.

c. Specificity

The methods of the present invention may be used for detecting the presence or absence of a particular pathogenic organism in a sample, as well as for detecting the presence of several closely related pathogenic organisms. For instance, as to the first and the second types of exemplary methods described above, the portion of a trigger nucleic acid to which a single-stranded nucleic acid probe anneals may be derived from a target nucleic acid or a portion thereof that is specific to a particular pathogenic organism to be detected. Alternatively, such a portion of a trigger nucleic acid may be derived from a target nucleic acid or a portion thereof that is substantially or completely conserved among several closely related pathogenic organisms, but absent in other more distantly related or unrelated pathogenic organisms.

As used herein, a target nucleic acid or a portion thereof that is "specific" to a particular pathogenic organism refers to a target nucleic acid or a portion thereof having a sequence that is present in the particular organism, not in any other organisms, including those closely related to the particular organism. In addition, as used herein, a region in a target nucleic acid that is "substantially conserved" among several closely related pathogenic organisms refers to a region in the target nucleic acid for which there

exists a nucleic acid molecule capable of hybridizing to the corresponding region in each of the several closely related organisms under appropriate conditions, but incapable of hybridizing to a similar region in the target nucleic acid from a more distantly related or unrelated organism under identical conditions. Also, as used herein, a region in a target nucleic acid that is "completely conserved" among several closely related pathogenic organisms refers to a region that has an identical sequence in the target nucleic acid from each of the several closely related pathogenic organisms.

Similarly, as to the above fourth type of exemplary methods, the portion of a target nucleic acid that is amplified with a primer pair may be a region that is specific for a particular pathogenic organism, or a region that is substantially or completely conserved among several closely related pathogenic organisms but absent in other distantly related or unrelated pathogenic organisms. In addition, the amplified portion of a target nucleic acid may be a variable region in the target nucleic acid among several closely related pathogenic organisms. As used herein, a "variable" region in a target nucleic acid refers to a region that has less than 50% sequence identity among the target nucleic acids from closely related organisms, but is surrounded by regions at each side having higher than 80% sequence identity among the target nucleic acids from the same closely related organisms. As used herein, percent sequence identity of two nucleic acids is determined using BLAST programs of Altschul *et al.* (*J. Mol. Biol. 215*: 403-10, 1990) with their default parameters. These programs implement the algorithm of Karlin and Altschul (*Proc. Natl. Acad. Sci. USA 87*:2264-8, 1990) modified as in Karlin and Altschul (*Proc. Natl. Acad. Sci. USA 90*:5873-7, 1993). BLAST programs are available, for example, at the web site http://www.ncbi.nlm.nih.gov.

Likewise, as to the above fifth type of exemplary methods, the overhang of a partially double-stranded nucleic acid probe may be at least substantially complementary to a region in a target nucleic acid specific to a pathogenic organism, or a region in a target nucleic acid that is substantially or completely conserved among several closely related pathogenic organisms. When the overhang is completely complementary to a target nucleic acid or a portion thereof from a particular organism, but also substantially complementary to the target nucleic acid or a portion thereof from one or more closely related organisms, one can vary hybridization stringencies to either detect the presence of the particular organism or to detect the presence of any one of the closely related organisms. For example, when a double-stranded nucleic acid probe is hybridized with nucleic acids from a biological sample under highly stringent conditions, nucleic acid amplification following the removal of unhybridized probes

using a portion of the probe molecule as a template may indicate the presence of the particular organism in the biological sample. On the other hand, when a probe is hybridized with nucleic acids from a biological sample under moderately or low stringent conditions, nucleic acid amplification (following the removal of unhybridized probe molecules) using a portion of the probe molecule as a template may indicate a presence of the particular organism and/or one or more organisms closely related to the particular organism. Adjusting stringencies of hybridization conditions is well known in the art and detailed discussions may be found, for example, Sambrook and Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, 2001.

In the embodiments where a template nucleic acid is provided by annealing a trigger nucleic acid to a single-stranded nucleic acid probe, the trigger nucleic acid or a portion thereof and a portion of the probe molecule located 3′ to the sequence of one strand of a NARS in probe may be substantially complementary, rather than completely complementary, to each other. For instance, when a trigger nucleic acid is derived from a region of a target nucleic acid that is substantially conserved among several closely related pathogenic organisms and the presence of any of the several organisms needs to be detected, a probe molecule substantially complementary to the trigger nucleic acid may be used. In such a circumstance, the primer extension reaction needs to be performed under conditions that are not too stringent to prevent the trigger nucleic acid from annealing to the probe molecule or prevent the trigger nucleic acid from being extended using a portion of the probe molecule as a template. However, such conditions need also be sufficiently stringent to prevent the probe molecule from non-specifically annealing to a nucleic acid molecule other than the trigger nucleic acid. Conditions suitable for nucleic acid amplification where a trigger nucleic acid or a portion thereof is substantially complementary to a portion of a probe molecule may be worked out by adjusting the reaction temperature and/or reaction buffer composition or concentration. Generally, similar to hybridization reactions, an increase in reaction temperatures increases the stringency of amplification reactions.

The present invention may be used to detect the presence of either RNA or DNA target sequences, or both. The DNA polymerase may be selected to have the ability to use either RNA or DNA as a template. For RNA target sequence detection, the biological sample, for instance, fixed cells or tissues, may be directly hybridized with nucleic acid probes. Because DNA molecules are typically present in the sample as double-stranded, they are not available to hybridize with nucleic acid probes. For DNA target sequence detection, the biological sample may be treated first with RNase or NaOH (about 0.1 M) to degrade potential RNA target sequences. The treated sample

can be then heated to denature double-stranded target nucleic acids to allow resulting single strands to hybridize with nucleic acid probes. If the heating step is included (without RNase treatment) prior to the hybridization between the sample and nucleic acid probes, both DNA and RNA target sequences are available for hybridization with the probes and thus may be detected. Such hybridization is preferred for most diagnostic applications of the invention because it gives the greatest number of probes that hybridize to the target sequences and thus the best yield of amplified single-stranded nucleic acid molecules.

In certain embodiments where the target nucleic acids are immobilized onto a solid support, the solid support with immobilized target nucleic acids may first be incubated with a blocking agent to block the nonspecific attachment of a probe to the solid support before being hybridizing with the probe. Such a blocking agent is known in the art and readily available to one of ordinary skill in the art (*see, e.g.*, Sambrook *et al.*, *supra*).

d. Single-Stranded Nucleic Acid Amplification

Single-stranded nucleic acid amplification is essentially as described above for genetic variation detection methods. To facilitate detection (described in more detail below), a labeled dideoxyribonucleoside triphosphate may be used in the amplification and subsequently incorporated into the amplified single-stranded nucleic acid. Labels suitable for incorporating into a nucleic acid fragment, and methods for the subsequent detection of the fragment are known in the art, and exemplary labels are described above for genetic variation identification methods. In certain embodiments, the label may be a radiolabel such as ^{32}P, ^{33}P, ^{125}I or ^{35}S, an enzyme capable of producing a colored reaction product such as alkaline phosphatase, fluorescent labels such as fluorescein isothiocyanate (FITC), biotin, avidin, digoxigenin, antigens, haptens or fluorochromes. The presence of the label may also facilitate retention of the amplified single-stranded nucleic acid molecules within the cells to allow *in situ* detection of target nucleic acids.

As described above, a single-stranded nucleic acid molecule is amplified using a portion of a template nucleic acid as a template. In certain embodiments, the amplified single-stranded nucleic acid may be relatively short and has at most 25, 20, 17, 15, 10, or 8 nucleotides. Such short length may be accomplished by appropriately designing single-stranded nucleic acid probes or ODNPs used in making template molecules. For instance, for the second type of providing template molecules (Figure 37), the single-stranded nucleic acid probe may be designed to have a short region 5' to

a sequence of the antisense strand of a NARS. For the fourth type of providing template molecules (Figures 39-41), the ODNP pair may be designed to be close to each other when the primers anneal to the target nucleic acid. For the fifth type of providing template nucleic acid molecules (Figure 6), the partially double-stranded probes may be designed to have a short region 5' to a sequence of the antisense strand of a NARS. The short length of an amplified single-stranded nucleic acid molecule may be advantageous because it increases amplification efficiencies and rates. In addition, it allows the use of a DNA polymerase that does not have a stand displacement activity. It also facilitates the detection of the amplified single-stranded nucleic acid molecule via certain technologies such as mass spectrometric analysis.

e. Detection/Characterization of Amplified Single-Stranded Nucleic Acid Molecules

The detection/characterization methods for amplified single-stranded nucleic acid molecules of the present invention include any method known in the art that is suitable for detecting and/or characterizing short nucleic acid fragments. Exemplary methods are also described above for genetic variation identification methods. Detection of amplified single-stranded nucleic acid molecules may be performed *in situ* or after the molecules have been released and/or removed from the biological sample. In certain embodiments wherein the target nucleic acids are first isolated from the biological sample and subsequently immobilized onto a solid support, the amplified single-stranded nucleic acid molecules may be transferred to another solid support (also referred to as "solid substrate") before their detection via techniques such as MALDI-TOF. Such a procedure can be easily multiplexed and thus used to form an array of amplified single-stranded nucleic acid molecules originated from target nucleic acids from multiple different biological samples. The use of the techniques that can simultaneously detect and/or characterize a plurality of amplified single-stranded nucleic acid molecules (*e.g.*, MALDI-TOF) increases the throughput of the diagnostic method of the present invention.

In certain embodiments, as described above, the amplified single-stranded nucleic acid molecules are labeled by incorporating labeled deoxynucleotide triphosphates. Certain such labeled molecules (*e.g.*, those labeled with radioisotopes, fluorescent moieties and dyes) may be detected directly using techniques such as radiography, fluorescence spectroscopy or spectrometry, and luminescence spectroscopy or spectrometry. Other labeled molecules (*e.g.*, enzymes capable of producing a colored reaction product, biotin, avidin, digoxigenin, antigens, haptens or

fluorochromes) may have to be detected indirectly. For instance, alkaline phosphatase may be detected by reaction with a substrate such as Vector Red/Vector Blue (Vector Labs, CA), 5-bromo-4-chloro-3-indolyl phosphate (BCIP)/nitro blue tetrzolium (NBT) (Sigma Chem. Co., MO) or Nuclear Fast Red (Sigma Chem. Co.). Biotin (or avidin) labels may be detected by binding to labeled avidin (or labeled biotin). Digoxigenin and hapten labels may be detected by a labeled anti-digoxigenin or anti-hapten antibody.

Fluorescence polarization (FP) may also be used to detect the amplification of single-stranded nucleic acid molecules. FP is based on the phenomenon that when a fluorescent molecule is excited by plane-polarized light, it emits polarized fluorescent light into a fixed plane if the tagged-molecules do not significantly rotate between excitation and emission. If the molecule is small enough and rotates and tumbles in space, however, fluorescence polarization is not observed fully by the detector.

The fluorescence polarization of a molecule is proportional to the molecule's rotational relaxation time (usually the time it takes to rotate through an angle of 68.5°), which is related to properties of the solution such as the viscosity, temperature, and molecular volume of the analyte or biomolecule. Therefore, if the viscosity and temperature are held constant, fluorescence polarization is directly proportional to molecular volume, which, in turn, is directly proportional to molecular weight. Larger tagged molecules rotate and tumble slowly in space and, accordingly, fluorescence polarization values can be obtained. In contrast, smaller molecules rotate and tumble faster and fluorescence polarization cannot be measured.

In one aspect, the present invention uses fluorescence polarization to detect fluorescence labeled nucleotides upon their incorporation into single-stranded oligonucleotides. Fluorescence labeled nucleoside triphosphates are small and rotate rapidly in solution. Thus, absent incorporation into a small nucleic acid fragment, fluorescent-tagged-nucleoside triphosphates are undetectable by fluorescence polarization. The amplification of single-stranded nucleic acid molecules described above that incorporates a fluorescence labeled nucleoside triphosphate into the nucleic acid molecules increases by about 20-fold the molecular weight of the fluorophore. Because of the increase in the molecular weight, the fluorescence polarization of the oligonucleotide can be detected and measured.

Various fluorescent dyes or probes may be employed in the present invention. Fluorescent dyes are identified and quantified most directly by their absorption and fluorescence emission wavelengths and intensities. Emission spectra

(fluorescence and phosphorescence) are much more sensitive and specific than absorption spectra. Other photophysical characteristics (like fluorescence anisotropy) are less preferred. The useful intensity parameters are quantum yield (QY) for fluorescence, and the molar extinction coefficient (ε) for absorption. QY is a measure of the total photon emission over the entire fluorescence spectral profile and the value of ε is specified at a given wavelength (usually the absorption maximum of the probe). A narrow optical bandwidth (<25 nm) is usually used for fluorescence excitation (via absorption), whereas the fluorescence detection bandwidth is more variable, ranging from full spectrum for maximal sensitivity to narrow band (~20 nm) for maximal resolution. Fluorescence intensity per probe molecule is proportional to the product of ε and QY. Commercially important and exemplary fluorochromes that are suitable for use in the present invention are fluorescein, tetramethylrhodamine, lissamine, Texas Red and BODIPYs.

Fluorescent labels are now commonly used for the detection of small nucleic acid fragments that have been separated by CE and HPLC. One group of labels that may be employed for this purpose is those based on near-infrared (near-IR) fluorescent dyes. In aqueous solution, these types of tags have a maximum absorption of light at >680 nm, followed by the emission of fluorescence at near-IR wavelengths (emission maximum, >700 nm). One advantage of using this type of fluorescence for detection is that it occurs in a spectral region where there is relatively little absorption or emission due to other compounds that might be present in biological samples. This, plus the fact that most near IR probes can be excited with commercially available lasers, provides this approach with low background signals and limits of detection that extend into the attomole range.

One of the limitations of fluorescence spectroscopy is the phenomenon of autofluorescence. One method to avoid autofluorescence is to employ fluorochromes that possess significantly longer delay times to emission. These fluorochromes are usually luminescent metal chelates that are attached at the 5′-end of an ODN probe or primer.

Terbium deoxyribonucleoside triphosphates are available that allow the incorporation of time-resolved fluorochromes into "natural" nucleic acids. These probes have the advantage of the large Stokes shift, narrow emission bands and long lifetimes. Time-resolved fluorescence spectroscopy is particularly useful in structural biology and is used to monitor molecular interactions and motions that occur in the picosecond-nanosecond time range. Time-resolved fluorescence spectroscopy is

beginning to dominate the analysis of biomolecular structure and dynamics, and may be used as a detection/characterization technique according to the present invention.

Deoxyribonucleoside analogs that may be incorporated into a small nucleic acid fragment of the present invention, to thereby afford an effective characterization means for the small nucleic acid, include but are not limited to: Fluorescein-12-dUTP, Coumarin-5-dUTP, Tetramethylrhodamine-6-dUTP, Texas Red®-5-dUTP, Napthofluorescein-5-dUTP, Fluorescein Chlorotriazinyl-4-dUTP, Pyrene-8-dUTP, Diethylaminocoumarin-5-dUTP, Cyanine 3-dUTP, Cyanine 5-dUTP, Coumarin-5-dCTP, Fluorescein-12-dCTP, Tetramethylrhodamine-6-dCTP, Texas Red®-5-dCTP, Lissamine™-5-dCTP, Napthofluorescein-5-dCTP, Fluorescein Chlorotriazinyl-4-dCTP, Pyrene-8-dCTP, Diethylaminocoumarin-5-dCTP, Cyanine 3-dCTP, Cyanine 5-dCTP, Coumarin-5-dATP, Diethylaminocoumarin-5-dATP, Fluorescein-12-dATP, Fluorescein Chlorotriazinyl-4-dATP, Lissamine™-5-dATP, Napthofluorescein-5-dATP, Pyrene-8-dATP, Tetramethylrhodamine-6-dATP, Texas Red®-5-dATP, Cyanine 3-dATP, Cyanine 5-dATP, Coumarin-5-dGTP, Fluorescein-12-dGTP, Tetramethylrhodamine-6-dGTP, Texas Red®-5-dGTP, and Lissamine™-5-dGTP.

Ribonucleoside analogs include but are not limited to: Fluorescein-12-UTP, Coumarin-5-UTP, Tetramethylrhodamine-6-UTP, Texas Red®-5-UTP, Lissamine™-5-UTP, Napthofluorescein-5-UTP, Fluorescein Chlorotriazinyl-4-UTP, Pyrene-8-UTP, Cyanine 3-UTP, Cyanine 5-UTP, Coumarin-5-CTP, Fluorescein-12-CTP, Tetramethylrhodamine-6-CTP, Texas Red®-5-CTP. Lissamine™-5-CTP, Napthofluorescein-5-CTP, Fluorescein Chlorotriazinyl-4-CTP, Pyrene-8-CTP, Cyanine 3-CTP, Cyanine 5-CTP, Coumarin-5-ATP, Fluorescein-12-ATP, Tetramethylrhodamine-6-ATP, Texas Red®-5-ATP, Lissamine™-5-ATP, Coumarin-5-GTP, Fluorescein-12-GTP, Tetramethylrhodamine-6-GTP, Texas Red®-5-GTP, and Lissamine™-5-GTP.

Dideoxy analogs include but are not limited to: Fluorescein-12-ddUTP, FAM-ddUTP, ROX-ddUTP, R6G-ddUTP, TAMRA-ddUTP, JOE-ddUTP, R110-ddUTP, Fluorescein-12-ddCTP, FAM-ddCTP, ROX-ddCTP, R6G-ddCTP, TAMRA-ddCTP, JOE-ddCTP, R110-ddCTP, Fluorescein-12-ddGTP, FAM-ddGTP, ROX-ddGTP, R6G-ddGTP, TAMRA-ddGTP, JOE-ddGTP, R110-ddGTP, Fluorescein-12-ddATP, FAM-ddATP, ROX-ddATP, R6G-ddATP, TAMRA-ddATP, JOE-ddATP, and R110-ddATP.

All of the above analogs can be further radiolabeled with ^{3}H, deuterium, ^{32}P, ^{14}C, ^{35}S and other radioisotopes.

Analogs can also be un-natural nucleoside analogs including, but not limited to, the following: 8-Bromo-2′-deoxyadenosine-TTP, 8-Oxo-2′-deoxyadenosine, Etheno-2′-deoxyadenosine-TTP, Etheno-2′-deoxyadenosine-TTP, N^{6}-Methyl-

2′deoxyadenosine-TTP, 2,6-Diaminopurine-2′-deoxyriboside-TTP, 8-Bromo-2′-deoxyguanosine-TTP, 7-Deaza-2′-deoxyguanosine-TTP, 2′-Deoxyisoguanosine-TTP, -Oxo-2′-deoxyguanosine-TTP, O^6-Methyl-2′-deoxyguanosine-TTP, S^6-DNP-2′-deoxythioguanosine-TTP, 3-Nitropyrrole-2′-deoxyriboside-TTP, 5-Propynyl-2′-deoxyuridine-TTP, 5-Fluoro-2′-deoxyuridine-TTP, 2′-deoxyuridine-TTP, 5-Bromo-2′-deoxyuridine-TTP, 5-Iodo-2′-deoxyuridine-TTP, and 4-Triazolyl-2′-deoxyuridine-TTP.

Currently, polarizing fluorometers and more than 50 fluorescence polarization immunoassays (FPIAs) are commercially available, many of which are routinely used in clinical laboratories for the measurement of therapeutics, metabolites, and drugs of abuse in biological fluids. (*See, e.g.*, Checovich et al., *Nature 375*:254-256 (1995); published erratum appears in *Nature 375*:520 (1995) both of which are incorporated herein by reference in their entirety).

Detection of amplified single-stranded nucleic acid molecules alone may be sufficient to indicate the presence of a target nucleic acid if only one nucleic acid probe of the present invention is used to hybridize to the biological sample where the target nucleic acid is present. However, when multiple nucleic acid probes are used to determine the presence, or the absence, of multiple target nucleic acids in a biological sample, amplified single-stranded nucleic acid fragments usually need to be characterized. Such characterization allows the identification of each amplified molecule and consequently the identification of its correlated target nucleic acid. As discussed above, any technique known in the art suitable for characterizing small nucleotide fragments may be used for characterizing the amplified nucleic acid molecules. In certain preferred embodiments, these molecules are characterized by LC/MS.

2. Kits for Detecting Target Nucleic Acids

The present invention also provides kits for detecting target nucleic acids in biological samples. Such kits generally comprise a partially double-stranded nucleic acid probe as described above or two single-stranded oligonucleotides that when annealed produce the partially double-stranded nucleic acid probe. They may also further comprise at least one, two, several, or each of the following components: (1) a nicking agent (*e.g.*, a NE or a RE) that recognizes the NARS in the probe and/or the buffer thereof; (2) a DNA polymerase for amplifying single stranded nucleic acid fragments and/or the buffer thereof; (3) dNTPs (some of the dNTPs may be labeled); (4) a control target nucleic acid and a control probe; (4) hybridization buffer; (5) washing buffer for removing unhybridized probes; (6) a liquid chromatography column;

(7) Buffer A for the chromatography column; (8) Buffer B for the chromatography column; (9) water; (10) a solid support coated with a chemical bindable to nucleic acid (*e.g.*, PEI ocated stainless steel pin); (11) microtubes or microwell plates; (12) a strand displacement facilitator (*e.g.*, 1M trehalose); and (13) an instruction booklet for using the kit. Components common to kits for detecting genetic variations are also described in the section regarding those kits.

3. Applications for Target Nuclei Acid Detection Methods of the Present Invention

As discussed in detail herein above, the present invention provides methods for detecting target nucleic acids in biological samples. These methods will find utility in a wide variety of applications wherein it is necessary to determine the presence, or the absence, of a target nucleic acid in a biological sample. Such applications include, but are not limited to, disease diagnosis and predisposition, crop cultivation and animal breeding, expression profiling of cell function and/or disease marker genes and identification and/or characterization of infectious organisms that cause infectious diseases in plants or animals and/or that are related to food safety.

The present invention is useful in quickly detecting the presence of any target nucleic acid of interest. In certain embodiments, the target nucleic acid is derived or originated from a pathogenic organism (*e.g.*, an organism that causes infectious diseases). Such pathogenic organisms include those that impose bio-threat, such as Anthrax and smallpox. In addition, the present methods may be used for the detecting the presence of a particular pathogenic organism as well as for detecting the presence of several closely related pathogenic organisms. The present invention may also be used to detect organisms that are resistant to certain antibiotics. For example, the present methods, compositions or kits may be used to detect certain pathogenic organisms in a subject that has been treated with an antibiotic or certain combinations of antibiotics. It may also be used to detect the presence, or the absence, of a nucleic acid associated with a particular feature (*e.g.*, disease resistance or susceptibility) and to predict the likelihood of a particular subject from which the sample was obtained to have the particular feature.

In addition, the present invention also provides methods for multiplex determination of the presence, or the absence, of multiple target nucleic acids in a sample. These methods are particularly valuable when the sample (*e.g.*, a biopsy specimen) is difficult to obtain, the amount of the sample is scarce, or the patient from which the sample was obtained suffers from a symptom that is common among

different disorders or diseases. These methods allow one to determine the presence, or the absence, of multiple nucleic acids associated with various disorders or diseases, using the precious sample, and thus facilitate disease diagnosis for the patient.

D. Preparation of Single-Stranded Nucleic Acid Probes and cDNA Libraries

In another aspect, the present invention provides methods, compounds and kits for preparing single-stranded nucleic acid molecules and cDNA libraries useful in various nucleic acid analyses. It uses a nucleic acid adaptor in certain embodiments and an oligonucleotide primer in other embodiments.

1. Methods for Preparing Single-Stranded Nucleic Acid Molecules and cDNA Libraries Using Nucleic Acid Adaptors

The present invention provides methods for preparing single-stranded nucleic acid molecules or probes useful in various nucleic acid analyses. In certain embodiments, the methods utilize a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS). The nucleic acid adaptor is linked to a double-stranded target nucleic acid fragment in the direction so that in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, the target nucleic acid is located 5′ to the NARS. Such a linkage allows the use of that strand of the target nucleic acid as a template for synthesizing single-stranded nucleic acid probes in the presence of the NA that recognizes the NARS and a DNA polymerase. In certain embodiments, the nucleic acid adaptor further comprises a type IIs restriction endonuclease recognition sequence (TRERS). The TRERS is located in the adaptor so that when the adaptor is linked to a double-stranded target nucleic acid fragment in the direction as described above, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS in the strand that does not contain the NS is located both within the double-stranded target nucleic acid fragment and 5′ to the position corresponding to the NS. Such a location of the TRERS in the adaptor allows the use of a portion of one strand of the target nucleic acid as a template nucleic acid for synthesizing single-stranded nucleic acid probes after the double-stranded target nucleic acid is ligated with the adaptor and the resulting ligation product digested by a type IIs restriction endonuclease that recognizes the TRERS. The presence of the TRERS enables the making of relatively short single-stranded nucleic acid probes.

