PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, G01N 33/543, 33/544, 33/548, 33/566

A1

(11) International Publication Number: WO 95/12689

(43) International Publication Date: 11 May 1995 (11.05.95)

(21) International Application Number: PCT/US94/12768

(22) International Filing Date: 4 November 1994 (04.11.94)

(30) Priority Data:
08/147,785 4 November 1993 (04.11.93) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/147,785 (CIP)
Filed on 4 November 1993 (04.11.93)

(71) Applicant (*for all designated States except US*): GENE CHECK, INC. [US/US]; Route 1, Box 236, Vermontville, NY 02989 (US).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): WAGNER, Robert, E., Jr. [US/US]; Route 1, Box 236, Vermontville, NY 02898 (US).

(74) Agents: LIVNAT, Shmuel et al.; Morrison & Foerster, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006 (US).

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

Published
With international search report.

(54) Title: IMMOBILIZED MISMATCH BINDING PROTEIN FOR DETECTION OR PURIFICATION OF MUTATIONS OR POLYMORPHISMS

(57) Abstract

A method for detecting mutations, such as a single base change or an addition or deletion of about one to four base pairs, is based on the use of an immobilized DNA mismatch-binding protein, such as MutS, which binds to a nucleic acid hybrid having a single base mismatch or unpaired base or bases, thereby allowing the detection of mutations involving as little as one base change in a nucleotide sequence. Such a method is useful for diagnosing a variety of important disease states or susceptibilities, detecting the presence of a mutated oncogene and for isolating or removing by affinity chromatography duplex DNA molecules containing mismatches such as error-containing molecules in PCR-amplified DNA samples. Also provided are kits useful for practicing the methods of the present invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

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

IMMOBILIZED MISMATCH BINDING PROTEIN FOR DETECTION OR PURIFICATION OF MUTATIONS OR POLYMORPHISMS

5

BACKGROUND OF THE INVENTION

Field of the Invention

The invention in the fields of molecular biology and medicine relates to a method for detecting mutations involving as little as one base change or a single base addition to, or deletion from, the wild-type DNA sequence, as well as methods for removing mismatch-containing DNA from batches of amplified DNA.

10

Description of the Background Art

Progress in human molecular and medical genetics depends on the efficient and accurate detection of mutations and sequence polymorphisms, the vast majority of which results from single base substitutions and small additions or deletions. Assays capable of detecting the presence of a particular mutation or mutant nucleic acid sequence in a sample are therefore of substantial importance in the prediction and diagnosis of disease, forensic medicine, epidemiology and public health. Such assays can be used, for example, to detect the presence of a mutant gene in an individual, allowing determination of the probability that the individual will suffer from a genetic disease. The ability to detect a mutation has taken on increasing importance in early detection of cancer or discovery of susceptibility to cancer with the discovery that discrete mutations in cellular oncogenes can result in activation of that oncogene leading to the transformation of that cell into a cancer cell (Nishimura, S. *et al.*,

15

20

25

30

Biochem. J. 243:313-327 (1987); Bos, J.L., *Cancer Res. 49*:4682-4689 (1989)).

The desire to increase the utility and applicability of such assays is often frustrated by assay sensitivity as well as complexity and cost. Hence, it would be highly desirable to develop more sensitive as well as simple and relatively inexpensive assays for detection of alterations in DNA.

Nucleic acid detection assays can be based on any of a number of characteristics of a nucleic acid molecule, such as its size, sequence, susceptibility to digestion by restriction endonucleases, *etc.* The sensitivity of such assays may be increased by altering the manner in which detection is reported or signaled to the observer. Thus, for example, assay sensitivity can be increased through the use of detectably labeled reagents, wherein the labels may be enzymes (Kourilsky *et al.*, U.S. Patent 4,581,333), radioisotopes (Falkow *et al.*, U.S. Patent 4,358,535; Berninger, U.S. Patent 4,446,237), fluorescent labels (Albarella *et al.*, EP 144914), chemical labels (Sheldon III *et al.*, U.S. Patent 4,582,789; Albarella *et al.*, U.S. Patent 4,563,417), modified bases (Miyoshi *et al.*, EP 119448), and the like.