The above method may be multiplexed to prepare multiple single-stranded nucleic acid probes. For such an application, the adaptor of the present invention (described in detail below) is linked to multiple target nucleic acids in the direction as described above. Each of the target nucleic acids is then used as a template

for synthesizing a single-stranded nucleic acid probe that is capable of specific hybridization to one strand of each of the target nucleic acids. The resulting mixture of single-stranded nucleic acid probes is useful in various nucleic acid analyses described in detail below.

In a related aspect, the present invention also provides methods for preparing and/or amplifying cDNA libraries. The general scheme for these methodsis similar to that for the multiplex preparation of single-stranded nucleic acid probes described above wherein the double-stranded target nucleic acid fragments are double-stranded cDNA molecules. Briefly, double-stranded nucleic acid molecules synthesized using mRNAs isolated from a biological sample are ligated to a nucleic acid adaptor that comprises a NARS so that in the strand that does not contain the NS of a NA that recognizes the NARS in the adaptor, the cDNA is located 5′ to the NARS. The ligation products, in combination, form a cDNA library specific to the mRNAs from the biological sample. This cDNA library may be amplified by first amplifying single-stranded cDNA molecules using one strand of the double-stranded cDNA molecules as templates in the presence of a NA that recognizes the NARS in the adaptor. Typically, the adaptor does not further comprise a TRERS. The absence of the TRERS in the adaptor allows for the amplification of relatively long single-stranded cDNA molecules using one strand of the cDNA molecules ligated to the nucleic acid adaptor as templates. The amplified single-stranded nucleic acid molecules may further function as templates for synthesizing or amplifying their complementary strands. The synthesis and/or amplification of both strands of the double-stranded cDNA molecules thus amplify the cDNA library of which the double-stranded cDNA molecules are components.

Such an amplification of cDNA libraries may be performed *in vitro* as compared to the amplification of conventional cDNA libraries (*e.g.*, cDNA libraries in λ phage) that requires *in vivo* nucleic acid amplification. In addition, the original cDNA library, especially when immobilized to a solid support, may be stored and used as templates for theoretically indefinite runs of library amplification. Because each run of library amplification uses the original cDNA library as templates, the amplification of cDNA libraries of the present invention reduces or eliminates the distortion of cDNA contents in a cDNA library prepared using an amplified cDNA library as templates when an original cDNA library is used up. Such distortion may be due to unequal amplification efficiencies among different cDNA clones during the library amplification. In addition, the cDNA libraries of the present invention and the method for amplifying such libraries decreases the amount of total RNA or mRNA required for

constructing a cDNA library, increases sample throughput, and reduces user manipulations.

a. Target Nucleic Acids and cDNA Molecules

Target nucleic acids useful in providing templates for synthesizing single-stranded nucleic acid probes can be any nucleic acid of interest. They include, but are not limited to, cDNAs, genomic DNAs, RNAs, and artificially synthesized nucleic acids. The target nucleic acids other than those artificially synthesized may be directly isolated from the organisms from which the target nucleic acids originate. Alternatively, they may be prepared from nucleic acid libraries or via recombinant nucleic acid techniques.

In certain preferred embodiments, the target nucleic acids are nucleic acids associated with, and/or characteristic of, pathogenic organisms, such as pathogenic bacteria, viruses, molds, parasites, and fungi, as found in biological fluids, tissues, foods, and water supplies. Exemplary pathogenic bacteria capable of providing a target nucleic acid sequence for synthesizing single-stranded nucleic acid probes include, but are not limited to, *Bacillus anthracis, Bacillus botulinus (Clostridium botulinum), Bacillus dysenterea, Bacillus enteritidis (Salmonella enteritidis), Bacillus pneumoniae (Diplococcus pneumoniae, Klebsiella pneumoniae), Bacillus tetani (Clostridium tetani), Bacillus typhosus (Salmonella typhosa), Chlamydia oculogenitalis, Clostridium perfringens, Hemophilus influenzae, Hemophilus pertussis, Mycobacteria tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Shigella paradysenteriae, Staphylococcus aureus, Streptococcus agalactiae, Streptococcus faecalis, Streptococcus hemolyticus, Streptococcus mitis, Streptococcus pyrogenes, Vibrio cholerae, Vibrio comma, Vibrio fetus, Vibrio jejuni, Vibrio metchnikovii, and Vibrio niger.* Exemplary viruses, including retroviruses, capable of providing a target nucleic acid sequence for synthesizing single-stranded nucleic acid probes include, but are not limited to, HSV-I, HSV-II, HIV, CMV, measles, polio, hepatitis B, hepatitis C, influenza, and the like. Nucleic acid sequences characteristic of each of these species are known to the art and suitable for use in whole or part as target nucleic acid sequences in the method of the present invention.

In certain other preferred embodiments, the target nucleic acids are nucleic acid sequences associated with, and/or characteristic of, diseases or disorders in plants, animals or humans. In some other embodiments, the target nucleic acids are nucleic acid sequences originated from normal (or control) plants, animals or humans that does not have the disease or disorder of interest. In yet other preferred

embodiments, the target nucleic acids are associated with, and/or characteristic of, desirable traits in plants or animals, such as disease resistance and high crop or cattle yield.

In some embodiments, the target nucleic acids are cDNAs derived from mRNA isolated from an organism or a biological sample of interest. In addition, such cDNAs may also be used for preparing and/or amplifying cDNA libraries according to the method of the present invention.

The methods for isolating mRNA and cDNA synthesis are well known in the art (*see, e.g.*, Sambrook *et al.*, *supra*; Chomczynski *et al.*, *Anal. Biochem. 162*: 156, 1987). Such methods generally comprise cell, tissue or sample lysis and RNA recovery by means of extraction procedures. These procedures can be done in particular by treatment with chaotropic agents such as guanidinium thiocyanate followed by RNA extraction with solvents such as phenol and chloroform. They may be readily implemented by using commercially available kits such as US73750 kit (Amersham) for total RNA isolation. mRNA molecules may be purified from total cellular RNA using oligo(dT) primers that bind the poly(A) tails of the mRNA molecules (*see*, Jacobson, *Metho. Enzymol. 152*: 254, 1987, incorporated herein by reference). In this regard, the preparation of mRNA can be carried out using commercially available kits such as US72700 kit (Amersham). Alternatively, random primers (*i.e.*, primers with random sequences) may be used for purifying mRNA from total cellular RNA (*see*, Singh *et al.*, *Cell 52*: 415, 1988; Vinson *et al.*, *Genes Dev. 2*: 801, 1988). Either the oligo(dT) primers or the random primers may be immobilized to facilitate the purification of mRNAs. In certain other embodiments, mRNA may be directly isolated from biological samples without first isolating total RNA.

The isolated/purified mRNAs may be then used as templates for synthesizing first strand cDNAs by reverse transcription according to conventional molecular biology techniques (*see, e.g.*, Sambrook *et al.*, *supra*). Reverse transcription is generally carried out using a reverse transcriptase and a primer.

Many reverse transcriptases have been described in the literature and are commercially available (*e.g.*, 1483188 kit, Boehringer). Exemplary reverse transcriptases include, but are not limited to, those derived from avian virus AMV (Avian Myeloblastosis Virus), from murine leukemia virus MMLV (Moloney Murine Leukemia Virus), from *Yhermus flavus* and *Thermus thermophilus* HB-8 (Promega, catalog number M1941 and M2101). The operating conditions that apply to these enzymes are well known to those of ordinary skill in the art.

The primers used for reverse transcription may be of various types. It may be a random oligonucleotide comprising 4 to 10 nucleotides, preferably a hexanucleotide. Use of this type of random primer has been described in the literature and allows random initiation of reverse transcription at different sites within the RNA molecules. Alternatively, a poly(dT) primer comprising 4 to 20-mers, preferably 15mers may be used. In certain embodiments, the primer used in isolating mRNA is also used in cDNA synthesis.

Second strand cDNA may be synthesized using an RNase H and a DNA polymerase. Alternatively, it may be synthesized by first ligating an adaptor sequence to a first strand cDNA molecule and extending a primer complementary to the adaptor sequence using the first strand cDNA as a template.

The synthesized cDNAs may be in solution or linked to a solid support, for example, via an immobilized primer for isolating mRNA and synthesizing cDNAs.

b. Nucleic Acid Adaptors

The present invention provides nucleic acid adaptors useful for synthesizing single-stranded nucleic acid probes as well as for preparing and/or amplifying cDNA libraries. The adaptor comprises a nicking agent recognition sequence (NARS) and may or may not contain the nicking site (NS). In certain embodiments, the adaptor further comprises a type IIs restriction endonuclease recognition sequence (TRERS), which may be located at either side of the NARS. The presence of the TRERS enables the production of relatively short oligonucleotide probes (*e.g.*, 15-50 nucleotides in length) particularly useful for probing arrays of similarly short oligonucleotides.

For amplifying a single-stranded nucleic acid molecule that may be used as a nucleic acid probe for detecting the presence of a target nucleic acid or as a first strand cDNA template for synthesizing a second strand of cDNA, the adaptor of the present invention must be ligated to the target nucleic acid (including a cDNA molecule or a fragment thereof) in such a direction that, in the strand of the ligation product that does not contain the NS, the target nucleic acid fragment is located 5′ to the sequence of the antisense strand of the NARS (*see, e.g.*, Figures 29A-D where a NERS is used as an exemplary NARS). Such an arrangement is required for the subsequently amplified probes to contain sequences complementary to a portion of one strand of the target nucleic acid, as only the 3′-OH at the NS generated by the NA may be extended and the resulting extension products may be further amplified. Thus, the ligation product of the adaptor and the target nucleic acid in the other direction is incapable of being template

for amplifying nucleic acid probes that can hybridize with the target nucleic acid (*see, e.g.*, Figures 29 E-H where a NERS is used as an exemplary NARS).

In addition to the directional linkage between the adaptor and the NERS, the cleavage site of a restriction endonuclease (RE) that recognizes the TRERS is also required to locate both (1) within the target nucleic acid and (2) 5′ to the position corresponding to the NS. Such a location of the TRERS allows the production of a partially double-stranded nucleic acid molecule with a 5′ overhang generated by the NA (which generates the recessed 3′ terminus corresponding to the 5′ overhang) and the RE (which generates the 5′ overhang) that contains a portion of the target nucleic acid. Thus, when this 5′ overhang is used as a template for synthesizing single-stranded nucleic acid probes in the presence of the NA and a DNA polymerase, the synthesized probes contains a sequence that is complementary to the portion of the target nucleic acid.

In certain preferred embodiments, the NARS is a nicking endonuclease recognition sequence (NERS) that is recognizable by a NE that nicks outside its recognition sequence (*e.g.*, N.BstNB I). In such embodiments, the direction from the TRERS to the cleavage site of a RE that recognizes the TRERS (indicated by the arrows above or under the TRERS in Figures 29A-29D) and the direction from the NERS to the NS of a NE that recognizes the NERS (indicated by the arrows above or under the NERS) needs to be the same in order to fulfill the above two requirements (Figures 29A and 29C).

The nucleic acid adaptors of the present invention may be prepared by any methods of making short nucleic acid fragments known in the art. For instance, each strand of the nucleic acid adaptors may be synthesized individually and subsequently annealed with its complementary strand. In certain embodiments, the nucleic acid adaptors may be immobilized to a solid support.

c. Ligation between Adaptors and Target Nucleic Acids

As described above, to produce a template for synthesizing single-stranded nucleic acid probes or to prepare and/or amplify a cDNA library, the nucleic acid adaptor of the present invention needs to be directionally linked to a target nucleic acid or a cDNA molecule or a fragment thereof so that in the strand of the resulting nucleic acid molecule that does not contain the NS, the target nucleic acid fragment or the cDNA molecule or the fragment thereof is located 5′ to the NARS.

Any method for directionally linking a nucleic acid adaptor with another double-stranded nucleic acid fragment known in the art may be used in this invention.

For instance, if a target nucleic acid (including a cDNA molecule or a fragment thereof) is immobilized to a solid support, then only one terminus of the target nucleic acid is available to be linked with the nucleic acid adaptor. The directional linkage between the available terminus of the target nucleic acid and the appropriate terminus of the nucleic acid adaptor (*i.e.*, the 3′ terminus of the strand of the adaptor that either the strand itself or its ligation product contains the NS) may be accomplished by designing the adaptor to have the terminus to be ligated to the target to be a blunt end and the other terminus to have either a 3′ or a 5′ overhang. The presence of the 3′ or 5′ overhang in the other terminus prevents that terminus from being ligated with the target nucleic acid or the cDNA molecule or the fragment thereof.

In certain embodiments, the target nucleic acid is provided by nucleic acid amplification using a primer. The primer may contain a restriction endonuclease recognition sequence (RERS) that produces a protruding end upon digestion with a restriction endonuclease. The protruding end may be ligated to a compatible protruding end that is purposefully designed at the appropriate terminus of the nucleic acid adaptor. These compatible protruding ends allow directional linkage between the target nucleic acid and the nucleic acid adaptor. Alternatively, both the terminus of the target nucleic acid and the terminus of the nucleic acid adaptor that are involved in the ligation between the target and the adaptor are blunt, while the other terminus of the target and the other terminus of the adaptor are protruding, but not compatible with each other. The incompatibility between the two protruding termini prevents the ligation between the target and the adaptor at these termini.

In certain preferred embodiment, template nucleic acids for synthesizing single-stranded nucleic acid probes or for amplifying cDNA libraries may be immobilized either via a terminus of a target nucleic acid (including a cDNA molecule or a fragment thereof) or via a terminus of a nucleic acid adaptor. Such immobilization enables easy separation of the template nucleic acids from other components of nucleic acid amplification reaction mixtures, including amplified nucleic acid molecules. The immobilized template nucleic acids may thus be easily washed and stored for future use.

d. Single-Stranded Nucleic Acid Amplification

Single-stranded nucleic acid amplification is essentially the same as described above for genetic variation detection methods. In addition, in certain embodiment, as the amplified single-stranded nucleic acid molecules are used as nucleic acid probes, they are generally labeled to facilitate subsequent detection in

nucleic acid analyses. Labels suitable for incorporating into a nucleic acid fragment and methods for the subsequent detection of the fragment are known in the art. Exemplary labels and detection methods are described above for genetic variation detection methods and for target nucleic acid detection methods.

In certain embodiments, the amplified single-stranded nucleic acids are relatively short (*e.g.*, 15-50 nucleotides in length). As described above, such short nucleic acid molecules are generally prepared using nucleic acid adaptors that contain a TRERS. In other embodiments, longer single-stranded nucleic acid molecules (*e.g.*, 51-20,000 nucleotides in length) are produced. For the amplification of such longer single-stranded nucleic acid molecules, a strand displacement facilitator (*e.g.*, trehalose) may be used to enhance the strand displacement ability of a DNA polymerase.

2. Methods for Preparing Single-Stranded Nucleic Acid Molecule Using Oligonucleotide Primers

In certain embodiments, the present invention provides methods for preparing single-stranded nucleic acid molecules using oligonucleotide primers. The methods are based on the discovery of the present invention that for certain nicking agents, only the sense strands of their double-stranded nicking agent recognition sequences, not both strands of the double-stranded nicking agent recognition sequences, are required for their nicking activities. An exemplary method where a target nucleic acid is single-stranded is illustrated in Figure 43. One of ordinary skill in the art appreciates that the present methods are also useful in amplifying a portion of one strand of a double-stranded target nucleic acid.

Referring to Figure 43, an oligonucleotide primer is used to amplify a single-stranded nucleic acid molecule using a portion of a single-stranded target nucleic acid as a template. The primer comprises, from 5' to 3', three regions: Region A, Region B and Region C. Region B consists of a sequence of the sense strand of a double-stranded nicking agent recognition sequence, where Region A and Region C are regions that are located directly 5' and 3' to Region B, respectively. The oligonucleotide primer is at least substantially complementary to the target nucleic acid so that under conditions that allow for the amplification of a single-stranded nucleic acid, the oligonucleotide primer is able to anneal to the target and extends from its 3' terminus in the presence of a DNA polymerase. The resulting extension product may be nicked in the presence of a nicking agent that recognizes the double-stranded nicking agent recognition sequence even though there may be one or more nucleotides in Region B of the oligonucleotide primer that do not form conventional base pairs with nucleotides in

the target nucleic acid. A "conventional base pair" is a base pair formed according to the standard Watson-Crick model (*e.g.*, G:C, A:T, and A:U) between a nucleotide of one strand of a fully or partially double-stranded nucleic acid and another nucleotide on the other strand of the nucleic acid. The nicked product that contains the 5' terminus may readily dissociate from the target nucleic acid if it is relatively short (*e.g.*, no longer than 18 nucleotides) or be displaced by the extension of the nicked product that contains the 3' terminus at the nicking site. If the nicking agent nicks outside its recognition sequence, the extension product retains Region B of the oligonucleotide primer (*i.e.*, the sequence of the sense strand of the nicking agent recognition sequence) and may thus re-nicked by the nicking agent. The above nicking-extension cycle may be repeated multiple times, resulting in the amplification of a single-stranded nucleic acid molecule that contains the 5' terminus at the nicking site.

In embodiments where there are one or more mismatches between Region B and its corresponding region in the target nucleic acid, the nicking activity of a nicking agent that recognizes Region B decreases with the increase in the number of the mismatches between Region B and its corresponding region in the target. For example, N.BstNB I is about half as active in nicking a duplex that comprises a sequence of the sense strand of its double-stranded recognition sequence but has one mismatch between the sense strand of its recognition sequence and its corresponding region in the opposite strand of the duplex as in nicking a duplex that comprises a double-stranded recognition sequence. The nicking activity of N.BstNB I decreases to about 10% to 20% of its maximum level when it nicks a duplex that comprises a sequence of the sense strand of its double-stranded recognition sequence but does not have any nucleotides in the other strand that form conventional base pairs with any of the nucleotides in the sense strand of the recognition sequence.

In certain embodiments, a nicking agent that nicks within its recognition sequence may also be used where the nucleotide(s) in Region B that does not form a conventional base pair with a nucleotide in the target is located 5' to the nicking site within Region B. After the duplex formed between the oligonucleotide primer and the target is nicked by the nicking agent within Region B, the 3' terminus at the nicking site may be extended to regerate Region B. Such regeneration allows for the repetition of the nicking-extension cycles. In addition, the mismatch(es) between Region B and the corresponding region in the target must not affect the extension from the 3' terminus at the nicking site. Generally, the more distance between the nicking site and the nucleotide(s) in Region B that does not form a conventional base pair, the less adverse effect the mismatch(es) has on the extension.

Region A facilitates or enables the annealing of the oligonucleotide primer to the target nucleic acid. In addition, it facilitates or enables the nicked product that contains the 3' terminus at the nicking site to remain annealing to the target and to extend from the 3' terminus in the presence of a DNA polymerase. In certain embodiments, Region A is at most 100, 75, 50, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, or 2 nucleotides in length. In some embodiments, there may be one or more nucleotides that do not form conventional base pairs in Region A with the nucleotides in the target nucleic acid.

An oligonucleotide primer may or may not have a Region C. If Region C is present, in certain embodiment, it may be at most 100, 75, 50, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 nucleotide(s) in length. There may be mismatch(es) between Region C and its corresponding region in a target nucleic acid. However, the presence of the mismatch(es) need still allow for the nicking of the duplex formed between the oligonucleotide primer and the target or the nicking of the extension product of the duplex. In addition, the presence of the mismatch(es) need still allow for the extension of the nicked product that contains the 3' terminus at the nicking site to extend from that terminus in the presence of a DNA polymerase. If Region C comprises a nicking site nickable by a nicking agent that recognizes Region B, generally, the nucleotides between the 5' terminus of Region C and the nicking site forms conventional base pairs with nucleotides in the target.

In certain embodiments, an oligonucelotide primer of the present invention is immobilized, preferably at its 5' terminus. There may be a linker between the solid phase to which the primer is attached and the 5' or 3' terminus of the primer. In addition, multiple oligonucleotide primers may be immobilized to a single solid phase to produce an array of oligonucleotide primers. The multiple oligonucleotide primers may have identical sequences at discrete locations. Alternatively, they may have different sequences at distinct locations of the array. Such an array may be used to amplify multiple single-stranded nucleic acid molecules with different sequences. In certain embodiments, the amplification reactions performed at different locations of an array are physically separated, such as in microwells of a plate, so that the amplification products at different location are not mixed with each other and may be characterized individually.

The present invention is useful to amplify a single-stranded nucleic acid using any nucleic acid of interest as a target nucleic acid. The target nucleic acid includes that described above in the section for preparing single-stranded nucleic acids

and cDNA libraries using nucleic acid adaptors. In certain embodiments, the target nucleic acid, as well as other nucleic acids in a sample, is immobilized.

In certain embodiments where it may be desirable to synthesize a relatively short single-stranded nucleic acid, the target nucleic acid may be first subject to enzymatic, chemical, or mechanic cleavages. Relatively short single-stranded nucleic acids include those that have at most 200, 150, 100, 75, 50, 40, 30, 25, 20, 18, 16, 14, 12, 10, 9, 8, 7, 6, 5 or 4 nucleotides. Enzymatic cleavages may be accomplished, for example, by digesting the nucleic acid molecule with a restriction endonuclease that recognizes a specific sequence within the target nucleic acid. Alternatively, enzymatic cleavages may be accomplished by nicking the target nucleic acid with a nicking agent that recognizes a specific sequence within the nucleic acid molecule. Enzymatic cleavages may also be oligonucleotide-directed cleavages according to Szybalski (U.S. Pat. No. 4,935,357). Chemical and mechanic cleavages may be accomplished by any method known in the art suitable for cleaving nucleic acid molecules such as shearing. The cleavage product, if double-stranded, may be first denatured and subsequently anneal to an oligonucleotide primer described above.

One exemplary embodiment of enzymatic cleavage of a target nucleic acid and subsequent amplification of a single-stranded nucleic acid that is complementary to a portion of the target is illustrated in Figure 46. An oligonucleotide primer that comprises a sequence of the sense strand of a double-stranded nicking agent recognition sequence is annealed to a first region of a single-stranded target nucleic acid, whereas a partially double-stranded nucleic acid is annealed to a second region of the target nucleic acid located 5' to the first region. The double-stranded nucleic acid molecule comprises a double-stranded recognition sequence of a type II restriction enzyme recognition sequence (TRERS) in the double-stranded portion and a 3' overhang that is at least substantially, preferably exactly, complementary to a portion of the second region of the target nucleic acid. Because type IIs restiction endonuclease cleaves a nucleic acid outside its double-stranded recognition sequence, the partially double-stranded nucleic acid molecule may be designed to cleave within the duplex formed between the 3' overhang of the partially double-stranded nucleic acid molecule and the second region of the target nucleic acid. Such cleavage results in a shorter fragment of the target nucleic acid to be used as a template to amplify a single-stranded nucleic acid fragment.

The above embodiment may be used to amplify a single-stranded nucleic acid fragment that is exactly complementary to any portion of a target nucleic acid. Because oen or more nucleotides in Region B of an oligonucleotide primer are not

required to form a conventional base pair with a nucleotide of a target nucleic acid, the primer may be designed to anneal to any region of the target. In addition, the above embodiment may be used to amplify a single-stranded nucleic acid molecule of any length. As discussed above, the target nucleic acid may be cleaved at a portion where the 3' overhang of a partially double-stranded nucleic acid molecule is annealed. By designing the 3' overhang of the partially double-stranded nucleic acid molecule, the cleavage site of the target nucleic acid by a type IIs restriction endonuclease that recognizes the TRERS of the partially double-stranded nucleic acid may be at any distance from the first region of the target. Thus, the amplified single-stranded nucleic acid using the portion of the target between the position corresponding to the nicking site nickable by a nicking agent that recognizes Region B and the cleavage site of the type IIs restriction endonuclease may be of any length of interest.

In certain embodiments, the double-stranded nicking agent recognition sequence of which the sense strand is present in Region B of an oligonucleotide primer may be identical to the double-stranded TRERS. For instance, Region B of the oligonucleotide primer may consist of the sequence "5'-GAGTC-3'" recognizable by a nicking endonuclease N.BstNB I, while the TRERS in the partially double-stranded nucleic acid molecule may be

5'-GAGTC-3'
3'-CTCAG-5'

recognizable by type IIs restriction endonuclease PleI and MlyI. In such embodiments, there need be mismatch(es) between Region B of the oligonucleotide primer and the corresponding region in the target nucleic acid. In other words, one or more nucleotides in Region B do not form conventional base pairs with nucleotides in the target. The presence of mismatches prevents the cleavage of the duplex formed between the oligonucleotide primer and the first region of the target by a type IIs restriction endonuclease that recognizes the TRERS.

3. Kits for Preparing Single-Stranded Nucleic Acid Molecules

The present invention also provides kits for preparing nucleic acid molecules or for preparing and/or amplifying cDNA libraries. Such kits generally comprise a nucleic acid adaptor as described above, two single-stranded oligonucleotides that when annealed produce the nucleic acid adaptor, or an oligonucleotide primer as described above. They may further comprise at least one, two, several, or each, of the following components: (1) a nicking agent (*e.g.*, a NE or a

RE) that recognizes the NARS in the nucleic acid adaptor and/or the buffer thereof; (2) a DNA polymersase and/or the buffer thereof; (3) dNTPs (some of the dNTPs may be labeled); (4) a poly(dT) oligonucleotide primer immobilized to a solid support; (5) reverse transcriptase and/or the buffer thereof; (6) DNA ligase and/or the buffer thereof; (7) microwell plates or tubes; (8) a strand displacement-facilitator (*e.g.*, 1M trehalose); and (9) an instruction booklet for using the kit. Detailed descriptions of many of the components are provided above and in related sections of the other kits of the present invention.

4. Uses for Single-Stranded Nucleic Acid Molecules and cDNA Libraries

As discussed in detail herein above, the present invention provides methods and compositions for preparing single-stranded nucleic acid molecules and for preparing and/or amplifying cDNA libraries. These molecules and cDNA libraries will find utility in a wide variety of applications wherein it is necessary to determine the presence, or the absence, of a target nucleic acid in a sample and wherein it is desirable to compare two nucleic acid populations. Such applications include, but are not limited to, the identification and/or characterization of infectious organisms that cause infectious diseases in plants or animals, or are related to food safety, and the identification and/or characterization of genes associated with diseases in plants, animals or humans, or with desirable traits in plants or animals such as high crop yields, increased disease resistance, and high nutrition values.

For instance, the nucleic acid molecule or probe of the present invention may be specific to a nucleic acid originating from a pathogen (*e.g.*, a virus, a bacterium, a fungus, or a parasite) and thus is useful for determining the presence, or the absence, of a pathogen in a biological sample of interest. Alternatively, it may be specific to a gene known to be associated with a particular trait (*e.g.*, disease resistance or susceptibility) and thus is useful for predicting the likelihood for a particular subject from which the sample was obtained to have the particular trait.

In addition, the present invention also provides methods for simultaneously preparing multiple nucleic acid molecules or probes, such as those of which target nucleic acids are a particular cDNA population. These molecules or probes may be used to hybridize an expression array or a cDNA library prepared according to the present invention that contains another cDNA population to determine the cDNA molecules common to both cDNA populations and those present in one population but not the other. Such a determination helps the identification and/or characterization of nucleic acid molecules associated with the trait that is possessed by

only one organism from which one cDNA population is isolated, but not the other organism from which the other cDNA population is prepared.

E. Detection of pre-mRNA Alternative Splicing

In one aspect, the present invention provides methods, compounds and kits for detecting the presence, or the absence, of alternative splicing of mRNAs.

1. Methods for Detecting Pre-mRNA Alternative Splicing

The present invention provides methods for detecting the presence, or the absence, of a junction between an upstream exon (referred to as "Exon A") and a downstream exon (referred to as "Exon B") of a gene in a mRNA molecule or a cDNA molecule. The methods utilize an ODNP pair to synthesize a template nucleic acid for amplifying a single-strand nucleic acid fragment. The ODNP pair may be composed of the following two ODNPs: a first ODNP that comprises a sequence complementary to a portion of the antisense strand of Exon A near the 5′ terminus of Exon A in the antisense strand, and a second ODNP of the ODNP pair that comprises a sequence complementary to a portion of the sense strand of Exon B near the 5′ terminus of Exon B in the sense strand. In addition, at least one ODNP of the ODNP pair must further comprise a sequence of a sense strand of a nicking agent recognition sequence (NARS). The ODNP pair is then combined with a target cDNA under a condition that produces a fragment encompassed by the first and the second ODNPs if both Exon A and Exon B are present in the cDNA. The extension product is then used as a template in the presence of a DNA polymerase and a nicking agent (NA) that recognizes the NARS for amplifying a single-stranded nucleic acid fragment. The amplified nucleic acid fragment is then characterized to determine the presence, or the absence, of the junction between Exon A and Exon B.

As used herein, a "sense strand" of a cDNA molecule refers to the strand that has an identical sequence as the mRNA molecule from which the cDNA molecule is derived except that the nucleotide "U" in the mRNA is substituted by the nucleotide "T" in the cDNA molecule. An "antisense strand" of a cDNA molecule, on the other hand, refers to the strand that is complementary to the mRNA molecule from which the cDNA molecule is derived.

A target cDNA molecule refers to a cDNA molecule that is derived from a gene of interest. In other words, it is the product of reverse transcription of an mRNA molecule resulting from the transcription of the gene of interest. The target cDNA

molecule may have a partial sequence (*i.e.*, reverse transcribed from a partial mRNA molecule), but preferably a full-length sequence.

An "exon," as used herein, is any segment of an interrupted gene that is represented in the mature RNA product. An "intron" refers to a segment of DNA that is transcribed, but removed from within the transcript by splicing together the sequences (exons) on either side of it.

An exon (Exon A) is "upstream" to another exon (Exon B) in a same gene when the sequence of the sense strand of Exon A is 5′ to the sequence of the sense strand of Exon B. Exon A and Exon B may be further be referred to as an upstream exon and a downstream exon, respectively.

A nucleic acid fragment encompassed by a first ODNP and a second ODNP refers to a double-stranded nucleic acid fragment that one strand consists of the sequence of the first ODNP, the complementary sequence of the second ODNP, and the sequence between the first ODNP and the complementary sequence of the second ODNP; while the other strand consists of the complementary sequence of the first ODNP, the sequence of the second ODNP, and the sequence between the complementary sequence of the first ODNP and the sequence of the second ODNP.

"Differential splicing," also referred to as "alternative splicing," is the production of at least two different mRNA molecules from a same transcript of a gene. For instance, a particular segment of the transcript may be present in one of the mRNA molecules, but be spliced out from other mRNA molecules.

In addition to detecting a junction between two particular exons, the present invention also provides methods for detecting, in a cDNA molecule or a cDNA population, the presence, or the absence, of each potential junction between any two exons of a gene. For the convenience of describing the methods, the gene is assumed to contain n exons (n is an integer equal to or more than 2), with Exon A as the most upstream exon and Exon N as the most downstream exon (Figure 31). The present method uses the following two sets of ODNPs to prepare template nucleic acid molecules for amplifying single-stranded nucleic acid fragments: (1) a first set of ODNPs that comprise n-1 ODNPs corresponding respectively to one of the exons except Exon N, each of the ODNPs comprising a sequence complementary to a portion of the sense strand of the corresponding exon near the 5′ terminus of that corresponding exon in that sense strand, and (2) a second set of the ODNPs that comprise n-1 ODNPs corresponding respectively to one of the exons except Exon A, each of the ODNPs comprising a sequence complementary to a portion of the antisense strand of the corresponding exon near the 5′ terminus of that corresponding exon in that antisense

strand. In addition, each of the ODNPs of at least one ODNP set further comprises a NARS.

In the methods of the present invention, the above two sets of ODNPs are combined with the target cDNA molecule to be characterized or a cDNA population containing the target cDNA molecule and subsequently extended to produce fragments encompassed by one ODNP from the first ODNP set and another ODNP from the second ODNP set. The resulting extension products are then used as templates for synthesizing single-stranded nucleic acid fragments in the presence of a DNA polymerase and a nicking agent that recognizes the NARS. The amplified single-stranded nucleic acid fragments are detected and/or characterized to thereby determine the presence, or the absence, of each potential junction.

The information of the presence, or the absence, of each potential junction between any two exons of a gene in a cDNA population is useful in detecting alternative splicing of the gene in the cDNA population. The presence of more than one type of junctions for any one exon of the gene at at least one terminus of the exon indicates the mRNA of the gene was differentially spliced at that terminus of that exon.

In addition, a similar method may be used to detect alternative splicing of a gene between two biological samples. The presence, or the absence, of each potential junction between any two exons of a gene in a cDNA molecule or a cDNA population from each biological sample is determined and subsequently compared. Any difference in the presence of any particular exon junctions between these two biological samples indicates the mRNA transcribed from the gene has been alternatively spliced between the two samples.

Furthermore, one of ordinary skill in the art will appreciate that the above method may be further multiplexed to detect alternative splicing of transcripts of multiple genes in a cDNA population or between two biological samples.

a. Biological Samples and cDNA Molecules or Populations

Biological samples of the present invention include any sample that originates from an organism and that may contain an mRNA molecule of interest. In particular, the biological sample can be any cell, organ, tissue, biopsy material, etc. Of interest are samples derived from mammals (especially human beings), plants, bacteria and lower eukaryotic cells such as yeasts, fungal cells. Exemplary biological samples include, but are not limited to, a cancer biopsy, neurodegenerative plaque, cerebral zone biopsy displaying neurodegeenrative signs, a skin sample, a blood cell sample, a colorectal biopsy, etc. Exemplary cells include muscular cells, hepatic cells, fibroblasts,

nervous cells, epidermal and dermal cells, blood cells such as B-, T-lymphocytes, mastocytes, monocytes, granulocytes and macrophages.

As described above, in certain embodiments, cDNA molecules or populations from two different biological samples are compared to detect mRNA differential splicing between these samples. For such a comparison, one sample may be from a subject that is suspected of having, or is at risk for having, a genetic disease or a pathogen infection while the other sample may be a healthy, control subject. Alternatively, these two samples may be from a same biological source but at different developmental stages. In certain embodiments, one sample may be from a subject that possesses a desirable trait (*e.g.*, disease resistance), while the other may be from a subject that does not have the same trait. In other embodiments, one sample is from a subject that has been treated with a chemical (*e.g.*, a drug or a toxic material) while the other is from an untreated, control subject.

The cDNA molecules to be characterized can be derived from any mRNA molecules of interest. In certain preferred embodiments, the gene corresponding to the cDNA is associated with a disease or a disorder, particularly a human disease or disorder. In other preferred embodiments, the gene is associated with a desirable trait of the organism from which it originates. In yet other preferred embodiments, the gene is involved in the development of the subject from which it is isolated. In some embodiments, the gene participates the responses of the organism from which it is isolated to an external stimulus (*e.g.*, light, drug, and stress treatment).

In certain embodiments, the cDNA molecules are immobilized to a solid support. The methods for immobilized cDNA molecules are known in the art. For example, cDNA molecules may be immobilized via an immobilized primer used during their synthesis.

mRNA molecules of interest may be isolated and the corresponding cDNA may be synthesized according to conventional methods known in the art (*see*, the above "Target Nucleic Acids" section for the preparation of single-stranded nucleic acid probes). cDNA molecules to be characterized may be of partial length, however, preferably, of full length. Full-length cDNA molecules may be isolated by screening cDNA library with portions of the cDNA molecules as probes using conventional methods (*see*, *e.g.*, Sambrook *et al.*, *supra*). Alternatively, PCR-based method may be used for such screenings (Wilfinger *et al.*, *BioTechniques 22*: 481-486, 1997).

b. Making of Double-Stranded Template Nucleic Acids

As noted above, the present invention provides methods for detecting pre-mRNA alternative splicing, including the detection of alternative splicing at a terminus of a particular exon of a gene in a cDNA molecule or a cDNA population, and at every terminus of every exon of a gene in a cDNA molecule or a cDNA population. In addition, the methods of the present invention may be multiplexed to simultaneously detect alternative splicing at a specific (or every) terminus of a particular (or every) exon of each of multiple genes of interest in a cDNA population.

The present method uses an ODNP pair or two ODNP sets to produce double-stranded nucleic acid fragment(s) (referred to herein as "template nucleic acid fragment(s)") that comprise the nucleotide sequences of target cDNA molecules that flank specific exon-exon junction(s). In addition, the template nucleic acid fragment further comprises a NARS, which allows the amplification of a single-stranded nucleic acid fragment in the presence of DNA polymerases and a NA that recognizes the NARS. In addition, the nicking site (NS) of the NA that recognizes the NARS is located 5′ to the specific exon-exon junction so that the amplified single-stranded nucleic acid fragment contains a nucleotide sequence flanking the exon-exon junction. The amplified single-stranded nucleic acid fragment is then characterized. Because a specific end to every exon may be assigned by relying on the homology of exon-intron junctions, one of ordinary skill in the art is able to predict the sequence flanking each potential exon-exon junction. The comparison between the characteristics of the amplified single-stranded nucleic acid fragment with those of the predicted sequence flanking a specific exon-exon junction enables the determination whether the specific exon-exon junction is present in the target cDNA molecule or the target cDNA population.

i. Design of Oligonucleotide Primers (ODNPs)

The template nucleic acid of the present invention for amplifying a single-stranded nucleic acid fragment containing a specific exon-exon junction of a target cDNA may be provided by first amplifying a fragment of the target cDNA using specifically designed ODNPs. Accordingly, in one aspect, the present invention provides an ODNP pair useful for producing the above template nucleic acid.

In certain embodiments, the ODNP pair is composed of the following two ODNPs: (1) a first ODNP that comprises a sequence complementary to a portion of the antisense strand of Exon A near the 5′ terminus of Exon A in the antisense strand, and (2) the second ODNP that comprises a sequence complementary to a portion of the

sense strand of Exon B near the 5′ terminus of Exon B in the sense strand. The complementarity between the first ODNP and the portion of the antisense strand of Exon A needs not be exact, but must be sufficient to allow the ODNP to specifically anneal to that portion of Exon A. Likewise, the complementarity between the second ODNP and the portion of the sense strand of Exon B needs not be exact, but must be sufficient to allow the ODNP to specifically anneal to that portion of Exon B. Such complementarity between the first ODNP and Exon A and between the second ODNP and Exon B allows the amplification of a sequence flanking the junction between Exon A and Exon B, if such a junction is present in the cDNA. A portion of a strand of an exon is near one of the termini of the exon if that portion is within 100, 90, 80, 70, 60, 50, 40, 35, 30, 25, 20, 15, or 10 nucleotides from that terminus in that strand. Such a spacing arrangement between the two ODNPs of the ODNP pair enables the amplification of relatively short fragment encompassed by the ODNP pair.

Besides the sequence complementarity between each ODNP of the ODNP pair and one strand of its corresponding exon, either the first ODNP or the second ODNP must further comprise a sequence of a sense strand of NARS. In certain preferred embodiments, both the first ODNP and the second ODNP further comprise a NARS. The presence of the NARS allows the amplified nucleic acid fragments encompassed by the first ODNP and the second ODNP to function as a template nucleic acid for amplifying a single-stranded nucleic acid fragment in the presence of a DNA polymerase and a NA that recognizes the NARS, as discussed in detail below.

The NARS of which one strand is present in the ODNP may be of various types. For instance, it may be a restriction endonuclease recognition sequence (RERS). Alternatively, it may be a nicking endonulease recognition sequence (NERS). If the NARS is a RERS, the amplification of the template nucleic acid is typically performed in the presence of a modified deoxyribonucleoside triphosphate so that the amplified template contains a hemi-modified RERS, rather than a RERS consisting of native nucleotides. The presence of the hemi-modified RERS allows the cleavage of only one strand of the template by a restriction endonuclease that recognizes the RERS. If the NARS is a NERS, the sequence of the sense strand of the NERS in the ODNP needs to be in the direction such that the nicking site produced by a NE that recognizes the NERS is located 5′ to the junction between Exon A and Exon B. Such an arrangement allows the subsequently amplified single-stranded nucleic acid fragment to comprise a sequence flanking the junction between Exon A and Exon B, as described in detail below. In some embodiments where only one ODNP of the ODNP pair contains a sequence of a sense strand of a NERS, the other ODNP that does not contain the

NERS sequence may comprise a sequence of one strand of a RERS. In these embodiments, the amplification of the template nucleic acid is performed in the absence of any modified deoxynucleoside triphosphate so that the RERS in the amplified template nucleic acid fragment is not hemi-modified. The RERS is preferably recognizable by a type IIs restriction endonuclease. The presence of such a RERS allows the production of a relatively small nucleic acid fragment to be used as a template for subsequently synthesizing a single-stranded nucleic acid fragment. The production of the relative small nucleic acid fragment may be achieved by digesting amplified template nucleic acid fragments encompassed by the two ODNPs of the ODNP pair with a type IIs restriction endonuclease that recognizes the RERS.

The present invention also provides ODNP sets useful for detecting the presence, or the absence, of each potential junction between any two exons of a gene in a cDNA molecule or a cDNA population. For the convenience of describing the ODNP sets, the gene of interest is assumed to contain n exons where n is an integer equal to or more than 2, and the most upstream exon and the most downstream exon are denoted as Exon A and Exon N, respectively. The following two sets of ODNPs are useful for preparing template nucleic acid molecules that together contain all the exon-exon junctions of the gene present in the cDNA molecule or the cDNA population to be characterized: (1) a first set of ODNPs that comprise n-1 ODNPs corresponding respectively to one of the exons except Exon A, each of the ODNPs comprising a sequence complementary to a portion of the sense strand of the corresponding exon near the 5′ terminus of that corresponding exon in the sense strand, and (2) a second set of ODNPs that comprise n-1 ODNPs corresponding respectively to one of the exons except Exon N, each of the ODNPs comprising a sequence complementary to a portion of the antisense strand of the corresponding exon near the 5′ terminus of that corresponding exon in the antisense strand. The complementarity between each ODNP of the first set and the portion of the sense strand of the corresponding exon need not be exact, but must be sufficient to allow each ODNP to specifically anneal to the complementary portion of the corresponding exon. Likewise, the complementarity between each ODNP of the second set and the portion of the antisense strand of the corresponding exon need not be exact, but must be sufficient to allow each ODNP to specifically anneal to the complementary portion of the corresponding exon. Such complementarity between each ODNP of both sets and its corresponding exon allows the amplification of a sequence flanking every exon-exon junction of the gene present in the cDNA molecule or the cDNA population to be characterized.

Besides the sequence complementarity between each ODNP of the two ODNP sets and one strand of the corresponding exon, each fragment synthesized using one ODNP from the first set and another ODNP from the second set as primers and the cDNA molecule(s) derived from the gene of interest as a template must contain a NARS, including but not limited to, a NERS or a hemi-modified RERS. In certain preferred embodiments, either each ODNP of the first set or each ODNP of the second set contain a sequence of a sense strand of a NARS. In some other preferred embodiments, each ODNP of both the first set and the second set contains a NARS.

Similar to the NARS of which one strand is present in the ODNP pairs for detecting the presence, or the absence, of a particular exon-exon junction, the NARS of which one strand is present in the ODNP sets may also be of various types. For instance, it may be a restriction endonuclease recognition sequence (RERS). Alternatively, it may be a nicking endonulease recognition sequence (NERS). If the NARS is a RERS, the amplification of the template nucleic acids is typically performed in the presence of a modified deoxyribonucleoside triphosphate so that the amplified templates contain a hemi-modified RERS, rather than a RERS consisting of native nucleotides. If the NARS is a NERS, the sequence of the sense strand of the NERS in the ODNP needs to be in the direction such that the nicking site produced by a nicking endonucolease that recognizes the NERS is located 5′ to an exon-exon junction. Such an arrangement allows the subsequently amplified single-stranded nucleic acid fragments to comprise the sequences flanking an exon-exon junction. In some embodiments where only the ODNPs of one ODNP sets contain a sequence of a sense strand of a NERS, the ODNPs of the other set that do not contain the NERS sequence may comprise a sequence of one strand of a RERS. In these embodiments, the amplification of the template nucleic acids is performed in the absence of any modified deoxyribonucleoside triphosphate so that the RERS in the amplified template nucleic acid fragments is not hemi-modified. The RERS is preferably recognizable by a type IIs restriction endonuclease. The presence of such a RERS enables the production of relatively small nucleic acid fragments by digesting, with a type IIs restriction endonuclease that recognizes the RERS, the amplified template nucleic acid fragments encompassed by one ONDP from the first set and another ODNP from the second set. The produced relatively small nucleic acid fragments may be used as templates for synthesizing short single-stranded nucleic acid fragments that comprise sequences that flank exon-exon junctions. The short lengths of the synthesized fragments facilitate their detection and/or characterization by certain separation and analytic techniques.

General techniques for designing sequence-specific primers are well known and also described above for genetic variation detection methods of the present invention.

ii. Nucleic Acid Hybridization, Extension, Amplification and Digestion

The above ODNP pair or sets are combined with a target cDNA molecule or a cDNA population that contains the target cDNA molecule under conditions that allow the specific annealing of the ODNPs with their complementary sequences in the cDNA molecule(s). Determining and providing such conditions are within the skill of one of ordinary skill in the art. Exemplary conditions are also provided in the Hybridization section for genetic variation detection methods of the present invention and in examples.

After annealing, the ODNPs is extended to produce a fragment that is encompassed by two ODNPs and comprises a portion of the target cDNA molecule using various methodologies known in the art and described in the related section for genetic variation detection methods of the present invention. For determining the presence, or the absence, of a particular exon-exon junction (*e.g.*, the junction between an upstream Exon A and a downstream Exon B), the target cDNA must contain both Exon A and Exon B for extension from both the ODNP complementary to a portion of Exon A (referred to as "the first ODNP") and the ODNP complementary to a portion of Exon B (referred to as "the second ODNP") (Figures 30A and 30B). If a junction between Exon A and Exon B is present in the target cDNA molecule, a small fragment encompassed by the first ODNP and the second ODNP that contains the junction between Exon A and Exon B will be amplified (Figure 30A). If a junction between Exon A and Exon B is absent in the target cDNA, but both Exon A and Exon B are present (*i.e.*, an additional segment is between Exon A and Exon B), a larger fragment encompassed by the first ODNP and the second ODNP that contains the additional segment in addition to portions of Exon A and Exon B will be amplified (Figure 30B).

For the determination of the presence, or the absence, of every potential junction between any two exons of a gene of interest in a target cDNA molecule or a cDNA population containing the target cDNA molecule, two sets of ODNPs as described above are extended to produce various template nucleic acid fragments. Because the two sets of ODNPs together contain the ODNPs that are capable of annealing to each strand of the target cDNA molecule near the 5′ terminus of every exon that may be linked to another exon in that strand, the amplified template nucleic acid fragments contain all the exon-exon junctions in the target cDNA molecule.

The above notion may be further explained by referring to Figure 31. Figure 31A illustrates an exemplary gene of interest that contains n exons, with Exon A as the most upstream and Exon N as the most downstream exon. Figures 31B and 31C illustrate two cDNA molecules ("Target cDNA I" and "Target cDNA II") resulting from alternative splicing of the same transcript of the gene of interest. In particular, Exon B is present in Target cDNA I, but not in Target cDNA II where Exon A and Exon C are linked directly together. When two sets of ODNPs specifically designed for the gene of interest as described above are combined with Target cDNA I under a condition that allows the ODNPs to anneal to the complementary sequences in their corresponding exons, nucleic acid fragments that are encompassed by the first ODNP (which anneals to the antisense strand of Exon A near the 5′ terminus of Exon A in the antisense strand) and the second ODNP (which anneals to the sense strand of Exon B near the 5′ terminus of Exon B in the sense strand) (denoted as "Template 1-2") and by the third ODNP (which anneals to the antisense strand of Exon B near the 5′ terminus of Exon B in the antisense strand) and the fourth ODNP (which anneals to the sense strand of Exon C near the 5′ terminus of Exon C in the sense strand) (denoted as "Template 3-4") as well as fragments encompassed by other ODNP pairs may be produced. These fragments comprise portions of Target cDNA I that contain exon-exon junctions. In other words, each fragment comprises a sequence flanking a particular exon-exon junction. On the other hand, when the two sets of ODNPs specifically designed for the gene of interest according to the present invention are combined with Target cDNA II to allow the annealing between the ODNPs and the complementary sequences in their corresponding exons, nucleic acid fragments encompassed by the first ODNP and the four ODNP (denoted as "Template 1-4") as well as fragments encompassed by other ODNP pairs may be amplified. Because of the absence of Exon B in Target cDNA II, Template 1-2 and Template 3-4 may not be produced. However, because there is no junction between Exon A and Exon B or between Exon B and Exon C, the absence of Template 1-2 or Template 3-4 does not contradict the notion that the amplification products using the two sets of ODNPs of the present invention as primers and a target cDNA molecule as a template together contain all the exon-exon junctions present in the target cDNA molecule.

The presence of multiple ODNPs in the amplification reaction mixture may result in multiple products from a given primer. For instance, in Figure 32B, besides amplifying Template 1-2 with the second ODNP, the first ODNP may also pair up with other ODNPs from the first ODNP set (*e.g.*, the fourth ODNP) to amplify larger nucleic acid fragments (*e.g.*, Template 1-B-2 which comprises a portion of Exon A, full

length Exon B and a portion of Exon C). However, because smaller fragments are amplified more efficiently than larger fragments, Template 1-2 molecules accumulate faster than the other larger template molecules with the increase of amplification cycles.

As noted above, in certain embodiments, the NARS of which one strand is present in the ODNPs of the present invention is a RERS. In certain embodiments, to obtain template nucleic acids that contain a hemimodified RERS, the ODNPs are extended in the presence of a modified deoxyribonucleoside triphosphate. Suitable modified deoxyribonucleoside triphosphates and the amplification of nucleic acid fragments in the presence of these modified deoxyribonucleoside triphosphates are known in the art and also described in the related section for the genetic variation detection methods of the present invention.