Most methods devised to attempt to detect genetic alterations consisting of one or a few bases involve hybridization between a standard nucleic acid (DNA or RNA) and a test DNA such that the mutation is revealed as a mispaired or unpaired base in a heteroduplex molecule. Detection of these mispaired or unpaired bases has been accomplished by a variety of methods. Mismatches have been detected by means of enzymes (RNaseA, MutY) which cut one or both strands of the duplex at the site of a mismatch (Myers, R.M. *et al.*, *Cold Spring Harbor Symp. Quant. Biol. 51*:275-284 (1986); Gibbs, R. *et al.*, *Science 236*:303-305 (1987); Lu, A.S. *et al.*, 1992, *Genomics 14*:249-255 (1992)). Duplexes without mismatches are not cut. By using radioactively labeled nucleic acid fragments to anneal to a test DNA, it is possible to use these enzymes to generate specific size fragments when a mutation is

present in the test DNA. The fragments are distinguished from uncut fragments by means of polyacrylamide gel electrophoresis. The major problems with these methods are that they require the use of RNA (RNase method) or have the ability to detect only a limited number of mismatches (MutY method).

Mismatch-containing DNA duplexes have also been distinguished from perfectly matched duplexes by means of denaturing gel electrophoresis. In this system, duplexes are run on a polyacrylamide gel in a denaturing gradient under conditions where mismatch-containing DNA denatures more readily than the identical duplex lacking a mismatch, such that the two kinds of duplexes migrate differently. This method, while sensitive and accurate, is extremely laborious and requires a high level of technical sophistication.

Two other methods of mutation detection depend on the failure to extend or join fragments of DNA when mismatches are present. Both require the use of standard DNA oligonucleotides that end precisely at the site of the mutation in question such that, when annealed to test DNA, it is the last base of the oligonucleotide which is mismatched. Mismatch detection depends either on (a) the inability of DNA polymerase to extend an oligonucleotide with a mismatched terminal base or (b) the inability of DNA ligase to join two oligonucleotides when there is a mismatch at the joint between them. Fragment length is determined by gel electrophoresis. Presence of longer fragments than the input oligonucleotides indicates that a mismatch, *i.e.*, mutation, was not present in the test DNA. These methods are also somewhat laborious, require that the exact location of the mutation be known and are difficult to interpret when the sample DNA is heterozygous for the mutation in question. Therefore, they are not practical for use in screening for polymorphisms.

A chemical method for cleavage of mismatched DNA (Cotton, R.G. *et al.*, *Proc. Natl. Acad. Sci. USA* *85*:4397-4401 (1988); Cotton, R.G., *Nuc. Acids Res*

17:4223-4233 (1989)) is based on chemical cleavage at a mismatch site in a DNA-DNA heteroduplex, using a number of agents, in particular osmium tetroxide and hydroxylamine. In this method DNA probes are prepared by restriction enzyme cleavage of DNA of interest. Plasmid DNA containing the sequence of interest is hybridized to labeled probe DNA (either end-labeled or internally labeled with ^{32}P). Hydroxylamine chemically modifies mismatched cytosines; osmium tetroxide modifies mismatched thymines. Piperidine is then used to cleave the DNA at the modified sites, followed by polyacrylamide gel electrophoresis (PAGE) and autoradiography to identify the cleavage products. This method is said to have the advantage of detecting all possible single base pair mismatches because, the method also results in cleavage at a matched base pair in the vicinity of a mismatch.

Publications from Caskey's laboratory (Caskey, C.T. *et al.*, European Patent Publication 333,465 (9/20/89); Grompe, M. *et al.*, *Proc. Natl. Acad. Sci. USA 86*:5888-5892 (1989)) disclose a method for localizing a mutation which utilizes PCR-amplified cDNA as a source of template for the mismatch cleavage reaction. This technique was successfully applied in studying ornithine transcarbamoylase (OTCase) deficiency patients to map mutations.

Kung *et al.*, U.S. Patent 4,963,658, discloses detection of single stranded DNA (ssDNA) by binding with a high-affinity ssDNA-binding protein, such as a topoisomerase or a DNA unwinding protein which itself can be bound to a label, such as β-D-galactosidase.

MISMATCH REPAIR SYSTEMS AND MISMATCH BINDING PROTEINS

DNA mismatch repair systems employ a family of proteins including proteins which recognize and bind to mismatch-containing DNA, which are designated mismatch binding proteins (MBPs). For reviews, see Radman, M. *et al.*, *Annu. Rev. Genet. 20*:523-538 (1986); Radman, M. *et al.*, *Sci. Amer.*, August 1988, pp. 40-46; Modrich, P., *J.*

Biol. Chem. *264*:6597-6600 (1989)). The MutS protein was identified as such a component of the *E. coli* mismatch repair system. See, for example, Lahue, R.S. *et al.*, *Science* *245*:160-164 (1989); Jiricny, J. *et al.*, *Nucl. Acids Res.* *16*:7843-7853 (1988); Su, S.S. *et al.*, *J. Biol. Chem.* *263*:6829-6835 (1988); Lahue, R.S. *et al.*, *Mutat. Res.* *198*:37-43 (1988); Dohet, C. *et al.*, *Mol. Gen. Genet.* *206*:181-184 (1987); and Jones, M. *et al.*, *Genetics* *115*:605-610 (1987). Analogous proteins are known in other bacterial species including MutS in *Salmonella typhimurium* (Lu, A.L. *et al.*, *Genetics* *118*:593-600 (1988); Haber L.T. *et al.*, *J. Bacteriol.* *170*:197-202 (1988); Pang, P.P. *et al.*, *J. Bacteriol.* *163*:1007-1015 (1985)) and the hexA protein of *Streptococcus pneumoniae* (Priebe S.D. *et al.*, *J. Bacteriol.* *170*:190-196 (1988); Haber *et al.*, *supra*).