Also as discussed above, in certain other embodiments, each ODNP of one ODNP set comprises a sequence of a sense strand of a NERS, while each ODNP of the other ODNP set comprises a sequence of one strand of a RERS. When these two sets of ODNPs are used as primers for synthesizing template nucleic acid fragments containing the sequences flanking exon-exon junctions in the absence of any modified deoxyribonucleotide triphosphate, the resulting template nucleic acids contain both the NERS and the RERS. These template nucleic acids may be digested with a restriction endonuclease that recognizes the RERS and cleaves at both strands of the template nucleic acids. Such digestion produces smaller template nucleic acids and consequently enables the amplification of shorter single-stranded nucleic acid fragments when the smaller template nucleic acids are used as templates for the above amplification.

c. Amplification and Detection of Single-Stranded Nucleic Acid Fragments

The template nucleic acids described above are used as templates for amplifying single-stranded nucleic acid fragments in the presence of a nicking agent that recognizes the NARS in the ODNP(s) used in the synthesis of the templates and a DNA polymerase. Such amplification is essentially as described above for genetic variation detection methods. The resulting single-stranded nucleic acid fragments are subsequently characterized to determine the presence, or the absence, of a particular exon-exon junction in a target cDNA molecule or in a cDNA population that contains the target cDNA molecule.

In the embodiments that two ODNP sets are used to synthesize template nucleic acids, as discussed above, although the majority of synthesized template nucleic acids contain only one exon-exon junction, some template nucleic acids contains more

than one exon-exon junctions (*e.g.*, Template 1-B-2 in Figure 31B). The presence of these relatively long templates in the reaction mixture for amplifying single-stranded nucleic acids may result in the amplification of longer single-stranded nucleic acids than those using templates containing only one exon-exon junction. In certain embodiments, the amplification of these longer single-stranded nucleic acids may be eliminated or reduced by performing the amplification reaction under conditions that inhibit or prevent the amplification of relatively long single-stranded nucleic acid molecules. For instance, a DNA polymerase that lacks a strand displacement activity may be used to amplify only short single stranded nucleic acid molecules. The lack of the strand displacement activity prevents the DNA polymerase from extending from a 3′ terminus at a nicking site produced by a nicking agent where the portion of the nicked strand containing the 5′ terminus at the nicking site still anneals to its complementary sequence of the unnicked strand. Only when the portion of the nicked strand containing the 5′ terminus at the nicking site is relatively short (*e.g.*, containing fewer than 17 nucleotides) so that it can readily dissociate from its complementary strand, can it be amplified by a DNA polymerase that does not have a strand displacement activity in the presence of a nicking agent.

The single-stranded nucleic acid molecules amplified from the template nucleic acids that contain only one exon-exon junction may preferably be at most about 100, 90, 80, 70, 60, 50, 40, 35, 30, 25, 20, 25, 10, 8, 6, 4, or 2 nucleotides in length and may be detected and/or characterized by any methods known in the art suitable for detecting or characterizing small single-stranded nucleic acid molecules. Exemplary methods are discussed above in relevant sections for genetic variation detection methods and for target nucleic acid detection methods of the present invention. The characteristics of the amplified single-stranded nucleic acid fragments (*e.g.*, the mass to charge ratio obtained by mass spectrometric analysis) are subsequently compared with those of single-stranded nucleic acid fragments predicted in view of the positions and compositions of the ODNPs used in preparing template nucleic acid fragments and with the assumption that the junction between the two exons to which the ODNPs are complementary is present. If the characteristics of the amplified and the predicted nucleic acid fragments are identical, the particular exon-exon junction that was assumed to be present in the target cDNA molecule is in fact present in that target cDNA molecule. The prediction of the sequence and the characteristics (*e.g.*, mass to charge ratio) of the single-stranded nucleic acid fragment that would be amplified is based on the knowledge about consensus sequences near exon-intron junctions (also referred to as "splicing junctions"). This knowledge allows one of ordinary skill in the art to

pinpoint the exon-intron junctions and thus predicts the exact locations of exon-exon junctions when the intron between the two exons has been spliced out.

2. Kits for Pre-mRNA Differential Splicing Detection

The present invention also provides kits for detecting pre-mRNA differential splicing. Such kits generally comprise the ODNP pairs described above that are useful for preparing a template nucleic acid that contains a sequence flanking an exon-exon junction if the junction is present in a target cDNA molecule. They may also further comprise at least one, two, several, or each of the following components: (1) a nicking agent (*e.g.*, a NE or a RE) that recognizes the nicking agent recognition sequence in at least one primer of the ODNP pair; (2) the buffer for nicking agent (1); (3) a RE that recognizes the RERS that is present in at least one primer of the ODNPs; (4) the buffer for RE (3); (5) a DNA polymerase useful for making the template nucleic acid (*i.e.*, extending from the 3′ termini of the above ODNP pair); (6) the buffer for the DNA polymerase (5); (7) a DNA polymerase for amplifying single-stranded nucleic acid fragments; (8) the buffer for DNA polymerase (7); (9) dNTPs; (10) a modified dNTP; (11) a control template and/or control oligonucleotide primers for amplifying the template nucleic acid; (12) a chromatography column; (13) Buffer A for the chromatography column; (14) Buffer B for the chromatography column; (15) water; (16) a strand displacement facilitator (*e.g.*, 1M trehalose); (17) microtiter plates or microwell plates; (18) oligonucleotide standards (*e.g.*, 6mer, 7 mer, 8mer, 12mer and 16mer); (19) an instruction booklet for using the kit; and (20) an access code for a software used in designing and/or ordering ODNP pairs. Detailed descriptions of many of the above components are provided above.

3. Applications of Pre-mRNA Differential Splicing Detection Methods

Alternative pre-mRNA splicing is an important mechanism for regulating gene expression in higher eukaryotes. By recent estimates, the primary transcripts of ~30% of human genes are subject to alternative splicing, often regulated in specific spatial/temporal patterns during normal development. In complex genes alternative splicing can generate dozens or even hundreds of different mRNA isoforms from a single transcript (Breitbart and Nadal-Ginard, *Annu. Rev. Biochem. 56*: 467-95, 1987; Missler and Sudhof, *Trends Genet 14*: 20-6, 1998; Gascard *et al.*, *Blood 92*:4404-14, 1998). In many cases the alternatively spliced exon encodes a protein domain that is functionally important for catalytic activity or binding interactions, the resulting proteins can exhibit different or even antagonistic activities.

As discussed in detail herein above, the present invention provides methods, compositions, and kits for detecting pre-mRNA alternative splicing, including the detection of alternative splicing at a terminus of a particular exon of a gene in a cDNA molecule or a cDNA population, and at every terminus of every exon of a gene in a cDNA molecule or a cDNA population. Due to the importance of pre-mRNA splicing, these methods, compositions and kits will find utility in a wide variety of applications such as disease diagnosis, prediposition, and treatment, crop cultivation and animal breeding, development regulations of plants and animals, drug development and manipulation of responses of an organism to external stimuli (*e.g.*, extreme temperatures, poison, and light).

For instance, the present method may be used to identify and/or characterize pre-mRNA splicing patterns unique to a pathological condition. Abnormal pre-mRNA splicings in many genes have been implicated in various diseases or disorders, especially in cancers. In small cell lung carcinoma, the gene of protein p130, which belongs to the retinoblastoma protein family is mutated at a consensus splicing site. This mutation results in the removal of exon 2 and the absence of synthesis of the protein due to the presence of a premature stop codon. Likewise, in certain non small cell lung cancers, the gene of protein p161NK4A, which is an inhibitor of cyclin depndant kinase cdk4 and cdk6, is mutated at a donor splicing site. This mutation results in the production of a truncated short half-life protein. In addition, WT1, the Wilm's tumor suppressor gene, is transcribed into several messenger RNAs generated by alternative splicings. In breast cancers, the relative proportions of different variants are modified in comparison to healthy tissue, hence yielding diagnostic tools or insights into understanding the importance of the various functional domains of WT1 in tumoral progression. A similar alteration process affecting ratios among different mRNA forms and protein isoforms during cell transformation is also found in neruofibrin NF1. Moreover, in head and neck cancer, one of the mechanisms by which p53 is inactivated involved a mutation at a consensus splicing site. Furthermore, an altered splicing pattern of the IRF-1 tumor suppressor gene transcript results in the inactivation of the tumor suppressor and an acceleration of exon skipping in IRF-1 mRNA is indicative of a number of hematopoietic disorders including overt leukemia from myelodysplastic syndrome, acute myeloid leukemia, and the myelodysplastic syndromes (U.S. Pat. No. 5,643,729).

The present method may be used to compare the splicing pattern of the transcript of a gene that is known or suspected to be associated with a disease (or disorder) condition, and to identify exons of which presence or absence is unique to the

disease (or disorder) condition or to identify the alteration in the ratio among different splicing variants unique to the disease (or disorder) condition. The identification of the exons that are absence in a disease (or disorder) condition may indicate that the domains encoded by the exons are important to the normal functions of healthy cells and that the signaling pathways involving such domains may be restored for therapeutical purposes. On the other hand, the identification of the exons uniquely present in a disease (or disorder) condition may be used as diagnostic tools and the domains encoded thereof be considered as screening targets for compounds of low molecular weight intended to antagonize signal transduction mediated by the domains. In addition, the antibodies with specific affinities to these domains may also be used as diagnostic tools for the disease (or disorder) condition.

The present method may also be used to identify and/or characterize the pre-mRNA differential splicing important in organism development. Alternative splicing plays a major role in sex determination in Drosophila, antibody response in humans and other tissue or developmental stage specific processes (Chabot, *Trends Genet. 12:* 472-8; Smith *et al., Annu. Rev. Genet. 23*: 527-77, 1989; Breitbart *et al., Cell 49:* 793-803, 1987). Thus, the present method may be used to compare pre-mRNA splicing patterns of a gene that is known or suspected to be involved in development regulation at different developmental stages. The identification and/or characterization of the presence of differential splicing in the gene may provide guidance in regulating the corresponding development process to obtain desirable traits (*e.g.*, bigger fruits, higher protein or oil content seeds, higher milk production).

The present method may also be used to identify and/or characterize the pre-mRNA differential splicing important in organisms' responses to various external stimuli. The pre-mRNA splicing pattern of a gene that is known or suspected to play a role in response to a particular stimulus (*e.g.*, pathogen attack) of an untreated organism may be compared with that of an organism subjected to the stimulus. The identification and/or characterization of the splicing pattern unique to the organism subjected to the stimulus may provide guidance in manipulating the corresponding response process to enhance (if the response is desirable) or to reduce/eliminate (if the response is undesirable) the response.

The following examples are provided by way of illustration and not limitation.

EXAMPLES

EXAMPLE 1

DETECTION OF OLIGONUCLEOTIDE FRAGMENTS WITH ELECTROSPRAY-LIQUID CHROMATOGRAPHY/MASS SPECTROMETRY

This example discloses the use of Electrospray-Liquid-Chromatography/Mass Spectrometry (ES-LC/MS) for determining the molecular weight of single-strand oligonucleotide (ODN) fragments of 4, 6, 8 and 10 nucleotides in length (Table 3).

Oligonucleotides (ODNs) were synthesized by Midland Certified reagents of Midland TX. The ODNs were diluted to a concentration of 0.5 nm/µl in 0.01 M Tris-HCl, 0.1 mM EDTA to create a stock solution. Each stock solution of ODN was subsequently diluted 1:10, 1:100 and 1:1000 in purified water. Five microliters of each dilution was injected into a electrospray-liquid-chromatography/mass spectrometry, time-of flight (ES-LC/MS-TOF) system using negative ion full scan, cone = 35 volts, source at 100°C, desolvation at 250°C, with a Xterra column, C8, 2.1 x 50 mm, with a flow rate of 300 µl/min., direct, running isocratic in water, methanol +0.05% TEA.

The chromatography was performed using the following system: a ProStar Helix System (catalog # Helixsys01) that is composed of two pumps, a column oven, a UV detector, a degasser, a mixer and an autoinjector. The column is a Varian Microsorb MV (catalog number R0086203F5), C18 packing with 5 µM particle size, with 300 Angstroms pore size, 4.6 mm x 50 mm. The column was run at 30°C to 40°C with a gradient of acetonitrile in 100 mM triethylamine acetate (TEAA) and 0.1 mM EDTA. The following HPLC method was used to separate the fragments on the column: Buffer A is 100 mM TEAA with 0.1 mM EDTA, Buffer B is 100 mM TEAA with 0.1 mM EDTA and 25% (V/V) acetonitrile, at 0-3 minutes there is a gradient of 20% B to 25% B, at 3.01 minutes to 4 minutes, there is a ramp to 45% B, at 4.01 to 4.5 minutes there is a ramp to 95% B, at 4.51 minutes there is 1 minute hold at 20% B to re-equilibrate the column. The column was run at 40°C by adjusting the column oven to 40°C. The flow rate was 1.5 ml per minute. About 200 nanogram of fragment was injected per 10 µl volume. The UV detector measures the effluent of the column.

TABLE 3

ODNs AND CORRESPONDING SEQUENCES AND MOLECULAR WEIGHTS

ODN	SEQUENCE	LENGTH	MOLECULAR WEIGHT (Daltons)	SEQ ID NO.
Fok0001	5′-ACGA-3′	4-mer	1181 Da	SEQ ID NO: 1
Fok0002	5′-ACGATG-3′	6-mer	1816 Da	SEQ ID NO: 2
Fok0003	5′-ACGATGCA-3′	8-mer	2418.6 Da	SEQ ID NO: 3
Fok0004	5′-GAACATCCAT-3′	10-mer	2996 Da	SEQ ID NO: 4

The molecular weights of the ODNs detailed in Table 3 were measured by ES-LC/MS and the lower limit of detection was determined.

For the 4-mer with an expected molecular weight of 1181 Daltons, the total extracted ion current between 1181.102 and 1181.7 was 0.41. Figure 7 shows the chromatogram of 1/10 dilution of the 4-mer with the 1181.4 peak normalized to 100%. The 1181.4 peak represents the singly charged species in which the mass/charge (m/z) ratio = 1. A doubly charged 4-mer with a mass of 590.2 appears at 80% the intensity of the 1181.4 species. The 590.2 mass represents an m/z value of 0.5 that of the singly charged species (*i.e.*, the 590.2 species has two charges). Also note that an n+0.5 and n+1 charge was seen at 590.7 and 591.2. This is typical with electrospray ionization with polymeric molecules.

At a 1/100 dilution, the mass spectra looks similar to the 1/10 dilution, with peaks of expected molecular weight 1181.4 and 590.2. The peak at 1203.5 is the Na+ adduct (+23 Daltons) of the 4-mer. The peaks at 621.2, 659.2, 868.3, 948.3 are either fragmentation products or background (Figure 8).

In Figure 9, the dilution at 1/1000 is shown which is about at the lower limit of detection for the ES-TOF. Peaks at 1181.5 and 590.2 are clearly visible above the background. More "background" is visible in the spectra as the lower limits of detection are being pushed.

For the 6-mer with an expected molecular weight of 1816 Daltons, the total extracted ion current between 906.6 and 907.5 was 0.35 (the doubly charged species). Figure 10A shows the chromatogram of 1/10 dilution of the 6-mer with the 906.7 peak normalized to 100% (this is the doubly charged species where m/z =0.5 (mass/charge = 0.5)). The 906.7 mass represents an m/z value of 0.5 that of the singly charged species (*i.e.*, the 906.7 species has two charges). Also note that an n+0.5 and n+1 charge is seen at 907.3 and 907.8. This is typical with electrospray ionization with

polymeric molecules. In Figure 10B, the mass spectra shows peaks of expected molecular weight 1815.6. The peak at 1837.5 is the Na+ adduct (+23 Daltons) of the 6-mer. Figure 11 shows the chromatogram of the 1/100 dilution of the 6-mer with the 906.8 peak normalized to 100%. In Figure 12, the dilution at 1/1000 is shown which is about at the lower limit of detection for the ES-TOF. The peak at 906.8 is clearly visible above the background. More "background" is visible in the spectra as the lower limits of detection are being pushed.

For the 8-mer with an expected molecular weight of 2418.6 Daltons, the total extracted ion current between 1207.602 and 1210.1 was 0.37 (the doubly charged species). Figures 13A and 13B show the chromatogram of 1/10 dilution of the 8-mer with the 1208.4 peak normalized to 100% (this is the doubly charged species where m/z =0.5 (mass/charge = 0.5). The 805.3 mass represents an m/z value of 0.33 that of the singly charged species (*i.e.*, the 805.3 species has three charges.). Also note that an n+0.5 and n+1 charge is seen at 805.9 and 806.3. This is typical with electrospray ionization with polymeric molecules. At a 1/100 dilution (Figure 14), the mass spectrum shows peaks of expected molecular weight 1207.9 (doubly charged) and 804.9 (triply charged).

For the 10-mer with an expected molecular weight of 2996 Daltons, the total extracted ion current between 1496.199 and 1500.182 was 0.37 (the doubly charged species). Figures 15A and 15B shows the chromatogram of 1/10 dilution of the 10-mer with the 1497.0 peak normalized to 100% (this is the doubly charged species where m/z =0.5 (mass/charge = 0.5). The 1497.0 mass represents an m/z value of 0.5 that of the singly charged species (*i.e.*, the 1497 species has two charges.). Also note that an n+0.5 and n+1 charge is seen at 748.5 and 749 for the species with 4 charges. This is typical with electrospray ionization with polymeric molecules. Figure 15B shows the mass spectra with a peak of expected molecular weight 2996. The peak at 3017.9 is the Na+ adduct (+23 Daltons) of the 10-mer. Figure 16 shows the 1/100 dilution of the 10-mer with the 1497.1 peak normalized to 100%. In Figure 17, the dilution at 1/1000 is shown which is about at the lower limit of detection for the ES-TOF. The peak at 1497.1 (doubly charged species) is clearly visible above the background as is the triply charged species (997.6) and the species with 4 charges (748). More "background" is visible in the spectra as the lower limits of detection are being pushed.

The lower limits of detection were calculated to be about 2.5 picomole of either the 4-, 6-, 8-, or 10-mer. Electrospray-TOF has great benefit over MALDI-

TOF because the electrospray can be quantified. The data indicate that as little as 1/100th of a 10-microliter PCR reaction could be detected using the ES-TOF.

EXAMPLE 2

SEPARATION AND IDENTIFICATION OF OLIGONUCLEOTIDE FRAGMENTS THAT DIFFER BY A SINGLE NUCLEOTIDE USING HPLC

This example describes the separation and identification of short DNA fragments by liquid chromatography. The detection of DNA fragments is by UV absorbance and the identification is by retention time on the column compared to standards.

The chromatography system is from Varian (Walnut Creek, CA) and is a ProStar Helix System (catalog # Helixsys01) that is composed of two pumps, a column oven, a UV detector, a degasser, a mixer and an autoinjector. The column is a Varian Microsorb MV (catalog number R0086203F5), C18 packing with 5 uM particle size, with 300 Angstroms pore size, 4.6 mm x 50 mm. The column was run at 30°C to 40°C with a gradient of acetonitrile in 100 mM triethylamine acetate (TEAA) and 0.1 mM EDTA. The type of gradient is described in the text.

The following genotyping fragments, each containing a specific Single Nucleotide Polymorphism were tested and successfully separated.

4-merA: 5′-ACGA-3′ (SEQ ID NO:1)
6-merA: 5′-ACGATG-3′ (SEQ ID NO:2)
8-merA: 5′-ACGACGCA-3′ (SEQ ID NO:5)
8-merB: 5′-ATGACGCA-3′ (SEQ ID NO:6)
8-merC: 5′-ACGATGCA-3′ (SEQ ID NO:3)
10-merA: 5′-GAATATCCAT-3′ (SEQ ID NO:7)
10-merB: 5′-GAATATCCAC -3′ (SEQ ID NO:8)
10-merC: 5′-GAACATCCAT –3′ (SEQ ID NO:4)

The polymorphisms in the 8-mers and 10-mers are underlined. The 8-mers B and C differ from 8-mer A by only a single base. The 10-mers B and C differ from 10-mer A by only a single base.

The following HPLC method was used to separate the fragments on the column: Buffer A is 100 mM TEAA with 0.1 mM EDTA, Buffer B is 100 mM TEAA with 0.1 mM EDTA and 25% (V/V) acetonitrile, 0-3 minutes there is a gradient of 20% B to 25% B, at 3.01 minutes to 4 minutes, there is a ramp to 45% B, at 4.01 to 4.5

minutes there is a ramp to 95% B, at 4.51 minutes there is 1 minutes hold at 20% B to re-equilibrate the column. The column was run at 40°C by adjusting the column oven to 40°C. The flow rate was 1.5 ml per minute. The injection volume was 10 microliters and 200 nanogram of fragment was injected per 10 microliter volume. Different combinations of the 4-mer, 6-mer, 8-mer and 10-mer were injected to determine the chromatographic behavior.

The first result is shown in Figure 18. In Trace 1 of Figure 18, all 8 fragments composed of the 4-mer, 6-mer, 8-mer and 10-mer were separated. All three 8-mers and all three 10-mers were separated even though they differed by only a single base. The fragments are single stranded. The order of elution in Trace 1 is (from left to right): 4-mer, 6-mer, 8-merB, 8-merA, 10-merA, 8-merC, 10-merB, 10-merC. In Trace 2, the 6-mer and 10-merC were co-injected and the elution times of the 6-mer and 10-merC were the same as seen in Trace 1. In Trace 3, the three 10-mers were co-injected and separated. The elution times of the three 10-mers were the same as seen in Trace 1. In Trace 4, the three 8-mers were co-injected and separated. The elution times of the three 8-mers were the same as seen in Trace 1. Trace 5 shows a single peak of 8-merA and Trace 6 shows a single Trace of 8-merB. Genotypes can be directly inferred from the retention times during the chromatography, even from fragments that differ by only a single base.

Figures 19A and 19B shows HPLC fractionation and detection of three 8-mers (Fig. 19A) and three 10-mers (Fig. 19B). In Figure 19A, the "T" allele at position 2 of the 1st 8-mer is discriminated from the "C" allele at position 2 of the 2nd 8-mer and the "T" allele at position 5 of the 2nd 8-mer from the "C" allele at position 5 of the 3rd 8-mer. In Figure 19B, the "T" allele at position 4 of the 3rd 10-mer is discriminated from the "C" allele at position 4 of the 2nd 10-mer and the "C" allele at position 10 of the 2nd 10-mer from the "T" allele at position 10 of the 1st 10-mer. Genotypes can be directly inferred from the retention times during the chromatography, even from fragments that differ by only a single base.

In Figure 20A, one 4-mer (4-merA), one 6-mer (6-merA), three 8-mers (8-merA, 8-merB and 8-merC) and three 10-mers (10-merA, 10-merB, and 10-merC) are separated. In Figure 20B, two 6-mers are shown eluting between 2 and 3 minutes. The 6-mers were generated by double Fok I digestion of a 41-mer which contained the forward and reverse Fok I recognition site which was separated by 6 nucleotides.

EXAMPLE 3

SEPARATION OF GENOTYPING FRAGMENTS BY LIQUID CHROMATOGRAPHY AND DETECTION WITH A UV DETECTOR AND TIME OF FLIGHT MASS SPECTROMETER

The following example describes the amplification of a specific sequence from the human genome in which the primers contain one or both of the Bsl I and N.BstNB I restriction endonuclease recognition sequences. The resulting amplicon contains a double cutting site that liberates a small oligonucleotide fragment, which is then subjected to a chromatography step and identified by mass to charge ratio.

The 50 µl PCR reactions were composed of 25 ng genomic DNA, 0.5 µM each forward and reverse primers, 10 mM Tris pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 µM each dNTP, 1 Unit DNA Polymerase (MasterAmp ™ Taq DNA Polymerase from Epicentre Technologies, Madison WI or Vent exo- Polymerase New England BioLabs, Beverly MA). Thermocycling conditions were as follows: 95°C for 3 minutes initial denaturation; 30 cycles of 92°C for 40 seconds, 60°C for 30 seconds, 72°C for 30 seconds. A MJ Research PTC-100 thermocycler (MJ Research, Watertown, MA) was used for all PCR reactions. Primers were purchased from MWG Biotech (High Point, NC).

After the thermocycling was complete, an enzyme mixture was prepared containing N.BstNB I, Bsl I and 10x N.BstNB I buffer (New England BioLabs Beverly, MA). The mixture was added to each well to contain a final concentration one unit each enzyme in 150 mM KCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.5. The reaction was carried out at 55C for more than 60 minutes. The reaction was injected directly without any further manipulation.

The chromatography system is an Varian Prostar Helix system composed of a binary pump, degasser, a column oven, a diode array detector, and thermostatted microwell plate autoinjector (Varian Inc. Walnut Creek, CA). The column is a Varian Microsorb MV, incorporating C18 packing with 3 µM particle size, with 300 Angstrom pore size, 2.1 mm x 50 mm (Varian Inc. Walnut Creek, CA). The column was run at 30°C with a gradient of acetonitrile in 5 mM Triethylamine acetate (TEAA). Buffer A is 5 mM TEAA, buffer B is 5 mM TEAA and 25% (V/V) acetonitrile. The gradient begins with a hold at 10%B for one minute then ramps to 50%B over 4 minutes followed by 30 seconds at 95%B and finally returning to 10%B for a total run time of six minutes. The column temperature was held constant at 30°C. The flow rate was 0.416 ml per minute. The injection volume was 10 microliters. Flow into the mass spectrometer was 200ul/min, half the LC flow was diverted to waste using a tee. The mass spectrometer is a Micromass LCT Time-of-Flight with an electrospray inlet

(Micromass Inc. Manchester UK). The samples were run electrospray negative mode with a scan range from 700 to 2300 amu using a 1 second scan time. Instrument parameters were: TDC start voltage 700, TDC stop voltage 50, TDC threshold 0, TDC gain control 0, TDC edge control 0, Lteff 1117.5, Veff 4600. Source parameters: Desolvation gas 862 L/hr, Capillary 3000V, Sample cone 25V, RF lens 200V, extraction cone 2V, desolvation temperature 250C, Source temperature 150C, RF DC offset 1 4V, FR DC offset 2 1V, Aperture 6V, acceleration 200V, Focus, 10V, Steering 0V, MCP detector 2700V, Pusher cycle time (manual) 60, Ion energy 40V, Tube lens 0V, Grid 2 74V, TOF flight tube 4620V, Reflectron 1790V.

The cytochrome 2D6 gene containing a specific Single Nucleotide Polymorphism (T→deletion) was tested and successfully separated and identified. A partial sequence of the gene surrounding the SNP is shown below with the location of the polymorphism italicized and in bold face and the region that functions as a template in an amplification reaction using a pair of internal primers (described below) underlined:

agcagaggcg cttctccgtg tccaccttgc gcaacttggg cctgggcaag aagtcgctgg agcagtgggt gaccgaggag gccgcctgcc tttgtgccgc cttcgccaac cactccggtg ggtgatgggc (SEQ ID NO:9)

Two primer pairs are used for amplifying the region of cytochrome 2D6 gene containing the SNP. The external primers are designed to amplify only cytochrome 2D6 gene, not its pseudogenes. The internal primers are designed to have a partial Bsl I recognition sequence and to amplify a small region of cytochrome 2D6 gene containing the SNP. One of the internal primers also contains an N.BstNB I recognition sequence located 5′ to Bsl I recognition sequence. The sequences of these two primer pairs as well as that of the final amplification product are shown below with the bases that will form Bsl I recognition sequence highlighted in bold face and those that is N.BstNB I recognition sequence italicized. Some or all the bases in Bsl I and/or N.BstNB I recognition sequence(s) may be mismatched with the template sequence.

External primer forward:	5′-GAG ACC AGG GGG AGC ATA-3′ (SEQ ID NO:10)
External primer reverse:	5′-GGC GAT CAC GTT GCT CA-3′ (SEQ ID NO:11)
Internal forward primer:	5′-TGGGCCTGGG*AGTC*AAGTCGCTGG**C**CCAG-3′ (SEQ ID NO:12)
Internal reverse primer:	5′-GGC CTC CTC GGT **C**CC CC-3′ (SEQ ID NO:13)

Final amplification product: tgggcctggg*agtc*aagtcgctgg**cc**ag*t*ggggggaccgaggaggcc
(SEQ ID NO: 14)

The final amplification product is then digested by Bsl I and N.BstNB I, resulting the release of a single-stranded nucleic acid fragment CGCTGGCCCAG*T*G (SEQ ID NO:15). The mass of this sequence, including a 3′ OH and 5′ PO_4 is 4031.6 amu.

Figure 21 shows a UV chromatogram of the products of the PCR reaction after being digested with N.BstNB I and Bsl I. The large peak is the first two minutes contains PCR components. The arrow points to the single-stranded nucleic acid fragment created by Bsl I and N.BstNB I. The larger peak at the end of the run contains excess primer and double stranded product.

Figures 22A and 22B are single ion chromatogram and total ion chromatogram, respectively, of the products of the PCR reaction after being digested with both N.BstNB I and Bsl I. The single ion chromatogram plots the intensity of the ion of interest for each scan recorded, selectively pulling it from the noise.

Figure 23 shows the mass spectrum of a single-stranded oligonucleotide excised from the double-stranded PCR product with double digestion of N.BstNB I and Bsl I. The fragment mass is 1342 amu, exactly three times smaller than the calculated mass of 4031amu. This is due to the presence of multiple charge states on the fragment. It is common for biomolecules to be multiply charged upon ionization, generally the fragments generated by this assay will hold two, three or four charges. The masses expected can be calculated ahead of time using the simple formula M/Z-1, where M is the mass of the fragment, Z is the number of charges (1, 2, 3, *etc.*) subtracting one for the loss of a proton during the ionization resulting in a negatively charged molecule. In this case the only charge state is the triple charged form (this is not always the case, one, two, or several charge states can be seen for a single oligonucleotide at the same time) thus 4031/3 -1=1342 amu.

EXAMPLE 4

SEPARATION OF GENOTYPING FRAGMENTS BY LIQUID CHROMATOGRAPHY AND DETECTION WITH A UV DETECTOR AND TIME OF FLIGHT MASS SPECTROMETER USING ISOTHERMAL AMPLIFICATION

The following example describes the amplification of a specific sequence from the human genome using a primer pair each containing an N.BstNB I recognition sequence. The resulting amplicon, when cleaved with N.BstNB I, produced

two double stranded nucleic acid fragments each containing a 3′ recessed terminus. The 3′ recessed termini were then filled using 5′→3′ exonuclease deficient Bst DNA polymerase. The filled-in fragments were than nicked again by N.BstNB I and the resulting 3′ recessed termini were filed again by Bst polymerase. The process of nicking followed by filling the 3′ recessed termini were repeated so that two short single-stranded nucleic acid fragments were amplified. Such short fragments were then subjected to a chromatographic separation and identified by their mass to charge ratios.

To amplify a human genomic sequence containing an SNP using the above primer pair, a 50 µl PCR reaction mixture was prepared to contain 25 ng genomic DNA, 0.5 µM each forward and reverse primers, 10 mM Tris pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 µM each dNTP, 1 Unit DNA Polymerase (MasterAmp ™ Taq DNA Polymerase from Epicentre Technologies, Madison WI or Vent exo- Polymerase New England BioLabs, Beverly MA). Thermocycling was performed as follows: 95°C for 3 minutes initial denaturation; 30 cycles of 92°C for 40 seconds, 60°C for 30 seconds, 72°C for 30 seconds. A MJ Research PTC-100 thermocycler (MJ Research, Watertown, MA) was used for all PCR reactions. Primers were purchased from MWG Biotech (High Point, NC).

After the thermocycling was complete, an enzyme mixture was prepared containing N.BstNB I, Bst polymerase and 10x N.BstNB I buffer (New England BioLabs Beverly, MA). The mixture was added to PCR mixtures so that each reaction container had 30 units of N.BstNB I and 2 units of Bst polymerase in 150 mM KCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.5. The nicking reaction and the filing of the resulting 3′ recessed termini were carried out at 60°C for 30 minutes. The reaction was injected directly without any further manipulation. The injection was performed at 4°C from a cooled plated holder.

The chromatography system is an Agilent 1100 system composed of a binary pump, degasser, a column oven, a diode array detector, and thermostatted microwell plate autoinjector (Palo Alto CA). The column is a Varian Microsorb MV, incorporating C18 packing with 3 µM particle size, with 300 Angstrom pore size, 2.1 mm x 50 mm (Varian Inc. Walnut Creek, CA). The column was run at 30°C with a gradient of acetonitrile in 5 mM Triethylamine acetate (TEAA). Buffer A is 5 mM TEAA, buffer B is 5 mM TEAA and 25% (V/V) acetonitrile. The gradient begins with a hold at 10%B for one minute, then ramps to 50%B over 4 minutes followed by 30 seconds at 95%B, and finally returning to 10%B for a total run time of six minutes. The column temperature was held constant at 30°C. The flow rate was 0.416 ml per minute. The injection volume was 10 microliters. Flow into the mass spectrometer was

200ul/min, half the LC flow was diverted to waste using a tee. The mass spectrometer is a Micromass LCT Time-of-Flight with an electrospray inlet (Micromass Inc. Manchester UK). The samples were rune electrospray negative mode with a scan range from 700 to 2300 amu using an one second scan time. Instrument parameters were: TDC start voltage 700, TDC stop voltage 50, TDC threshold 0, TDC gain control 0, TDC edge control 0, Lteff 1117.5, Veff 4600. Source parameters were: desolvation gas 862 L/hr, capillary 3000V, sample cone 25V, RF lens 200V, extraction cone 2V, desolvation temperature 250C, source temperature 150C, RF DC offset 1 4V, FR DC offset 2 1V, aperture 6V, acceleration 200V, Focus, 10V, Steering 0V, MCP detector 2700V, Pusher cycle time (manual) 60, Ion energy 40V, Tube lens 0V, Grid 2 74V, TOF flight tube 4620V, Reflectron 1790V.

A cytochrome P450 2D6 gene containing a non-bi-allelic SNP (*i.e.*, G→C or A) was tested and a fragment containing this SNP were successfully separated and identified. The sequence of the primer pair used to amplify the fragment as well as PCR amplification product are shown below with the bases that constitutes N.BstNB I recognition sequence highlighted in bold face and the SNP indicated by "N" that can be either G, C or A.

Forward primer: 5′-ACC-CAG-CTG-GAT-GAG-**GAG-TC**A-ACT-GAG-CAC-3′ (SEQ ID NO:15)

Reverse primer: 5′-GGG-CTG-GCT-GG**G-AGT-C**AG-GTC-ATC-C -3′ (SEQ ID NO:16)

Amplification product:

5′-acc-cag-ctg-gat-gag-**gag-tc**a-act-gag-cac-**n**gg-atg-acc-t**ga-ctc**-cca-gcc-agc-cc-3′

3′-tgg-gtc-gac-cta-ctc-**ctc-ag**t-tga-ctc-gtg-***n***′cc-tac-tgg-act-gag-ggt-cgg-tcg-gg-5′

Digestion of N.BstNB I produces the following two double-stranded DNA fragment each having a 3′ recessed terminus:

5′- ACC-CAG-CTG-GAT-GAG-**GAG-TC**A-ACT-3′

3′- TGG-GTC-GAC-CTA-CTC-**CTC-AG**T-TGA-CTC-GTG-***N***′CC-TAC-5′

and

5′-GAG-CAC-***N***GG-ATG-ACC-TGA-CTC-CCA-GCC-AGC-CC-3′

3′-TGG-ACT-GAG-GGT-CGG-TCG-GG-5′

In the presence of Bst polymerase, two single-stranded nucleic acid fragments are amplified and then released:

BOTTOM STRAND: 5′-CAT-CC***N***-GTG-CTC-3′
TOP STRAND: 5′-GAG-CAC-***N***′GG-ATG-3′

The mass of the amplified single-stranded sequences containing various single nucleotide polymorphisms (SNPs), including their 3′ OH and 5′ PO_4 is shown in Table 4.

TABLE 4

MASS OF AMPLIFIED SINGLE-STRANDED SEQUENCES CONTAINING VARIOUS SNPS

M/Z for "A" allele	Top strand	Bottom strand
1	3799	3652
2	1899.5	1826
3	1266.3	1217
4	949	913
M/Z for "G" allele	Top strand	Bottom strand
1	3775.4	3677.4
2	1887.7	1838.7
3	1258.5	1225.8
4	943.9	919.4
M/Z for "C" allele	Top strand	Bottom strand
1	3815.5	3637.3
2	1907.8	1836.6
3	1271.8	1212.4
4	953.0	909.3

The results from the genotyping of two families (28 individuals) are shown in Table 5, which indicates that all the members tested are homozygous for a rare variant allele (*i.e.*, A at the SNP site).

TABLE 5

GENOTYPING DATA OF TWENTY EIGHT INDIVIDUALS

Individual	A allele top	bottom	G allele top	Bottom	C allele top	Bottom

Individual	A allele top	bottom	G allele top	Bottom	C allele top	Bottom
1	1265	1216	none	none	none	none
2	1265	1216	none	none	none	none
3	1265	1216	none	none	none	none
4	1265	1216	none	none	none	none
5	1265	1216	none	none	none	none
6	1265	1216	none	none	none	none
7	1265	1216	none	none	none	none
8	1265	1216	none	none	none	none
9	1265	1216	none	none	none	none
10	1265	1216	none	none	none	none
11	1265	1216	none	none	none	none
12	1265	1216	none	none	none	none
13	1265	1216	none	none	none	none
14	1265	1216	none	none	none	none
15	1265	1216	none	none	none	none
16	1265	1216	none	none	none	none
17	1265	1216	none	none	none	none
18	1265	1216	none	none	none	none
19	1265	1216	none	none	none	none
20	1265	1216	none	none	none	none
21	1265	1216	none	none	none	none
22	1265	1216	none	none	none	none
23	1265	1216	none	none	none	none
24	1265	1216	none	none	none	none
25	1265	1216	none	none	none	none
26	1265	1216	none	none	none	none
27	1265	1216	none	none	none	none
28	1265	1216	none	none	none	none

Figures 24 and 25 show exemplary UV chromatograms and extracted ion currents for the top and bottom stands of two individuals. The M/Z values of 1265 and 1216 are one Dalton lower than shown in Table 5 because the mass spectrometer was run in negative mode.

Figure 26 shows the UV chromatogram and the extracted ion currents for a no template control.

Figure 27 shows the mass spectra for the top fragment (m/z =3 at 1265 amu) and for the bottom fragment (m/z=3 at1216 amu).

EXAMPLE 5

MEASUREMENT OF A SINGLE NUCLEOTIDE POLYMORPHISM USING AN ISOTHERMAL AMPLIFICATION ASSAY AND AN LIQUID CHROMATOGRAPHY/MASS SPECTROMETRY ANALYSIS

This example describes the amplification of a genomic fragment containing a single nucleotide polymorphism (SNP) using a primer pair each containing a sequence of a sense strand of an N.BstNB I recogntion sequence. The resulting amplicon, when cleaved with N.BstNB I, produced two double-stranded nucleic acid fragments each containing a 3′ recessed terminus. The 3′ recessed termini were then filled using exo⁻ Vent polymerase. The filled-in fragments were then nicked again by N.BstNB I and the resulting 3′ recessed termini were filled again by exo⁻ Vent polymerase. The process of nicking followed by filling the 3′ recessed termini were repeated so that two short single-stranded nucleic acid fragments were amplified. Such short fragments were then subjected to a chromatographic separation and identified by their mass to charge ratios via a mass spectrometric analysis.

To amplify a genomic sequence containing an SNP, the following two primers were designed and purchased from MWG Biotech (High Point, NC) where the sequence of the sense strand of N.BstNB I recognition sequence is underlined:

Forward primer: 5′-tcccatggaaagagagtcaaag-3′

Reverse primer: 5′-tgaaataccgagcttgagtcaaaa-3′

0.4 μmol of the above forward and reverse primers were used in a PCR mixture assembled at 4°C. The total reaction volume of 50 microliters per well in a 96-well microtiter plate (Fisher, St Louis MO) also contains 400 micromolar dNTPs, 5 nanogram of genomic DNA (Coriell DNAs from Coriell Institute, Bethesda MD), 1x PCR buffer (EpiCentre Technologies, Austin TX), 1x enhancer solution (EpiCentre Technologies, Austin TX), and 1 unit Taq polymerase. The plate was sealed using a polypropylene plate sealer and then transferred to a hot (95°C) thermocycler (MWG, NC). The plate was then subjected to a thermocycling procedure known as a touchdown procedure starting at 70°C and ending at 60°C over 30 cycles. At the end of

30 cycles, the plate was taken to 4°C. The PCR process produced an oligonucleotide fragment that is 61 nucleotides in length.

The amplified oligonucleotide fragment was further used as a template for amplifying two single-stranded nucleic acid fragments in the presence of exo$^-$ Vent polymerase and N.BstNBI nicking endonuclease. The following mixture was combined and then 15 microliters of the mixture was added to each well in the microtiter plate.

250 ul 10x Thermopol buffer (NEB, Beverly, MA)
125 ul 10x N.BstNBI (NEB)
100 ul 25 mM dNTPs (NEB)
1000 ul 1 M trehalose (Sigma, St. Louis, MO)
250 units N.BstNBI nicking enzyme (NEB)
50 units exo- Vent DNA polymerase (NEB)
1020 ul ultrapure water

The plate was resealed at 4°C, and then the plate was heated to 60°C for 1 hour.

In this example, the SNP is biallelic and is an A to T substitution. In other words, for this SNP, the abundant allele is A whereas the rare allele is T. For the abundant allele (*i.e.* allele A), the two amplified single-stranded nucleic acid fragments are as follows with the SNP and its complementary nucleotide indicated by capital letters:

Top strand: 5′-ggagaaacacAtttg-3′
Bottom strand: 3′-cctctttgtgTaaac-5′

Likewise, for the rare allele (*i.e.*, allele T), the two amplified single-stranded nucleic acid fragments are as follows with the SNP and its complementary nucleotide indicated by capital letters:

Top strand: 5′-ggagaaacacTtttg-3′
Bottom strand: 3′-cctctttgtgAaaac-5′

The mass/charge calculations for the above four fragments when they are analyzed by mass spectrometry via electrospray in negative mode are as follows:

	Top strand allele A	Bottom strand allele A	Top strand allele T	Bottom strand allele T
Mass in Daltons	4705.10	4598.03	4696.09	4607.04
Mass/charge =2	2351.55	2298.02	2347.05	2302.52
Mass/charge =3	1567.37	1531.68	1564.36	1534.68
Mass/charge =4	1175.28	1148.51	1173.02	1150.76

Upon the completion of the single-stranded nucleic acid amplification reaction, the reaction mixture was loaded onto the LC/MS (Micromass LTD, Manchester UK and Beverly, MA, USA), where mass spectrometric analysis was conducted by a LCT time-of-flight using electrospray in the negative mode.

The chromatography system used is an Agilent HPLC-1100 composed of a binary pump, degasser, a column oven, a diode array detector, and thermostatted microwell plate autoinjector (Palo Alto, CA). The column is a Waters Xterra, incorporating C18 packing with 3 uM particle size, with 300 Angstrom pore size, 2.1 mm x 50 mm (Waters Inc. Milford, MA). The column was run at 30°C with a gradient of acetonitrile in 5 mM Triethylamine acetate (TEAA). Buffer A is 5 mM TEAA whereas buffer B is 5 mM TEAA and 25% (V/V) acetonitrile. The gradient begins with a hold at 20%B for three minutes then ramps to 50%B for 30 seconds followed by 1.5 minutes at 80%B for a total run time of six minutes. The column temperature was held constant at 30°C. The flow rate was 0.5 ml per minute. The injection volume was 10 microliters. Flow into the mass spectrometer was 200ul/min, half the LC flow was diverted to waste using a tee.

The mass spectrometer used is a Micromass LCT Time-of-Flight with an electrospray inlet (Micromass Inc. Manchester UK). The samples were run in electrospary negative mode with a scan range from 700 to 2000 amu using a one second scan time. Instrument parameters were: TDC start voltage 700, TDC stop voltage 200, TDC threshold 0, TDC gain control 0, TDC edge control 0, Lteff 1118.3, Veff 4600. Source parameters: Desolvation gas 862 L/hr, Capillary 3000V, Sample cone 30V, RF lens 250V, extraction cone 3V, desolvation temperature 325°C, Source temperature 150°C, RF DC offset 1 5V, FR DC offset 2 2V, Aperture 0V, accelaration 200V, Focus, 1V, Steering 0V, MCP detector 2700V, Pusher cycle time (manual) 60, Ion energy 40V, Tube lens 44V, Grid 2 50V, TOF flight tube 4620V, and Reflectron 1785V.

The sequence that was excised and analyzed was given for the abundant and the rare alleles along with their respective expected mass to charge ratios. Generally, mass to charge ratios larger than 14,000 were not recorded.

The following extracted ion currents were monitored:

1534.6 + 1150.7 (for the bottom strand of the T allele)
1564.3 + 1173.2 (for the top strand of the T allele)
1531.6 + 1148.5 (for the bottom strand of the A allele)
1567.3 + 1175.2 (for the top strand of the A allele)

The results for Individual No. 1 (Sample No. 1 from Coriell) are shown in Figure 32. The top panel shows the extracted ion current chromatogram for the bottom strand of the T allele. The second panel shows the extracted ion current chromatogram for the top strand of the T allele. The third panel shows the extracted ion current chromatogram for the bottom strand of the A allele. The fourth panel shows the extracted ion current chromatogram for the top strand of the A allele. The bottom panel shows the chromatogram for the total ion current. As indicated in Figure 32, the ion currents for the top or bottom strand of the T or A allele were as follows:

1534.6 + 1150.7 (for the bottom strand of the T allele):	0 units
1564.3 + 1173.2 (for the top strand of the T allele):	0 units
1531.6 + 1148.5 (for the bottom strand of the A allele):	17.23 units
1567.3 + 1175.2 (for the top strand of the A allele):	17.99 units

Accordingly, Individual No. 1 is homozygous for the A allele.

The results for Individual No. 2 (Sample No. 1 from Coriell) are shown in Figure 33. The top panel shows the extracted ion current chromatogram for the bottom strand of the T allele. The second panel shows the extracted ion current chromatogram for the top strand of the T allele. The third panel shows the extracted ion current chromatogram for the bottom strand of the A allele. The fourth panel shows the extracted ion current chromatogram for the top strand of the A allele. The bottom panel shows the chromatogram for the total ion current. As indicated in Figure 33, the ion currents for the top or bottom strand of the T or A allele were as follows:

1534.6 + 1150.7 (for the bottom strand of the T allele):	33.56 units
1564.3 + 1173.2 (for the top strand of the T allele):	35.98 units
1531.6 + 1148.5 (for the bottom strand of the A allele):	44.82 units
1567.3 + 1175.2 (for the top strand of the A allele):	55.95 units

Accordingly, Individual No. 2 is heterozygous for the T and A alleles.

EXAMPLE 6

MEASUREMENT OF ALLELIC FREQUENCIES USING AN ISOTHERMAL AMPLIFICATION ASSAY AND AN LIQUID CHROMATOGRAPHY/MASS SPECTROMETRY ANALYSIS

This example describes the amplification of a genomic fragment containing a single nucleotide polymorphism (SNP) in a complex, pooled sample using a primer pair each containing a sequence of a sense strand of an N.BstNB I recogntion sequence. Each of the resulting amplicons, when cleaved with N.BstNB I, produced two double-stranded nucleic acid fragments each containing a 3′ recessed terminus. The 3′ recessed termini were then filled using exo⁻ Vent polymerase. The filled-in fragments were then nicked again by N.BstNB I and the resulting 3′ recessed termini were filled again by exo⁻ Vent polymerase. The process of nicking followed by filling the 3′ recessed termini were repeated so that two short single-stranded nucleic acid fragments were amplified. Such short fragments were then subjected to a chromatographic separation and identified by their mass to charge ratios via a mass spectrometric analysis. The identification of these short fragments containing same or different alleles allows for the calculation of the frequency of a particular allele in the tested sample.

To amplify a genomic sequence containing an SNP, the following two primers were designed and purchased from MWG Biotech (High Point, NC) where the sequence of the sense strand of N.BstNB I recognition sequence is underlined:

Forward primer: 5′-tcccatggaaagagagtcaaag-3′
Reverse primer: 5′-tgaaataccgagcttgagtcaaaa-3′

In this example, the SNP is biallelic and is an A to T substitution. In other words, for this SNP, the abundant allele is A whereas the rare allele is T.

Five nanograms of genomic DNA from each of 200 different individuals obtained from the Coriell Institute (Bethesda, MD) were pooled together in one 1.5 ml centrifuge tube. All 200 individuals were homozygous for the A allele. To the pooled DNA sample, 5 ng of DNA from an individual that was previously determined to be homozygous for the T allele was added. Therefore, the pooled sample contained 200 parts of genomic DNA that was homozygous for the A allele and 1 part that was homozygous for the T allele. The sample was dessicated at -20°C in a rotovap and then rehydrated in 10 microliters of 0.01 M Tris-HCl, pH 7.6.

The following PCR mixture was assembled at 4°C. A total reaction volume of 50 microliters per well in a 96-well microtiter plate (Fisher, St Louis MO)

contained the following: 400 micromolar forward and reverse primers, 400 micromolar dNTPS, 1005 nanogram of genomic DNA (Coriell DNAs from Coriell Institute, Bethesda MD), 1x PCR buffer (EpiCentre Technologies, Austin TX), 1x enhancer solution (EpiCentre Technologies, Austin TX), 1 unit Taq polymerase. The plate was sealed using a polypropylene plate sealer. The plate reaction was prepared and sealed at 4°C and then transferred to a hot (95°C) thermocycler (MWG, NC). The plate was subjected to a thermocycling procedure known as a touchdown procedure starting at 70°C and ending at 60°C over 30 cycles.