Purified MutS protein binds DNA containing mispaired bases, but does not bind DNA without mismatches or single-stranded DNA. The MutS-DNA interaction does not result in any degradation or modification of the DNA. None of the above references disclose the possibility of using a MBP or immobilized MBP as part of a mutation detection assay or for purposes of removing mismatched DNA from amplified DNA samples.

SUMMARY OF THE INVENTION

The present inventor has conceived of the use of immobilized mismatch binding protein (MBP), for example, the MutS protein of *E. coli*, for (1) the detection of genetic mutations or genomic polymorphisms, (2) the purification of amplified DNA samples by removing contaminating sequences and sequences containing errors introduced during the amplification processes, and (3) the identification of specific alleles in multi-allelic systems.

The nucleic acid, preferably DNA, being analyzed can be obtained from any source, including blood cells,

tumor tissues, cells in culture or any tissue, and can be obtained from any species including humans. The DNA may be labeled by any of a variety of well-known methods, using colorimetric, chemiluminescent or radioactive markers. In fact, it is not necessary to label test DNA at all.

For detection of mutations or polymorphism, the assay can be performed with a labeled competing oligonucleotide. For purification of amplified DNA, no label is required. For allele identification, the label must be in a synthesized single-stranded oligonucleotide probe.

The methods of the present invention depend on the creation of mismatches in the test DNA which are revealed by denaturing the test DNA and allowing it to reanneal. When testing for heterozygosity or for polymorphism within a test DNA sample, the test DNA can simply be self-annealed, resulting in formation of mismatches when the single strands reanneal with a strand descended from the other parental chromosome. If no heterozygosity exists, no mismatches will be formed. In this case, the label can be in the primers used for amplification or may be added to the termini of the test DNA if amplification is not required.

A similar procedure and labeling scheme is used to remove sequences containing errors introduced during amplification of DNA or minority sequence species. In these cases, the material which does not bind to the MBP is recovered, and contains only those duplex sequences without mismatches. These sequences will therefore be greatly enriched for the majority sequence in the amplified population. When the starting material contains only one sequence, the unbound material will contain those sequences which are identical to the starting material, while those sequences containing errors introduced during amplification will, provided they are relatively rare, have formed mismatches which are retained by the immobilized MBP.

To detect homozygous mutations, it in necessary to anneal the test DNA in the presence of known wild-type

sequences. Such sequences can be synthesized artificially or created during amplification by adding known wild-type sequences to the starting material before amplification. When annealing is performed in the presence of known wild-type sequences, the assay will detect both homozygous and heterozygous mutations.

For allele identification, it is necessary to add a labeled single-stranded probe DNA to the test DNA after amplification. The probe sequence must be identical to the allele of interest such that no mismatches are formed when it anneals to DNA of that allele. The sequence of the probe is selected so that it forms mismatches when annealed to the DNA of any other allele. When test DNA is annealed to such a probe (with the test DNA in excess such that the probe sequences all anneal to form a duplex) and exposed to excess immobilized MBP, the presence of unbound label indicates that the allele in question is present in the test DNA sample.

Thus, the present invention is directed to a method for detecting a mutation from a non-mutated sequence of a target polynucleotide, preferably DNA, in a sample, comprising:

(a) incubating a detectably labeled polynucleotide or oligonucleotide from the sample with an immobilized mismatch-binding protein under conditions in which mismatch-containing polynucleotide molecules bind to the immobilized protein; and

(b detecting the binding of any mismatch-containing polynucleotide from the sample to the mismatch-binding protein,

whereby the presence of detectably labeled polynucleotide or oligonucleotide bound to the mismatch-binding protein is indicative of a mutation in the sequence of the target polynucleotide.