Each of the amplified oligonucleotide fragment was further used as a template for amplifying two single-stranded nucleic acid fragments in the presence of exo$^-$ Vent polymerase and N.BstNB I nicking endonuclease. The following mixture was combined and then 15 microliters of the mixture was added to each well in the microtiter plate.

250 ul 10x Thermopol buffer (NEB, Beverly, MA)
125 ul 10x N.BstNBI (NEB)
100 ul 25 mM dNTPs (NEB)
1000 ul 1 M trehalose (Sigma, St. Louis MO)
250 units N.BstNBI nicking enzyme (NEB)
50 units exo- Vent DNA polymerase (NEB)
1020 ul ultrapure water

The plate was resealed at 4°C, and then the plate was heated to 60°C for 1 hour.

For the abundant allele (*i.e.* allele A), the two amplified single-stranded nucleic acid fragments are as follows with the SNP and its complementary nucleotide indicated by capital letters:

Top strand: 5′-ggagaaacacAtttg-3′
Bottom strand: 3′-cctctttgtgTaaac-5′

Likewise, for the rare allele (*i.e.*, allele T), the two amplified single-stranded nucleic acid fragments are as follows with the SNP and its complementary nucleotide indicated by capital letters:

Top strand: 5′-ggagaaacacTtttg-3′
Bottom strand: 3′-cctctttgtgAaaac-5′

The mass/charge calculations for the above four fragments when they are analyzed by mass spectrometry via electrospray in negative mode are as follows:

	Top strand allele A	Bottom strand allele A	Top strand allele T	Bottom strand allele T
Mass in Daltons	4705.10	4598.03	4696.09	4607.04
Mass/charge =2	2351.55	2298.02	2347.05	2302.52
Mass/charge =3	1567.37	1531.68	1564.36	1534.68
Mass/charge =4	1175.28	1148.51	1173.02	1150.76

Upon the completion of the single-stranded nucleic acid amplification reaction, the reaction mixture was loaded onto the LC/MS (Micromass LTD, Manchester UK and Beverly, MA, USA), where mass spectrometric analysis was conducted by a LCT time-of-flight using electrospray in the negative mode.

The chromatography system used is an Agilent HPLC-1100 composed of a binary pump, degasser, a column oven, a diode array detector, and thermostatted microwell plate autoinjector (Palo Alto, CA). The column is a Waters Xterra, incorporating C18 packing with 3 uM particle size, with 300 Angstrom pore size, 2.1 mm x 50 mm (Waters Inc. Milford, MA). The column was run at 30°C with a gradient of acetonitrile in 5 mM Triethylamine acetate (TEAA). Buffer A is 5 mM TEAA whereas buffer B is 5 mM TEAA and 25% (V/V) acetonitrile. The gradient begins with a hold at 20%B for three minutes then ramps to 50%B for 30 seconds followed by 1.5 minutes at 80%B for a total run time of six minutes. The column temperature was held constant at 30°C. The flow rate was 0.5 ml per minute. The injection volume was 10 microliters. Flow into the mass spectrometer was 200ul/min, half the LC flow was diverted to waste using a tee.

The mass spectrometer used is a Micromass LCT Time-of-Flight with an electrospray inlet (Micromass Inc. Manchester UK). The samples were run in electrospary negative mode with a scan range from 700 to 2000 amu using a one second scan time. Instrument parameters were: TDC start voltage 700, TDC stop voltage 200, TDC threshold 0, TDC gain control 0, TDC edge control 0, Lteff 1118.3, Veff 4600. Source parameters: Desolvation gas 862 L/hr, Capillary 3000V, Sample cone 30V, RF lens 250V, extraction cone 3V, desolvation temperature 325°C, Source temperature 150°C, RF DC offset 1 5V, FR DC offset 2 2V, Aperture 0V, accelaration 200V, Focus, 1V, Steering 0V, MCP detector 2700V, Pusher cycle time (manual) 60, Ion energy 40V, Tube lens 44V, Grid 2 50V, TOF flight tube 4620V, and Reflectron 1785V.

The sequence that was excised and analyzed was given for the abundant and the rare alleles along with their respective expected mass to charge ratios. Generally, mass to charge ratios larger than 14,000 were not recorded.

The following extracted ion currents were monitored:

1534.6 + 1150.7 (for the bottom strand of the T allele)
1564.3 + 1173.2 (for the top strand of the T allele)
1531.6 + 1148.5 (for the bottom strand of the A allele)
1567.3 + 1175.2 (for the top strand of the A allele)

The results for the pooled sample are shown in Figure 34. The top panel shows the extracted ion current chromatogram for the bottom strand of the T allele. The second panel shows the extracted ion current chromatogram for the top strand of the T allele. The third panel shows the extracted ion current chromatogram for the bottom strand of the A allele. The fourth panel shows the extracted ion current chromatogram for the top strand of the A allele. The bottom panel shows the chromatogram for the total ion current. As indicated in Figure 34, the ion currents for the top or bottom strand of the T or A allele were as follows:

1534.6 + 1150.7 (for the bottom strand of the T allele):	10.0 units
1564.3 + 1173.2 (for the top strand of the T allele):	3.6 units
1531.6 + 1148.5 (for the bottom strand of the A allele):	151.9 units
1567.3 + 1175.2 (for the top strand of the A allele):	166.3 units

These results indicate that 13.6 units of signal (*i.e.*, the sum of signal from the bottom and the top stands of the T allele) were observed for the T allele whereas 318.2 units of signal (*i.e.*, the sum of signal from the bottom and the top stands of the A allele) for the A allele. The frequency of the T allele in this DNA pool is thus 0.004 (*i.e.*, 13.6 divided by 318.2) or about 1 part in 200 parts.

Figure 35 shows the diode array trace (UV trace) for the amplified oligonucleotide fragments between 2.5 minutes and 5 minutes (the upper panel) and the total ion current (the bottom panel). The genotyping fragments have a collective retention time of 3.82 minutes in the diode array trace and 3.93 minutes in the total ion current chromatogram.

EXAMPLE 7

cDNA SYNTHESIS ON A SOLID SUPPORT AND PREPARATION OF SINGLE-STRANDED NUCLEIC ACID PROBE USING NICKING AGENTS AND DNA POLYMERASES

This example describes cDNA synthesis on a solid support and subsequent preparation of single-stranded nucleic acid probes using the synthesized cDNA as templates in the presence of a nicking endonuclease and a DNA polymerase. The design of this amplification scheme was to address the following problems in current cDNA-dependent technologies: high input of RNA required, low number of sample throughput, many user manipulations, organic extractions and precipitations, and poor adaptability.

Cell Stimulation and Total RNA Isolation

Jurkat line JRT 3.5 was stimulated for 6 hours at a cell density of $1x10^6$ cells/ml in serum-free RPMI medium (Life Technologies, Gaithersburg, MD) in the presence of 10 ng/ml phorbol-12-myristate-13 acetate (Calbiochem, San Diego, CA) and 100 ng/ml ionomycin (Calbiochem). Cells were pelleted, washed in 1xPBS (Life Technologies), re-pelleted. Each fraction containing 10^6 cells was lysed in 0.5ml buffer containing 4M guanidine isothiocyanate, 1% N-lauryl sarcosine, and 25mM sodium citrate pH 7.1 (Fisher Scientific. Pittsburgh, PA). One-tenth volume of 2M sodium acetate, pH 4.2 (Fisher Scientific) was added followed by one volume of water saturated phenol (Amresco. Solon, OH). After mixing, one-fourth volume chloroform:isoamyl alcohol (29:1) (Fisher Scientific) was added and the solution was mixed vigorously, then incubated on ice for 10 minutes. The lysate was then spun, the aqueous phase removed and extracted with an equal volume of chloroform:isoamyl alcohol. The aqueous phase was then pooled and the RNA precipitated with 2 volumes of ethanol (Quantum Chemical Corp. Tuscola, IL). After centrifugation, the ethanol was decanted and the RNA was air-dried briefly, then re-suspended in RNase-free water to a concentration of between 1 and 5 mg/ml.

Besides using total RNA prepared above by by acid-guanidinium-phenol extraction as a source for capturing mRNA, a simultaneous lysis-mRNA capture protocol using a chaotrope, such as GuSCN or guanidinium HCl, for both cell lysis and hybridization may also be employed without first isolating total RNA. This alternative approach permits the lysis of a small number of cells and avoids losses associated with organic phase extractions and serial ethanol precipitations inherent in standard RNA preparations.

Preparation of Oligo(dT)-Solid Tips (or Beads)

To capture mRNA, solid tips (or beads) to which the following oligonucleotide with an Asc I recognition sequence underlined that comprises oligo(dT) was prepared:

5'- ACTACTGATCAGGCGCGCCTTTTTTTTTTTTTTTTTT -3'
spacer Asc I $(poly\ dT)_{20}$

The above oligonucleotides were synthesized on a commercial synthesizer using standard cyanoethyl-N,N-diisorpropylamino-phosphoramidite (CED phosphoramidite) chemistry. Amine tails were incorporated onto the 5'-end using the commercially available N-monomethoxytritylaminohex-6-yloxy-CED-phosphoramidite. Alternatively, oligonucleotides may be commercially purchased (Midland Certified Reagents, Midland, TX).

Oligonucleotide tips (or beads) were prepared as previously described (Van Ness et al., *Nucl. Acids Res. 19*:3345, 1991). The oligonucleotide tips (or beads) contained 0.1 to 1.2 μg/tip (or bead) of covalently immobilized oligonucleotides.

Capturing Polyadenylated mRNA

Polyadenylated mRNA was captured on the above-prepared beads by first heating the mixture to 65-70°C, adding a high salt hybridization solution and then placing on a moving platform such as a rotary mixer. Sufficient mixing was accomplished by a variety of instruments including continual vortexer, orbital shaker, rotary rocker and hybridization oven equipped with a rotator.

Capture duration is dependent on the quantity of input RNA. At ambient temperature, 10 μg total RNA was enough to achieve 90% saturation of a bead in approximately 2 hours. In a typical resting cell, this would correspond to approximately 100 ng $poly(A)^+$ mRNA bound per bead. 40 μg total RNA from the same source gave the same level of saturation in 30 minutes. In another series of experiments, a stimulated human T-cell line was used as the RNA source and saturation was reached in 1 hour under the same capture conditions using a 10 μg total RNA input. Several RNA sources have been captured including mouse, hamster and human cell lines and all tend to fall within this range of 90% capture in 1 to 2 hours per 10 μg total RNA.

In one experiment, 10 μg total cellular RNA isolated from stimulated Jurkat line JRT 3.5 cells as described above was mixed with the beads bearing the covalently linked oligonucleotide that comprises poly(dT) and diluted in enough

RNase-free water to cover the bead, in a sterile 1.5 ml microfuge tube (Fisher Scientific). The RNA and bead were incubated at 65°C for 5 minutes. An equal volume of 2X mRNA hybridization buffer consisting of 50mM Tris pH 7.5, 1M NaCl (Fisher Scientific) and 20 µg/ml acetylated-BSA (New England Biolabs, Beverly, MA) was added to each tube and the tubes rocked gently for 2 hours at room temperature. The supernatant was removed and the bead was then washed three times in 1X mRNA hybridization buffer.

First Strand cDNA Synthesis

After capture of poly(A)$^+$ mRNA, the beads were washed in hybridization buffer three times to remove unbound RNA. Reverse transcription was done in a 30 µl volume using MMLV-reverse transcriptase (MMLV-RT) and an optimized buffer system for 1 to 2 hours at 42°C on a hybridization oven rotator rack. As with the capture step, efficient first stand synthesis requires constant mixing of the reagents. Initial experiments yielded 50-100 ng cDNA per solid support. The products of reverse transcription have been assayed indirectly by autoradiography of labeled double-stranded cDNA cleaved from the solid support and sized by agarose gel electrophoresis. In several experiments, the size span of copied mRNA species was comparable to, and often better than, conventional methods. The size distribution of the cDNA ranges from 0.5 Kb to 20 Kb, with the average being approximately 2.0 Kb.

In one experiment, the captured mRNA from stimulated Jurkat line JRT 3.5 cells prepared as described above was suspended in a reverse transcription mix consisting of 1X MMLV-reverse transcriptase buffer, 1mM dNTP mix, 2mM DTT (Life Technologies), 20 units Rnasin (Promega. Madison, WI) and 10ug/ml acetylated-BSA (New England Biolabs) followed by the addition of 600 units MMLV-reverse transcriptase (Life Technologies). This reaction was rocked gently at 42°C for 2 hours to synthesize the first strand of cDNA.

First strand cDNA synthesis is the first branch-point in solid support cDNA methodology. The remaining RNA template can be digested with RNaseH, hydrolyzed in NaOH or removed by heat denaturation, leaving a single-stranded cDNA template. The single-stranded cDNA template can then be used for oligonucleotide-directed second strand cDNA synthesis, PCR, random-primed probe production or gene-expression studies using labeled oligonucleotides.

Second Strand cDNA Synthesis

Upon the synthesis of a first strand cDNA, a second strand cDNA may be synthesized using an RNase H and a DNA polymerase. Alternatively, an adapter may be ligated onto the 3′ end of the first strand of cDNA and a complementary primer may be then hybridized to the adaptor. The second strand cDNA may subsequently be synthesized using the first strand cDNA as a template.

In one experiment, after reverse transcription, the bead was washed three times to remove reactants and enzyme. Second strand synthesis was done in a 40 µl volume using 1 unit RNase H per 25 units E.coli DNA polymerase I. The reaction was incubated at room temperature on a rotary rocker for 6 hours or overnight. Unevenly extended ends are "polished" for the subsequent ligation step by removing the second strand reaction and adding T4 DNA polymerase and dNTPs. This incubation proceeded for 30 minutes at 37°C on the rotator in a hybridization oven. Products were visualized directly by running the reaction in the presence of a ^{32}P-labelled dNTP, and either boiling the second strand products or by cleavage from the support with the restriction enzyme AscI. The radiolabelled products were run on a gel and placed on film for visialization. From a typical input of 10 µg total RNA, it was possible to recover 50-120 ng double-stranded cDNA. A thermostable DNA polymerase may also be used with an enzyme possessing RNaseH activity, such as MMLV-RT, after reverse transcription. Experiment comparing RNaseH/DNA pol I versus Tth I DNA polymerase in second strand synthesis demonstrated little difference in either the quantity of the AscI-cleaved product or the content of that product tested by PCR amplification of selected genes such as GAPDH and IL-2. The use of TthI DNA polymerase reduced the incubation time from 6 hours at room temperature to 1 hour at 70°C. TthI polymerase also showed reverse transcriptase activity in the presence of Mn^{2+}, but incubations needed to be kept very short due to the possibility of RNA hydrolysis at high temperature in the presence of divalent cations. DNA polymerization at high temperature also has the advantage of melting regions of secondary structure, which may be beneficial to the synthesis of cDNA from highly structured mRNAs.

In another experiment, to the reaction mixture for the first strand cDNA synthesis using mRNA from stimulated Jurkat line JRT 3.5 as templates (as described above), 1 unit RNase H (Boehringer-Mannheim, Indianapolis, IN) was then added. The reaction was allowed to continue for another 0.5 hour. The supernatant was removed and each bead is washed three times in a TE buffer containing 10mM Tris pH 8 and 1mM EDTA pH 8 (Fisher Scientific). Remaining RNA template was removed by boiling the beads in the TE buffer with 0.01% SDS (Fisher Scientific).

The synthesis of second strand cDNA represents a second branch-point in the solid support cDNA technology. Here, the choice can be made to ligate adapters to the cDNA that can support processes such as full-length single-stranded cDNA probes, library production, in vitro transcription and 5′ RACE.

Scaling Down

By minimizing loss of material during the many manipulations involved in standard cDNA library production, it should be possible to scale down the input RNA levels 10-100 fold. It has been shown that roughly 100 ng double-stranded cDNA can be synthesized on a solid support from 10 μg total RNA, which is 50-fold less starting material than called for in commercially available kits. It may be possible to reduce the above scale as much as 10-fold without significantly changing the protocol described above, but below the reduced scale, it will likely be necessary to include some sort of amplification step. One approach is to follow cDNA synthesis with an adapter ligation using an adapter that comprises an N.BstNB I recognition sequence. The resulting ligation product may be used as a template for synthesizing a single-stranded nucleic acid molecule that is identical to one strand of the double-stranded cDNA molecule, which in turn may be used as a cDNA probe or a template for synthesizing its complementary strand to thereby produce a double-stranded cDNA molecule.

Adapter Ligation

To facilitate the construction of a solid phase cDNA library or the production of cDNA probes, hemi-phosphorylated adapters were ligated to the double-stranded cDNA in a 30 μl volume with a 5-10:1 molar ratio of adapter:cDNA ends in a buffer containing 10%PEG. The adapter-cDNA mixture was incubated with T4 DNA ligase overnight at room temperature on a rotary rocker. The solid support was then washed 3 times to remove excess linkers. After ligation, the 5′-hydroxyl group on the adapter was phosphorylated with T4 polynucleotide kinase and ATP for 1 hour at 37°C on a hybridization oven rotating rack. This reaction was stopped by washing the bead three times in TE. If the cDNA is used for another application, the phosphorylation step will be unnecessary.

As discussed above, under certain circumstances, synthesized double-stranded cDNA was ligated to an adaptor that comprises an N.BstNB I recognition sequence. Such a ligation product may be used as a template to prepare single-strand cDNA probes or to further make double-stranded cDNA molecules.

cDNA Probe Manufacturing Using cDNA Synthesized on Solid Support as Templates

An N.BstNBI adapter has been synthesized and ligated to double-stranded cDNA on beads produced from 10 μg total RNA. The beads were placed in a standard Cetus PCR tube (ABI, Foster City, CA) and incubated at 60°C for 1 hour. The bound double-stranded cDNAs were continually nicked, producing many copies of one-strand of the double-stranded cDNA molecules. After incubation, the products were heat denatured and run on an agarose gel to visualize the range of lengths of the single-stranded cDNA. Sizes ranged from roughly 500 bp to larger than 20 Kb, which was in good agreement with the sizes of double-stranded cDNAs that were labelled, cleaved from beads, separated by electrophoresis and visualized by autoradiography in parallel with control experiments.

The reaction buffer for N.BstNB I (from NEB) was 10mM Tris-HCI, 10 mM $MgCl_2$, 150 mM KCl, 1 mM dithiothreitol (pH 7.5 at 25°C). The nicking reaction was carried out at 55°C-60°C. The nicking endonuclease may be heat inactivated by incubating at 80°C for 20 minutes.

The Polymerases tested in these experiments were exo$^-$ Vent, exo$^-$ Deep Vent and Bst (NEB). The reaction buffer for the polymerases was 1X ThermoPol Buffer containing 10 mM KCl, 20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, dNTPs (400 micromolar), and DNA. exo$^-$ Vent and exo$^-$ Deep Vent DNA Polymerases are maximally active at from 72°C to 80°C and supplied at a concentration of 2,000 units/ml.

To determine whether or not the amplified single-stranded cDNA was representative of the original mRNA population, PCR primers were designed for several genes known to be present at a high level, low level or not at all in a given type of cells. The design of the primers was such that all products would be approximately the same length (400-600 bp) and would be situated at or near the 5′ terminus of the cDNA so that a good approximation of the quality of the first strand synthesis may be obtained. Stimulated human Jurkat T-cell line was used as the RNA source. The following genes were analyzed: IL-2, IL-4, GM-CSF, GAPDH, c-fos, CTLA4 and the Werner's helicase. Mouse guanylate kinase was used as negative controls. All the PCR reactions using primers specific to the above genes gave products with expected sizes except those using primers specific to CTLA4 or mouse guanylate kinase. The product for IL-2 was confirmed by Northern blot where 50 ng of the putative IL-2 was ^{32}P-labelled and hybridized to a blot containing immobilized RNA from both stimulated (PMA and Ionomycin) Jurkat cells and unstimulated Jurkat cells. After a high stringency wash, the

unstimulated RNA showed almost no signal whereas the stimulated sample showed an intense signal consistent with the size expected for IL-2 message.

EXAMPLE 8

LINEAR AMPLIFICATION OF AN OLIGONUCLEOTIDE-

This example illustrates linear amplification of an oligonucleotide from a template duplex. The template duplex is formed by annealing two oligonucelotides to each other as shown below. The recognition sequence of N.BstNB I is shown below:

ITATOP: 5'-ccgatctagtgagtcgctc-3'
NbBT16: 3'-ggctagatcactcagcgagtcaaggtcagcatacc-5'

In the presence of a DNA polymerase, the recess of the above duplex is filled in to provide the following extension product:

5'-ccgatctagtgagtcgctcagttccagtcgtatgg-3'
3'-ggctagatcactcagcgagtcaaggtcagcatacc-5'

In the presence of N.BstNB I, the above extension product is nicked to produce the following nicked products:

5'-ccgatctagtgagtcgctc-3' + 5'-agttccagtcgtatgg-3'
3'-ggctagatcactcagcgagtcaaggtcagcatacc-5'

The above extension and nicking cycle may be repeated multiple times, resulting in the amplification of the fragment: 5'-agttccagtcgtatgg-3'. This fragment may be detected and characterized by liquid chromatography and mass spectrometry. It has a mass to charge ratio of 3 at 1663.1, a mass to charge ratio of 4 at 1247.1, and a mass to charge ratio of 5 at 997.1 daltons.

The following reaction was assembled at 4°C:

740 ul deionized nuclease free water
110 ul 10X N.BstNB I buffer (NEB)
55 ul 10X N.BstNB I buffer (NEB)
1ul of 1 picomole/ul of NBbt16 oligonucleotide
1ul of 1 picomole/ul of ITATOP oligonucleotide

80 ul of 2000 units /ml of N.BstNB I (NEB)
24 ul of 5000 units/ml 9°Nm™ DNA polymease (NEB)
16 ul 25 mM dNTPs (NEB)

The reaction mixture was divided into 20 50 ul aliquots in PCR tubes. The tubes were placed at 60°C on an MJ thermocycler and incubated for the indicated times. The samples were then subjected to the following liquid chromatography mass spectrometry analysis.

The column buffers are as follows: Buffer A contains 0.05 M dimethylbutylamine acetate, pH 7.6, while Buffer B contains 0.05 M dimethylbutylamine acetate, pH 7.6, 50% acetontrile.

A shallow gradient of acetonitrile is used to elute the oligonucleotides and clean up the sample. The analysis portion of the gradient starts at 5% acetonitrile and increases to 15% over about 90 seconds, followed by a wash that quickly pushes a "plug" of 45% acetonitrile onto the column for just a few seconds followed by a return to starting conditions of 5% acetonitrile.

The column used is Guard column Xterra 2. x 20 mm, 3.5 micron. MSC18. In front of the column is a frit in a frit holder (Upchurch A356 frit holder with Upchurch A701 Peek Prefilter Frit 0.5 micron).

The fractions from liquid chromatography were injected into mass spectrometer (Micromass LCT Time-of-Flight with an electrospray inlet, Micromass Inc., Manchester UK). The injection volume was 10 microliters. The samples were run electrospray negative mode with a scan range from 800 to 2000 amu. The time course results of the relative mass units at 1247.1 daltons are shown in the following table:

Time	Relative Mass Units
1	16
2	33
3	49
4	63
5	82
6	98
7	116
8	123
9	156
10	177

Time	Relative Mass Units
12.5	208
15	255
20	310
30	512
45	730
60	955
75	1233
90	1553

EXAMPLE 9

NUCLEIC ACID AMPLIFICATION USING TEMPLATE NUCLEIC ACID COMPRISING MISMATCHES IN NICKIGN AGENT RECOGNIZATION SEQUENCE

The following oligonucleotides were synthesized and obtained from MWG (MWG Biotech Inc., High Point, NC). The oligonucleotides were placed in 0.01 M Tris-HCl and 0.001 M EDTA at 100 pmoles per microliter. The sequence of the sense strand of the double-stranded recognition sequence of N.BstNB I is underlined whereas the nucleotide(s) that is different from the nucleotide at the corresponding position(s) of the antisense strand of the double-stranded recognition sequence of N.BstNB I is italicized

B-1: 5' CC TAC GAC TGG AAC AGA CTC ACC TAC GAC TGG A- 3'
B-2: 5' CC TAC GAC TGG AAC A*AT* *AAA* ACC TAC GAC TGG A- 3'
B-3: 5' CC TAC GAC TGG AAC AGA *TTC* ACC TAC GAC TGG A- 3'
B-4: 5' CC TAC GAC TGG AAC AGA *CAC* ACC TAC GAC TGG A- 3'
B-5: 5' CC TAC GAC TGG AAC AG*T* CTC ACC TAC GAC TGG A- 3'
B-6: 5' CC TAC GAC TGG AAC AGA *AAC* ACC TAC GAC TGG A- 3'
B-7: 5' CC TAC GAC TGG AAC AG*T* *AAC* ACC TAC GAC TGG A- 3'

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
T-1a: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
T-1b: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC - 5'
T-1c: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG AC- 5'
T-1d: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG A- 5'
T-1e: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG - 5'
T-1f: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CT- 5'

```
T-1g: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG C- 5'
T-1h: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG- 5'
T-1i: 3' GG ATG CTG ACC TTG TCT GAG TGG AT- 5'
T-1j: 3' GG ATG CTG ACC TTG TCT GAG TGG A- 5'
T-1k: 3' GG ATG CTG ACC TTG TCT GAG TGG - 5'
T-1l: 3' GG ATG CTG ACC TTG TCT GAG TG- 5'
T-1m: 3' GG ATG CTG ACC TTG TCT GAG T- 5'
T-1n: 3' GG ATG CTG ACC TTG TCT GAG - 5'
```

The following mixture was combined and then 25 microliters of the mixture was added to each well in the microtiter plate.

250 ul 10x Thermopol buffer (NEB Biolabs, Beverly, MA)
125 ul 10x N.BstNBI (NEB Biolabs, Beverly, MA)
100 ul 25 mM dNTPs (NEB Biolabs, Beverly, MA)
1000 ul 1 M trehalose (Sigma, St. Louis, MO)
250 units N.BstNBI nicking enzyme (NEB Biolabs, Beverly, MA)
50 units Vent exo- DNA polymerase (NEB Biolabs, Beverly, MA)
1020 ul ultra pure water

25 microliters of each respective duplex was then added to the microtiter plate. The duplex was formed by first diluting two oligonucleotide primers and placing them in the following solution at a final concentration of 1 pmole per microliter: 1x Thermopol buffer (New England Biolabs, Beverly, MA) and 0.5x N.BstNBI buffer. The 1x Thermopol buffer consists of 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl pH8.8, 0.1% Triton X-100, 2 mM $MgSO_4$, whereas the 1x N.BstNBI buffer consists of 150 mM KCl, 10 mMTris-HCl, 10 mM $MgCl_2$, 1 mM DTT. The mixture was then heated to 100°C for 1 minute and then held at 50°C for 10 minutes to allow the duplexes to form. The plate was resealed at 4°C, and then heated to 60°C for 1 hour.

The following duplexes were tested:

#1 (perfect base pairing)

```
T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-1: 5' CC TAC GAC TGG AAC AGA CTC ACC TAC GAC TGG A- 3'
```

#2 (complete mismatching)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#3 (single mismatch)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-3: 5' CC TAC GAC TGG AAC AGA TTC ACC TAC GAC TGG A- 3'

#4 (single mismatch)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-4: 5' CC TAC GAC TGG AAC AGA CAC ACC TAC GAC TGG A- 3'

#5 (single mismatch)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-5: 5' CC TAC GAC TGG AAC AGT CTC ACC TAC GAC TGG A- 3'

#6 (2 mismatches)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-6: 5' CC TAC GAC TGG AAC AGA AAC ACC TAC GAC TGG A- 3'

#7 (3 mismatches)

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8a.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8b.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC - 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8c.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG AC- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8d.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG A- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8e.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG - 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8f.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CT- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8g.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG C- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8h.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG - 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8i.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG AT- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8j.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG A- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8k.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG - 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8l.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TG- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8m.

T-1: 3' GG ATG CTG ACC TTG TCT GAG T- 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#8n.

T-1: 3' GG ATG CTG ACC TTG TCT GAG - 5'
B-7: 5' CC TAC GAC TGG AAC AGT AAC ACC TAC GAC TGG A- 3'

#9a.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9b.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC - 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9c.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG AC- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9d.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG A- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9e.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG - 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9f.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CT- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9g.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG C- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9h.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG - 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9i.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG AT- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9j.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG A- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9k.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG - 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9l.

T-1: 3' GG ATG CTG ACC TTG TCT GAG TG- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9m.

T-1: 3' GG ATG CTG ACC TTG TCT GAG T- 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

#9n.

T-1: 3' GG ATG CTG ACC TTG TCT GAG - 5'
B-2: 5' CC TAC GAC TGG AAC AAT AAA ACC TAC GAC TGG A- 3'

The plate was loaded onto the LC/MS (Micromass LTD, Manchester UK and Beverly, MA, USA) that is a LCT time-of-flight uisng electrospray in the negative mode. The conditions were as follows:

The chromatography system was an Agilent HPLC-1100 composed of a binary pump, degasser, a column oven, a diode array detector, and thermostatted microwell plate autoinjector (Palo Alto, CA). The column was a Waters Xterra,

incorporating C18 packing with 3 uM particle size, with 300 Angstrom pore size, 2.1 mm x 50 mm (Waters Inc. Milford, MA). The column was run at 30C with a gradient of acetonitrile in 5 mM Triethylamine acetate (TEAA). Buffer A was 5 mM TEAA, buffer B was 5 mM TEAA and 25% (V/V) acetonitrile. The gradient began with a hold at 10%B for one minute then ramped to 50%B over 4 minutes followed by 30 seconds at 95%B and finally returned to 10%B for a total run time of six minutes. The column temperature was held constant at 30C. The flow rate was 0.416 ml per minute. The injection volume was 10 microliters. Flow into the mass spectrometer was 200ul/min, half the LC flow was diverted to waste using a tee. The mass spectrometer wass a Micromass LCT Time-of-Flight with an electrospray inlet (Micromass Inc. Manchester UK). The samples were run in electrospary negative mode with a scan range from 700 to 2300 amu using a 1 second scan time. Instrument parameters were: TDC start voltage 700, TDC stop voltage 50, TDC threshold 0, TDC gain control 0, TDC edge control 0, Lteff 1117.5, Veff 4600. Source parameters: Desolvation gas 862 L/hr, Capillary 3000V, Sample cone 25V, RF lens 200V, extraction cone 2V, desolvation temperature 250C, Source temperature 150C, RF DC offset 1 4V, FR DC offset 2 1V, Aperture 6V, accelaration 200V, Focus, 10V, Steering 0V, MCP detector 2700V, Pusher cycle time (manual) 60, Ion energy 40V, Tube lens 0V, Grid 2 74V, TOF flight tube 4620V, Reflectron 1790V.

The following extracted ion currents were monitored: 1144.7 daltons plus or minus 1 dalton around 1144.7 for the following fragment to be released:

```
3' GG ATG CTG ACC-5'
```

from the following duplex, as well as the other duplexes listed above:

```
T-1: 3' GG ATG CTG ACC TTG TCT GAG TGG ATG CTG ACC T- 5'
B-1: 5' CC TAC GAC TGG AAC AGA CTC ACC TAC GAC TGG A- 3'
```

The results are shown in the table below:

Duplex Names	Number of Mismatches Within Double-Stranded N.BstNB1 Recognition Sequence	Relative Mass Units Observed
1	0	121.0
2	5	18.5
3	1	66.7

Duplex Names	Number of Mismatches Within Double-Stranded N.BstNB1 Recognition Sequence	Relative Mass Units Observed
4	1	61.5
5	1	63.0
6	2	45.0
7	3	21.2
8a	3	23.4
8b	3	28.3
8c	3	11.5
8d	3	29.2
8e	3	14.6
8f	3	17.8
8g	3	20.8
8h	3	12.3
8i	3	14.9
8j	3	18.3
8k	3	19.3
8l	3	15.6
8m	3	18.3
8n	3	12.5
9a	5	21.3
9b	5	17.8
9c	5	19.2
9d	5	15.3
9e	5	14.0
9f	5	15.9
9g	5	28.3
9h	5	22.7
9i	5	23.9
9j	5	21.4
9k	5	22.6
9l	5	22.5
9m	5	13.5
9n	5	14.3

All of the above U.S. patents, U.S. patent application publications, U.S. patent applications, foreign patents, foreign patent applications and non-patent

publications referred to in this specification and/or listed in the Application Data Sheet, are incorporated herein by reference, in their entirety.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed is:

1. A method, comprising

(a) contacting a partially or fully double-stranded template nucleic acid molecule with a nicking agent (NA), the template comprising a nicking agent recognition sequence (NARS) recognizable by the NA;

(b) if the template does not comprise a nicking site (NS) nickable by the NA, extending the template to provide the NS;

(c) nicking the template or the extension product thereof at the NS;

(d) extending the nicked product of step (c) from the 3' terminus at the NS; and

(e) repeating steps (c) and (d) to thereby amplify a single-stranded nucleic acid fragment.

2. The method of claim 1 wherein the single-stranded nucleic acid fragment of (e) has no more than 25 nucleotides.

3. The method of claim 1 wherein step (c) is performed in the presence of a DNA polymerase selected from the group consisting of exo$^-$ Vent, exo$^-$ Deep Vent, exo$^-$ Bst, exo$^-$ Pfu, exo$^-$ Bca, Taq, the Klenow fragment of DNA polymerase I, T5 DNA polymerase, Phi29 DNA polymerase, phage M2 DNA polymerase, phage phiPRD1 DNA polymerase, Sequenase, PRD1 DNA polymearse, 9°Nm^{TM} DNA polymerase, and T4 DNA polymerase.

4. The method of claim 1 wherein steps (b), (c), (d) and (e) are performed under an identical isothermal condition.

5. The method of claim 1 wherein the NA is a nicking endonuclease (NE).

6. The method of claim 1 wherein the template nucleic acid molecule is immobilized to a solid support.

7. The method of claims 1-6 wherein the template nucleic acid further comprises a genetic variation located 3' to, and on the same strand of the molecule as the NS, such that the genetic variation is incorporated into the amplified single-stranded nucleic acid fragment.

8. The method of claims 1-7 wherein the template nucleic acid is formed by a method comprising

(i) forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP, and a target nucleic acid comprising a genetic variation, wherein

(a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then

the first ODNP comprises a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation, and optionally comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS),

or

(b) if the target nucleic acid is a single-stranded nucleic acid, then

the first ODNP comprises a sequence of a sense sequence of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially identical to the target nucleic acid located 5' to the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation, and optionally comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS);

(ii) extending the first and the second ODNPs to produce a fragment having both the NERS and the RERS; and

(iii) optionally cleaving the extension product of step (ii) with a restriction endonuclease that recognizes the RERS.

9. The method of claims 1-7 wherein the template nucleic acid is formed by a method comprising:

(i) forming a mixture of a first ODNP, a second ODNP, and a target nucleic acid comprising a genetic variation, wherein

(a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then

the first ODNP comprises a nucleotide sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation,

or

(b) if the target nucleic acid is a single-stranded nucleic acid, then

the first ODNP comprises a sequence of one strand of a first restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the target nucleic acid located 5' to the genetic variation, and

the second ODNP comprises a sequence of one strand of a second RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation;

(ii) extending the first and the second ODNPs in the presence of deoxyribonucleoside triphosphates and at least one modified deoxyribonucleoside triphosphate to produce a fragment having both the first and the second RERSs.

10. The method of claims 1-7 wherein the template nucleic acid is formed by a method comprising

(i) forming a mixture of a first ODNP, a second ODNP and a target nucleic acid comprising a genetic variation, wherein

(a) if the target nucleic acid is a double-stranded nucleic acid molecule having a first strand and a second strand, then

the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation,

or

(b) if the target nucleic acid is a single-stranded nucleic acid, then

the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation;

the first and the second ODNPs each further comprise a sequence of a sense strand of a NERS;

(ii) extending the first and the second ODNPs to produce a fragment having two NERSs.

11. The method of claim 1 wherein the template nucleic acid molecule further comprises either a 5' overhang in the strand that contains the NS or a 3' overhang in the strand that does not contain the NS, wherein an overhang comprises a nucleic acid sequence at least substantially complementary to a target nucleic acid.

12. The method of claim 11 further comprising:

(i) admixing the template nucleic acid with nucleic acid molecules in a biological sample, where the biological sample may contain the target nucleic acid, under conditions where the target nucleic acid, if present in the biological sample, would hybridize to the overhang of the template nucleic acid; and

(ii) removing unhybridized template nucleic acid from the admixture of step (i) before performing step (a); and

(iii) combining the hybridized template nucleic acid with a NA.

13. The method of claim 1 wherein the double-stranded template nucleic acid molecule further comprises a type IIs restriction endonuclease recognition sequence (TRERS) and is provided by linking a nucleic acid adaptor to a double-stranded target nucleic acid so that the amplified single-stranded nucleic acid molecule comprises a portion of the target nucleic acid, the nucleic acid adaptor comprising a type IIs restriction endonuclease recognition sequence (TRERS) and the NARS where in the stand that does not contain the

NS, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the double-stranded target nucleic acid fragment and 5' to the position corresponding to the NS.

14. The method of claims 1-7 wherein the template nucleic acid further comprises a junction between an upstream exon (Exon A) and a downstream exon (Exon B) located 3' to the NS such that the nucleotides adjunct to the junction at both sides of the junction are incorporated into the amplified single-stranded nucleic acid fragment.

15. The method of claim 14 wherein the template nucleic acid is formed by a method comprising

(i) admixing a first oligonucleotide primer (ODNP), a second ODNP, and a cDNA, wherein

the first ODNP comprises a sequence at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand,

the second ODNP comprises a sequence at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand, and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking agent recognition sequence (NARS); and

(ii) extending the first ODNP and the second ODNP to provide the template nucleic acid that comprises the first ODNP and the second ODNP.

16. A method for amplifying a single-stranded nucleic acid molecule, comprising:

(A) forming a mixture comprising:

(i) a target nucleic acid;

(ii) an oligonucleotide primer that

(a) comprises a sequence of the sense strand of a double-stranded nicking agent recognition sequence recognizable by a nicking agent that nicks outside the recognition sequence, and

(b) is at least substantially complementary to the target nucleic acid; and

(B) amplifying a single-stranded nucleic acid molecule using a portion of the target nucleic acid as a template in the presence of the nicking agent.

17. The method of claim 16 wherein the double-stranded nicking agent recognition sequence is recognizable by N.BstNB I.

18. The method of claim 16 wherein a nucleotide in the sequence of the sense strand of the double-stranded nicking agent recognition sequence does not form a conventional base pair with another nucleotide of the target nucleic acid when the oligonucleotide primer anneals to the target nucleic acid.

19. The method of claim 16 wherein two or more nucleotides in the sequence of the sense strand of the double-stranded nicking agent recognition sequence do not form conventional base pairs with nucleotides of the target nucleic acid when the oligonucleotide primer anneals to the target nucleic acid.

20. A method comprising:

(A) forming a mixture comprising:

(i) a target nucleic acid;

(ii) an oligonucleotide primer that

(a) comprises a sequence of the sense strand of a NARS recognizable by a nicking agent (NA) that nicks outside the NARS, and

(b) is at least substantially complementary to a first region of the target nucleic acid; and

(iii) a partially double-stranded nucleic acid that

(a) comprises a double-stranded type IIs restriction endonucelase recognition sequence, and

(b) a 3' overhang that is at least substantially complementary to a second region of the target nucleic acid located 5' to the first region;

under conditions that allow for hybridization between the oligonucleotide primer and the first region of the target and between the 3' overhang of the partially double-stranded nucleic acid and the second region of the target;

(B) digesting the target nucleic acid hybridized to the oligonucleotide primer and the partially double-stranded nucleic acid in the second region;

(C) amplifying a single-stranded nucleic acid molecule using a portion of the target nucleic acid digested in step (B) as a template in the presence of the nicking agent.

21. A method for identifying a genetic variation at a defined position in a target nucleic acid, comprising

(a) providing a partially or fully double-stranded template nucleic acid that comprises a portion of the target nucleic acid that includes the genetic variation at the defined position, and further comprises a nicking site (NS) located at 5' to the genetic variation;

(b) amplifying, in the presence of a DNA polymerase and a nicking agent (NA) that nicks at the NS, a single-stranded nucleic acid fragment that comprises the genetic variation, and whose 5' terminus was at the NS; and

(c) characterizing the single-stranded nucleic acid fragment to thereby identify the genetic variation.

22. The method of claim 21 wherein the genetic variation is a single nucleotide polymorphism.

23. The method of claim 21 wherein the single-stranded nucleic acid fragment contains no more than 25 nucleotides.

24. The method of claim 21 wherein the target nucleic acid is genomic nucleic acid or cDNA.

25. The method of claim 21 wherein the NA is a nicking endonuclease (NE).

26. The method of claim 21 wherein step (b) is performed under an isothermal condition.

27. The method of 21 wherein the amplified single-stranded nucleic acid fragment contains no more than 17 nucleotides.

28. The method of claim 21 wherein the characterizing of step (c) is performed at least partially by the use of a technique selected from the group consisting of mass spectrometry, liquid chromatography, fluorescence polarization, and electrophoresis.

29. A method for identifying a genetic variation at a defined location in a target nucleic acid according to claim 21, comprising

(a) forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP and the target nucleic acid, wherein

(i) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand, then

the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation,

or

(ii) if the target nucleic acid is a single-stranded nucleic acid, then

the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and

the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation,

the first and the second ODNPs each further comprising a nucleotide sequence of a sense strand of a nicking endonuclease recognition sequence (NERS);

(b) extending the first and the second ODNPs to produce an extension product comprising two NERSs;

(c) amplifying a single-stranded nucleic acid fragment using one strand of the extension product of step (b) as a template in the presence of one or more nicking endonucleases (NEs) that recognizes the NERS(s); and

(d) characterizing the single-stranded fragment of step (c) to thereby identify the genetic variation.

30. A method for determining the ratio of the number of a first nucleic acid molecule to the number of a second nucleic acid molecule in a nucleic acid population, wherein the sequences of the first and the second nucleic acid molecules are identical except at a defined location, comprising:

(a) admixing a first oligonucleotide primer (ODNP), a second ODNP and the nucleic acid population, wherein

(i) if the first and the second nucleic acids are double stranded, each having a first strand and a second strand,

a) the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the first strand of the first nucleic acid located 3' to the defined location, and

(b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the second strand of the first nucleic acid located 3' to the defined location,

or,

(ii) if the first and the second nucleic acids are single stranded,

a) the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleic acid sequence of the first nucleic acid located 5' to the defined location, and

b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the first nucleic acid located 3' to the defined location; and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking agent recognition sequence (NARS);

(b) extending the first and the second ODNPs to produce fragments encompassed by the first and the second ODNPs;

(c) amplifying single-stranded nucleic acid fragments in the presence of a DNA polymease and a nicking agent (NA) that recognizes the NARS; and

(d) determining the ratio of the number of the single-stranded nucleic acid fragments of step (c) derived from the first nucleic acid to that derived from the second nucleic acid.

31. The method of claim 30 wherein the first ODNP or the second ODNP comprise a sequence of one strand of a RERS, but do not comprise a sequence of a sense strand of a NERS; and fragments produced by step (b) do not contain a hemimodified RERS.

32. The method of claim 30 wherein the amplified fragments of step (b) are digested by a RE that recognizes the RERS before performing step (c).

33. The method of any claim 30 wherein the amplification product of step (c) is a single-stranded nucleic acid fragment containing no more than 35 nucleotides.

34. A method for determining the presence, or the absence, of a target nucleic acid in a sample, comprising

(a) admixing the nucleic acid molecules of the sample with a partially double-stranded oligonucleotide probe to allow for hybridization of the probe to the target nucleic acid, the probe comprising

(i) a nicking agent recognition sequence (NARS) cleavable by a nicking agent (NA) to produce a nicking site (NS), and

(ii) a 5' overhang in the strand that either the strand or an extension product thereof contains the NS, or

a 3' overhang in the strand that neither the strand nor an extension product thereof contains the NS,

wherein the overhang comprises a nucleic acid sequence at least substantially complementary to a target nucleic acid,

(b) separating unhybridized probe from the unhybridized probe;

(c) performing an amplification reaction in the presence of the NA, a DNA polymerase, and the hybridized probe, to provide an amplification product; and

(d) detecting the presence, or the absence, of the amplification product of step (c) to thereby determine the presence, or the absence, of the target nucleic acid in the sample.

35. The method of claim 34 wherein the target nucleic acid is a nucleic acid molecule originating from a member selected from the group consisting of a bacterium, a virus, a fungus and a parasite.

36. The method of claim 34 wherein the NA is N.BstNB I.

37. The method of claim 34 wherein step (c) is carried out in an isothermal condition.

38. The method of claim 34 wherein step (c) is carried out in the presence of a labeled deoxynucleoside triphosphate, and the labeled deoxynucleoside triphosphate is a deoxynucleoside triphosphate linked to a label selected from the group consisting of a radiolabel, an enzyme, a fluorescent dye, digoxigenin, and biotin.

39. The method of claim 34 wherein the amplification product, if present, contains no more than 35 nucleotides.

40. The method of claim 34 wherein the nucleic acid molecules of the sample are immobilized to a solid support.

41. The method of claim 34 wherein the amplification product of step (c) is immobilized to a solid substrate before being detected in step (d).

42. A method for determining the presence or the absence of a target nucleic acid in a sample, comprising

(A) forming a mixture comprising:

(i) the nucleic acid molecules of the sample;

(ii) a single-stranded nucleic acid probe comprising from 3' to 5':

(a) a first sequence that is at least substantially complementary to the target nucleic acid,

(b) a sequence of the antisense strand of a nicking agent recognition sequence (NARS), and

(c) a second sequence;

(iii) a nicking endonuclease (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s);

(B) maintaining the mixture at conditions that amplify a single-stranded nucleic acid fragment using the single-stranded nucleic acid probe as a template if the target nucleic acid is present in the sample, wherein the single-stranded nucleic acid fragment is at most 25 nucleotides in length; and

(C) detecting the presence or the absence of the single-stranded nucleic acid fragment amplified in step (B) to determine the presence, or the absence, of the target nucleic acid in the sample.

43. A method for determining the presence or the absence of a target nucleic acid in a sample, comprising

(A) form a mixture comprising:

(i) the nucleic acid molecules of the sample;

(ii) a single-stranded nucleic acid probe comprising from 3' to 5':

(a) a sequence that is at least substantially complementary to the target nucleic acid, and

(b) a sequence of the sense strand of a nicking agent recognition sequence (NARS),

(iii) a nicking endonuclease (NA) that recognizes the first NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s);

(B) maintaining the mixture at conditions that amplify a single-stranded nucleic acid fragment that is at least substantially complementary to the target nucleic acid if the target nucleic acid is present in the sample, wherein the single-stranded nucleic acid fragment is at most 25 nucleotides in length; and

(C) detecting the presence or the absence of the single-stranded nucleic acid fragment amplified in step (B) to determine the presence, or the absence, of the target nucleic acid in the sample.

44. A method for determining the presence or absence of a target nucleic acid that comprises a nicking agent recognition sequence (NARS) in a sample, the method comprising:

(A) forming a mixture comprising:

(i) the nucleic acid molecules of the sample,

(ii) a nicking agent (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s);

(B) maintaining the mixture at conditions that amplify a single-stranded nucleic acid fragment using a portion of the target nucleic acid as a template if the target nucleic acid is present in the sample; and

(C) detecting the presence or absence of the single-stranded nucleic acid fragment amplified in step (B) to determine the presence or absence of the target nucleic acid in the sample.

45. A method for determining the presence or absence of a target nucleic acid that comprises a nicking agent recognition sequence (NARS) in a sample, the method comprising:

(A) forming a mixture comprising:

(i) the nucleic acid molecules of the sample,

(ii) a single-stranded nucleic acid probe that is substantially identical to one strand of the target nucleic acid and comprise a sequence of the antisense strand of the NARS,

(iii) a nicking agent (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxynucleoside triphosphate(s);

(B) maintaining the mixture at conditions that amplify a single-stranded nucleic acid fragment using a portion of the probe as a template if the target nucleic acid is present in the sample; and

(C) detecting the presence or absence of the single-stranded nucleic acid fragment of step (B) to determine the presence or absence of the target nucleic acid in the sample.

46. The method of claims 42-45 wherein the single-stranded nucleic acid fragment of step (B) is at most 25 nucleotides in length.

47. The method of claims 42-46 wherein the nucleic acid molecules of the sample are immobilized.

48. The method of claims 42-46 wherein the single-stranded nucleic acid probe is immobilized to a solid support.

49. The method of either of claims 1 or 34 wherein the NARS contains a mismatched base pair.

50. A method for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, comprising:

(a) admixing a first oligonucleotide primer (ODNP), a second ODNP, and the cDNA, wherein

the first ODNP comprises a sequence at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand,

the second ODNP comprises a sequence at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand, and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking agent recognition sequence (NARS);

(b) performing an extension reaction under a condition that produces a fragment encompassed by the first ODNP and the second ODNP if both Exon A and Exon B are present in the cDNA;

(c) performing an amplification reaction in the presence of a DNA polymerase and a nicking agent (NA) that recognizes the NARS; and

(d) characterizing an amplification product of step (c), if the amplification product is present, to thereby determine the presence, or the absence, of the junction between Exon A and Exon B.

51. The method of claim 50 wherein both the first ODNP and the second ODNP comprise a sequence of a sense strand of a NARS.

52. The method of claim 50 wherein the NERS is recognizable by N.BstNB I.

53. The method of claim 50 wherein the first ODNP or the second ODNP does not contain a sequence of a sense strand of a NERS, but comprises a sequence of one strand of a RERS; and the fragment produced by step (b) does not contain a hemimodified RERS.

54. The method of claim 50 wherein step (c) is performed under an isothermal condition.

55. The method of claim 50 wherein the amplification product, if present, is a single-stranded nucleic acid fragment containing no more than 35 nucleotides.

56. An oligonucleotide primer (ODNP) pair for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, comprising

(a) the first ODNP comprises a sequence at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand, and

the second ODNP comprises a sequence at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand; and

wherein at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking agent recognition sequence (NARS).

57. The ODNP pair of claim 56 wherein the NARS is a nicking endonuclease recognition sequence (NERS).

58. The ODNP pair of claim 56 wherein both the first ODNP and the second ODNP comprise a sequence of a sense strand of a NARS.

59. The ODNP pair of claim 58 wherein the NARS is a NERS.

60. The ODNP pair of claim 56 wherein the first ODNP comprises a sequence of a sense strand of a NERS, and the second ODNP comprises a sequence of one strand of a RERS.

61. The ODNP pair of claim 56 wherein the first ODNP comprises a sequence of one strand of a RERS, and the second ODNP comprises a sequence of a sense strand of a NERS.

62. The ODNP pair of claim 56 wherein the second ODNP is at least 12 nucleotides in length.

63. A method of making a template nucleic acid molecule for synthesizing a single-stranded cDNA molecule, comprising

a. providing a double-stranded cDNA molecule; and

b. ligating the double-stranded cDNA molecule with a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS) so that the single-stranded nucleic acid molecule amplified in the presence of a nicking agent (NA) that recognizes the NARS and a DNA polymerase using the ligation product as a template comprises at least a portion of one strand of the double-stranded cDNA molecule.

64. The method of claim 63 wherein said portion is at least 8 nucleotides in length.

65. The method of claim 63 wherein said portion is at least 12 nucleotides in length.

66. The method of claim 63 wherein either the double-stranded cDNA molecule or the nucleic acid adaptor is immobilized via the terminus that is not involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor.

67. The method of claim 63 wherein the NA is a nicking endonuclease (NE).

68. A method for making a cDNA library, comprising:

a. providing mRNA molecules isolated from a biological sample;

b. preparing double-stranded cDNA molecules using the mRNA as templates; and

c. ligating the double-stranded cDNA molecules with a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS) so that the single-stranded nucleic acid molecules amplified in the presence of a nicking agent that recognizes the NARS and a DNA polymerase using the ligation products as templates comprise at least a portion of one strand of the double-stranded cDNA molecules.

69. The method of claim 68 wherein said portion is at least 8 nucleotides in length.

70. A method for identifying a single nucleotide polymorphism (SNP) at a defined position in a target nucleic acid, comprising

(a) providing a partially or fully double-stranded template nucleic acid that comprises a portion of the target nucleic acid that includes the SNP at the defined position, and further comprises a nicking site (NS) located at 5' to the genetic variation;

(b) amplifying, in the presence of a DNA polymerase and a nicking endonuclease (NE) that nicks at the NS and under an isothermal condition, a single-stranded nucleic acid fragment that comprises the genetic variation, contains no more than 17 nucleotides, and whose 5' terminus was at the NS; and

(c) characterizing the single-stranded nucleic acid fragment using at least partially liquid chromatography and/or mass spectrometry to thereby identify the SNP.

71. The method of claim 70 wherein the NE is N.BstNB I.

72. The method of claim 70 wherein step (b) is performed at 50°C-70°C.

73. A method for determining the presence, or the absence, of a target nucleic acid in a biological sample from a mammalian subject, comprising

(a) admixing the nucleic acid molecules of the sample with a partially double-stranded oligonucleotide probe to allow for hybridization of the probe to the target nucleic acid, the probe comprising-

(i) a nicking endonuclease recognition sequence (NERS) cleavable by a nicking endonuclease (NE) to produce a nicking site (NS), and

(ii) a 5' overhang in the strand that either the strand or an extension product thereof contains the NS, or

a 3' overhang in the strand that neither the strand nor an extension product thereof contains the NS,

wherein the overhang comprises a nucleic acid sequence at least substantially complementary to a target nucleic acid,

(b) removing unhybridized probe from the admixture of step (a);

(c) performing an amplification reaction under an isothermal condition in the presence of the NE and a DNA polymerase to provide an amplification product that contain no more than 17 nucleotides; and

(d) detecting the presence, or the absence, of the amplification product of step (c) using at least partially liquid chromatography and/or mass spectrometry to thereby determine the presence, or the absence, of the target nucleic acid in the sample.

74. The method of claim 73 wherein the amplification product of step (c) contains no more than 12 nucleotides.

75. A method for the multiplex identification of genetic variations at defined locations in target nucleic acids each having first and second strands, comprising,

(a) for each target nucleic acid,

(i) forming a mixture of a first oligonucleotide primer (ODNP), a second ODNP, and the target nucleic acid, wherein

a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand,

the first ODNP comprises a sequence of a sense strand of nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

the second ODNP comprises an optional sequence of one strand of a restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation,

or

b) if the target nucleic acid is a single-stranded nucleic acid,

the first ODNP comprises a sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and

the second ODNP comprises an optional sequence of one strand of a RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation;

(ii) extending the first and the second ODNPs to produce a fragment encompassed by the first ODNP and the second ODNP, wherein the extension reaction may be performed either individually for each target nucleic acid or jointly for more than one target nucleic acids;

(b) optionally cleaving the extension products with a RE that recognizes the RERS;

(c) amplifying single-stranded nucleic acid fragments using one strand of each extension product of step (b) or each digestion product of step (c) as template in the presence of a nicking endonuclease that recognizes the NERS; and

(d) characterizing the single-stranded fragments to thereby identify the genetic variations.

76. The method of claim 75 wherein the nucleotide sequences of the first ODNPs at least substantially complementary to first strands of the target nucleic acids if the target nucleic acids are double-stranded, or identical to the target nucleic acids if the target nucleic acids are single-stranded, are at least 12 nucleotides in length.

77. The method of claim 75 wherein the nucleotide sequences of the second ODNPs at least substantially complementary to second strands of target nucleic acids if the target nucleic acids are double-stranded, or to the target nucleic acids if the target nucleic acids are double-stranded, are at least 12 nucleotides in length.

78. The method of claim 75 wherein the RERS is recognizable by an interrupted restriction endonuclease.

79. The method of claim 75 wherein the NERS is 5'GAGTC3'.

80. The method of claim 75 wherein step (c) is performed under an isothermal condition.

81. The method of claim 75 wherein the single-stranded nucleic acid fragments of step (c) contains no more than 17 nucleotides.

82. The method of claim 75 further comprising combining products of at least two extension reactions of step (a)(ii) before performing step (c).

83. The method of claim 75 wherein at least one of the genetic variations is a single nucleotide polymorphism.

84. A kit for identifying a genetic variation at a defined location in a target nucleic acid, comprising:

(a) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand,

(i) a first oligonucleotide primer (ODNP) comprises a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located at 3' to the complement of the genetic variation, and

(ii) a second ODNP comprises an optional sequence of one strand of a restriction endonuclease recognition sequence (RERS) and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation;

or,

(b) if the target nucleic acid is a single-stranded nucleic acid,

(i) the first ODNP comprises a sequence of a sense strand of a NERS and a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and

(ii) the second ODNP comprises an optional sequence of one strand of a RERS and a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation.

85. The kit of claim 84 where in the double-stranded nucleic acid molecule amplified using the first and the second ODNPs as primers and the target nucleic acid as a template, the distance between a nicking site (NS) produced by a nicking endonuclease (NE) that recognizes the NERS in one strand and the location corresponding to a cleavage site produced by a restriction endonuclease (RE) that recognizes the RERS in the other strand is no more than 35 nucleotides.

86. The kit of claim 84 wherein the distance between the NS in one strand and the location corresponding to the cleavage site of the RE in the other strand is no more than 25 nucleotides.

87. A partially double-stranded oligonucleotide probe comprising

(a) a nicking agent recognition sequence (NARS) that is cleavable by a nicking agent (NA) to produce a nicking site (NS);

(b) either a 5' overhang in the strand that contains the NS or a 3' overhang in the strand that does not contain the NS, wherein each overhang comprises a nucleic acid sequence at least substantially complementary to a target nucleic acid; and

(c) a sequence within the strand that neither the strand nor the extension product thereof contains the NS, the sequence located at 5' to the position corresponding to the NS, the sequence uniquely correlating to the target nucleic acid to which the overhang is at least substantially complementary.

88. The oligonucleotide probe of claim 87 wherein the target nucleic acid is one strand of a denatured double-stranded nucleic acid.

89. The oligonucleotide probe of claim 87 wherein the double-stranded nucleic acid is genomic nucleic acid or cDNA.

90. The oligonucleotide probe of claim 87 wherein the double-stranded region is 12-20 nucleotides in length.

91. The oligonucleotide probe of claim 87 wherein the 5' overhang or the 3' overhang is 8-20 nucleotides in length.

92. An oligonucleotide primer (ODNP) pair comprising a first ODNP and a second ODNP, wherein

(a) the first ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of a single-stranded target nucleic acid located 3' to a defined location;

(b) the second ODNP comprises a nucleotide sequence that is at least substantially complementary to a nucleotide sequence of the complement of the target nucleic acid at a location 3' to the at least substantially complementary nucleotide of the nucleotide at the defined position;

(c) the first and the second ODNPs further comprise a first constant recognition sequence (CRS) of a first strand and a second CRS of a second strand of an interrupted restriction endonuclease recognition sequence (IRERS), respectively, but not a complete IRERS, the complete IRERS being a double-stranded nucleic acid having the first and the second strands and comprising the first and the second CRS linked by a variable recognition sequence (VRS), and

(d) either the first or the second ODNP further comprises a NERS located at 5' to the first CRS or the second CRS.

93. A method for determining the ratio of the number of a first nucleic acid molecule to the number of a second nucleic acid molecule in a nucleic acid population, wherein the sequences of the first and the second nucleic acid molecules are identical except at a defined location, comprising:

(a) admixing a first oligonucleotide primer (ODNP), a second ODNP and the nucleic acid population, wherein

(i) if the first and the second nucleic acids are double-stranded, each having a first strand and a second strand,

a) the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the first strand of the first nucleic acid located 3' to the defined location, and

(b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the second strand of the first nucleic acid located 3' to the defined location,

or,

(ii) if the first and the second nucleic acids are single-stranded,

a) the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleic acid sequence of the first nucleic acid located 5' to the defined location, and

b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the first nucleic acid located 3' to the defined location; and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS);

(b) extending the first and the second ODNPs to produce fragments encompassed by the first and the second ODNPs;

(c) amplifying single-stranded nucleic acid fragments under an isothermal condition in the presence of a DNA polymease and a nicking endonuclease (NE) that recognizes the NERS, wherein the single-stranded nucleic acid fragments each contain no more than 17 nucleotides; and

(d) determining the ratio of the number of the single-stranded nucleic acid fragments of step (c) derived from the first nucleic acid to that derived from the second nucleic acid using at least partially liquid chromatography and/or mass spectrometry.

94. A method for determining the allelic frequency of a target nucleic acid molecule with a gene variation at a defined location in a nucleic acid population, comprising:

(a) admixing a first oligonucleotide primer (ODNP), a second ODNP and the nucleic acid population, wherein

(i) if the target nucleic acid is double-stranded, having a first strand and a second strand,

a) the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the first strand of the target nucleic acid located 3' to the defined location, and

(b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the second strand of the target nucleic acid located 3' to the defined location,

or,

(ii) if the target nucleic acids are single-stranded,

a) the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleic acid sequence of the target nucleic acid located 5' to the defined location, and

b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleic acid sequence of the target nucleic acid located 3' to the defined location; and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS);

(b) extending the first and the second ODNPs to produce a fragment encompassed by the first and the second ODNPs;

(c) amplifying single-stranded nucleic acid fragment(s) under an isothermal condition in the presence of a DNA polymease and a nicking endonuclease (NE) that recognizes the NERS, wherein the single-stranded nucleic acid fragment(s) each contain no more than 17 nucleotides; and

(d) determining the percentage of the single-stranded nucleic acid fragment(s) of step (c) that contains the genetic variation of the target nucleic acid (Product A) in all the single-stranded nucleic acid fragment(s) of step (c) that are identical to the portions of Product A other than the genetic variation using at least partially liquid chromatography and mass spectrometry to thereby determine the allelic frequency of the target nucleic acid.

95. A nucleic acid adaptor comprising

(a) an optional type IIs restriction endonuclease recognition sequence (TRERS); and

(b) a nicking agent recognition sequence (NARS),

wherein when the adaptor is linked to a double-stranded target nucleic acid fragment so that, in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, the target nucleic acid fragment is located 5' to the NARS, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS in said strand, if the TRERS is present, is located both

(i) within the double-stranded target nucleic acid fragment, and

(ii) 5' to the position corresponding to the NS.

96. The nucleic acid adaptor of claim 95 wherein the NA is a nicking endonuclease (NE).

97. The nucleic acid adaptor of claim 95 where in the strand containing the NS, the TRERS is located 5' to the NARS.

98. The nucleic acid adaptor of claim 95 where in the strand containing the NS, the TRERS is located 3' to the NARS.

99. The nucleic acid adaptor of claim 95 wherein the TRERS is recognizable by Bpm I or Mme I.

100. The nucleic acid adaptor of claim 95 wherein the adaptor has a 3' overhang for ligating with the target nucleic acid fragment.

101. The nucleic acid adaptor of claim 95 wherein the adaptor has a 5' overhang for ligating with the target nucleic acid fragment.

102. The nucleic acid adaptor of claim 95 wherein the adaptor has a blunt end for ligating with the target nucleic acid fragment.

103. A nucleic acid adaptor comprising

(a) a type IIs restriction endonuclease recognition sequence (TRERS); and

(b) a nicking agent recognition sequence (NARS),

wherein when the adaptor is linked to a double-stranded target nucleic acid fragment so that, in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS, the target nucleic acid fragment is located 5' to the NARS, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS in said strand is located both

(i) within the double-stranded target nucleic acid fragment, and

(ii) 5' to the position corresponding to the NS.

104. A template nucleic acid molecule for synthesizing a single-stranded cDNA molecule, comprising

(a) a double-stranded cDNA molecule or a fragment thereof; and

(b) a nucleic acid adaptor that comprises a nicking agent recognition sequence (NARS) ligated to the double-stranded cDNA molecule or the fragment thereof so that the single-stranded cDNA molecule amplified in the presence of a nicking agent (NA) that recognizes the NARS and a DNA polymerase using the ligation product as a template comprises at least a portion of one strand of the double-stranded cDNA or the fragment thereof, said portion being at least 8 nucleotides in length.

105. The template nucleic acid molecule of claim 104 wherein at least one strand of the template nucleic acid molecule is immobilized to a solid support via a terminus of either the double-stranded cDNA molecule or the nucleic acid adaptor that is not involved in the ligation between the double-stranded cDNA molecule and the nucleic acid adaptor.

106. The template nucleic acid molecule of claim 104 wherein a linker molecule is present between the solid phase and the terminus of the template nucleic acid molecule.

107. The template nucleic acid molecule of claim 106 wherein the linker comprises a sequence of one strand of a restriction endonuclease recognition sequence (RERS).

108. A method, comprising

(a) providing a double-stranded nucleic acid comprising

(i) an optional type IIs restriction endonuclease recognition sequence (TRERS),

(ii) a nicking agent recognition sequence (NARS), and

(iii) a target nucleic acid fragment,

wherein in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS,

(1) the target nucleic acid fragment is located 5' to the NARS, and

(2) if the TRERS is present, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS;

(b) optionally cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS; and

(c) amplifying a single-stranded nucleic acid fragment in the presence of a NA that recognizes the NARS.

109. A method, comprising

(a) providing a double-stranded nucleic acid comprising

(i) a type IIs restriction endonuclease recognition sequence (TRERS),

(ii) a nicking endonuclease recognition sequence (NERS), and

(iii) a target nucleic acid fragment,

wherein in the strand that does not contain the nicking site (NS) of a nicking endonuclease (NE) that recognizes the NERS,

(1) the target nucleic acid fragment is located 5' to the NERS, and

(2) the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS;

(b) cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS; and

(c) amplifying a single-stranded nucleic acid fragment under an isothermal condition in the presence of a NA that recognizes the NARS, wherein the single-stranded nucleic acid fragment contains no more than 25 nucleotides.

110. A method for making a plurality of single-stranded nucleic acid probes, comprising

(a) providing double-stranded nucleic acids each comprising

(i) an optional type IIs restriction endonuclease recognition sequence (TRERS),

(ii) a nicking agent recognition sequence (NARS), and

(iii) a target nucleic acid fragment,

wherein in the strand that does not contain the nicking site (NS) of a nicking agent (NA) that recognizes the NARS,

(1) the target nucleic acid fragment is located 5' to the NARS, and

(2) if the TRERS is present, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS;

(b) optionally cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS; and

(c) amplifying single-stranded nucleic acid fragments in the presence of a NA that recognizes the NARS.

111. A method for making a plurality of single-stranded nucleic acid probes, comprising

(a) providing double-stranded nucleic acids each comprising

(i) a type IIs restriction endonuclease recognition sequence (TRERS),

(ii) a nicking endonuclease recognition sequence (NERS), and

(iii) a target nucleic acid fragment,

wherein in the strand that does not contain the nicking site (NS) of a nicking endonuclease (NE) that recognizes the NERS,

(1) the target nucleic acid fragment is located 5' to the NERS, and

(2) if the TRERS is present, the cleavage site of a type IIs restriction endonuclease that recognizes the TRERS is located both within the target nucleic acid fragment and 5' to the position corresponding to the NS;

(b) cleaving the double-stranded nucleic acid with a type IIs restriction endonuclease that recognizes the TRERS; and

(c) amplifying single-stranded nucleic acid fragments under an isothermal condition in the presence of a NE that recognizes the NERS, wherein the single-stranded nucleic acid fragments each contain no more than 25 nucleotides.

112. A method for detecting the presence, or the absence, of a junction between an upstream exon (Exon A) and a downstream exon (Exon B) of a gene in a cDNA molecule, comprising:

(a) admixing a first oligonucleotide primer (ODNP), a second ODNP, and the cDNA, wherein

the first ODNP comprises a sequence at least substantially complementary to a portion of the antisense strand of Exon A near the 5' terminus of Exon A in the antisense strand,

the second ODNP comprises a sequence at least substantially complementary to a portion of the sense strand of Exon B near the 5' terminus of Exon B in the sense strand, and

at least one of the first ODNP and the second ODNP further comprises a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS);

(b) performing an extension reaction under a condition that produces a fragment encompassed by the first ODNP and the second ODNP if both Exon A and Exon B are present in the cDNA;

(c) performing an amplification reaction under an isothermal condition in the presence of a DNA polymerase and a nicking endonuclease (NE) that recognizes the NERS; and

(d) characterizing an amplification product of step (c), if the amplification product is present, using at least partially liquid chromatography and/or mass spectrometry to thereby determine the presence, or the absence, of the junction between Exon A and Exon B.

113. A method for detecting, in a cDNA molecule or a cDNA population, the presence, or the absence, of each potential junction between any two exons of a gene that comprises n exons, wherein n is an integer equal or more than 2, Exon A is the most upstream exon, and Exon N is the most downstream exon; comprising:

(a) admixing a first set of oligonucleotide primers (ODNPs), a second set of ODNPs, and the cDNA molecule or the cDNA population, wherein

the first set of ODNPs comprise n-1 ODNPs corresponding respectively to one of the exons except Exon A, each of the ODNPs comprising a sequence at least substantially complementary to a portion of the sense strand of the corresponding exon near the 5' terminus of said exon in said strand,

the second set of ODNPs comprise n-1 ODNPs corresponding respectively to one of the exons except Exon N, each of the ODNPs comprising a sequence at least substantially complementary to a portion of the antisense strand of the corresponding exon near the 5' terminus of said exon in said strand, and

each of the ODNPs of at least one ODNP set further comprises a sequence of a sense strand of a nicking agent recognition sequence (NARS);

(b) amplifying fragments encompassed by one ODNP from the first ODNP set and another ODNP from the second ODNP set;

(c) performing an amplification reaction in the presence of a DNA polymerase and a nicking agent (NA) that recognizes the NARS; and

(d) detecting the presence, or the absence, of amplification product(s) of step (c) to thereby determine the presence, or the absence, of each potential junction.

114. A method for detecting, in a cDNA molecule or a cDNA population, the presence, or the absence, of a junction between each two exons of a gene that comprises n exons, wherein n is an integer equal or more than 2, Exon A is the most upstream exon, and Exon N is the most downstream exon; comprising:

(a) admixing a first set of oligonucleotide primers (ODNPs), a second set of ODNPs, and the cDNA molecule or the cDNA population, wherein

the first set of ODNPs comprise n-1 ODNPs corresponding respectively to one of the exons except Exon A, each of the ODNPs comprising a sequence at

least substantially complementary to a portion of the sense strand of the corresponding exon near the 5' terminus of said exon in said strand,

the second set of ODNPs comprise n-1 ODNPs corresponding respectively to one of the exons except Exon N, each of the ODNPs comprising a sequence at least substantially complementary to a portion of the antisense strand of the corresponding exon near the 5' terminus of said exon in said strand, and

each of the ODNPs of at least one ODNP set further comprises a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS);

(b) amplifying fragments encompassed by one ODNP from the first ODNP set and another ODNP from the second ODNP set;

(c) performing an amplification reaction under an isothermal condition in the presence of a DNA polymerase and a nicking endonuclease (NE) that recognizes the NERS; and

(d) detecting the presence, or the absence, of amplification product(s) of step (c) using at least partially liquid chromatography and/or mass spectrometry to thereby determine the presence, or the absence, of each potential junction.

115. A method for detecting alternative splicing of a gene in a cDNA population, comprising

(a) determining the presence, or the absence, of each potential junction between any two exons of the gene according to the method according to claims 114; and

(b) indicating the presence of alternative splicing of the gene in the cDNA population, if more than one junctions are present for at least one exon of the gene at at least one terminus of the exon.

116. A kit for identifying a genetic variation at a defined location in a target nucleic acid having first and second strands, comprising:

(A) a first oligonucleotide primer (ODNP) and a second oligonucleotide primer (ODNP) wherein

(1) if the target nucleic acid is a double-stranded nucleic acid having a first strand and a second strand,

(a) the first ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the first strand of the target nucleic acid located 3' to the complement of the genetic variation, and

(b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the second strand of the target nucleic acid located 3' to the genetic variation;

or

if the target nucleic acid is a single-stranded nucleic acid,

(a) the first ODNP comprises a nucleotide sequence at least substantially identical to a nucleotide sequence of the target nucleic acid located 5' to the genetic variation, and

(b) the second ODNP comprises a nucleotide sequence at least substantially complementary to a nucleotide sequence of the target nucleic acid located 3' to the genetic variation;

(2) wherein the first and the second ODNPs each further comprise a sequence of a sense strand of a nicking endonuclease recognition sequence (NERS); and

(3) wherein the distance between the a nicking site (NS) produced by a nicking endonuclease that recognizes the NERS in one strand and the location corresponding to a NS produced by the NE in the other strand is no more than 25 nucleotides;

(B) a nicking endonuclease that recognizes the nicking endonuclease recognition sequence;

(C) a DNA polymerase;

(D) a liquid chromatography column;

(E) Buffer A that comprises water and an ammonium salt of a secondary or tertiary amine complexed with an organic or inorganic acid;

(F) Buffer B that comprises Buffer A and orgainic solvent; and

(G) a strand displacement facilitator.

117. A method, comprising

(a) forming a mixture of a partially or fully double-stranded template nucleic acid molecule comprising a nicking agent recognition sequence (NARS), a nicking agent (NA) that recognizes the NARS, a DNA polymerase, and one or more deoxyribonucleoside triphosphates;

(b) maintaining the mixture of step (a) under conditions that allows for the amplification of a single-stranded nucleic acid fragment, wherein the single-stranded nucleic acid fragment is capable of dissociating from the template nucleic acid in the absence of any strand displacement activity of the DNA polymerase or a strand displacement facilitator.

118. A method, comprising

(a) contacting a partially or fully double-stranded template nucleic acid with a nicking agent (NA), the template comprising a nicking agent recognition sequence (NARS);

(b) if the template does not comprise a nicking site (NS) nickable by the NA, extending the template to provide the NS;

(c) nicking the template or the extension product thereof at the NS to provide a nicked product that comprises a new 3' terminus at the NS and a nicked product that comprises a new 5' terminus at the NS; wherein the melting temperature of the duplex formed between the nicked product comprising the new 5' terminus at the NS and the strand of the template or the extension product thereof that comprises the sequence of the antisense strand of the NARS is higher than that of the duplex formed between the nicked product comprising the new 3' terminus at the NS and the same strand of the template or the extension product thereof;

(d) extending the nicked product of step (c) that comprises the new 3' terminus at the NS;

(e) repeating steps (c) and (d) to amplify a single-stranded nucleic acid fragment.

119. A method according to any of claims 63, 73, 117, or 188, wherein the NARS includes at least one mismatched base pair.

120. The probe of claim 87 wherein the NARS includes at least one mismatched base pair.