Also provided is a method for detecting a mutation from a non-mutated sequence of a double stranded target mammalian polynucleotide in a sample, comprising:

(a) denaturing any double stranded polynucleotide in the sample followed by allowing DNA strands to reanneal;

(b) incubating the denatured and reannealed double stranded nucleotide of step (a) with a mismatch-binding protein immobilized on a solid support, either

(i) in the presence of a detectably labeled mismatch-containing oligonucleotide capable of binding to the MBP; or

(ii) wherein the MBP was preincubated with and allowed to bind a detectably labeled mismatch-containing oligonucleotide; and

(c) detecting the amount of detectably labeled mismatch-containing oligonucleotide bound to the mismatch-binding protein

whereby the presence of a mutation in the double stranded mammalian polynucleotide of the sample results in a decrease in the binding of the detectably labeled oligonucleotide to the mismatch-binding protein.

A preferred MBP in the above methods is the *E. coli* MutS protein or a functional derivative thereof.

Preferred solid supports on which the mismatch binding protein is immobilized include, but are not limited to, modified cellulose, polystyrene, polypropylene, polyethylene, dextran, nylon, polyacrylamide and agarose. A most preferred solid support is a nitrocellulose membrane.

A preferred detectable label for the detectably labeled poly- or oligonucleotides in the above methods is biotin.

Also provided is a method for the removal from an amplified DNA sample of minority sequences and sequences containing sequence errors introduced during the process of amplification, which method comprises:

(a) subjecting the amplified DNA sample to conditions of denaturation followed by reannealing, such that minority sequence or error-containing

sequences form mismatch-containing DNA duplexes, thereby generating a mixture containing mismatched duplexes;

(b) incubating the mixture of step (a) with an immobilized mismatch-binding protein so that mismatch-containing duplexes bind to the MBP; and

(c) removing the immobilized MBP to which mismatch-containing DNA has bound from the amplified DNA sample,

thereby removing the sequences containing sequence errors.

In another embodiment is provided a method for identifying a specific allele in a multi-allelic system in a sample of amplified DNA, comprising:

(a) mixing a detectably labeled oligonucleotide probe which is perfectly complementary to the DNA sequence of the specific allele with an excess of amplified test DNA under conditions of denaturation followed by annealing such that, after denaturing and annealing, every copy of the probe will be found in a duplex DNA;

(b) incubating the mixture of step (a) with an excess of immobilized MBP such that all mismatch-containing DNA is retained on the immobilized MBP;

(c) removing said immobilized MBP to which has bound any mismatch-containing DNA from the amplified test DNA; and

(d) detecting the presence of the detectably labeled probe in the sample from which the immobilized MBP has been removed,

wherein, the presence of labeled DNA in the sample indicates that the probe is perfectly complementary to an allele in the test DNA.

In the above methods, the immobilized MBP may be (a) in a form which is removable by centrifugation, or (b) immobilized onto a column support material wherein the flow through material is devoid of mismatch-containing duplexes,

or (c) immobilized on a filter support such the filtrate is devoid of mismatch-containing duplexes.

The present invention is also directed to a kit useful for detecting a mutation from a non-mutated sequence of a target polynucleotide sequence in a sample, the kit being adapted to receive therein one or more containers, the kit comprising:

(a) a first container containing an immobilizable mismatch-binding protein;

(b) a second container containing a solid support capable of immobilizing the MBP; and

(c) a third container or a plurality of containers containing a reagent or reagents capable of detecting the binding of a detectably labeled mismatch-containing nucleic acid hybrid to the mismatch-binding protein.

Also provided is a kit useful for detecting a mutation from a non-mutated sequence of a target polynucleotide sequence in a sample, the kit being adapted to receive therein one or more containers, the kit comprising:

(a) a first container containing a mismatch-binding protein immobilized on a solid support; and

(b) a second container or a plurality of containers containing a reagent or reagents capable of detecting the binding of a detectably labeled mismatch-containing nucleic acid hybrid to the mismatch-binding protein.

In the above kits, the MBP is preferably MutS or a functional derivative thereof. The solid support is preferably selected from the group consisting of natural cellulose, modified cellulose, most preferably nitrocellulose, or polystyrene, polypropylene, polyethylene, dextran, nylon, polyacrylamide and agarose.

Also provide is a mismatch binding protein,

preferably the MutS protein or a functional derivative thereof, immobilized on a solid support, such that said immobilized mismatch binding protein is capable of binding to mismatch-containing polynucleotide molecules.

5 **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 shows the results of a direct assay of mismatches using nitrocellulose-bound MutS. Increasing amounts of biotinylated mismatch-containing DNA (upper 2 lines) or mismatch-free DNA (lower 2 lines) were added to 10 the reaction mixtures.

Figure 2 shows the results of a competition assay of mismatched duplexes using nitrocellulose-bound MutS protein. Increasing amounts of unlabeled mismatch-containing 30-mer (upper 2 lines) or mismatch-free 30-mer 15 (lower 2 lines) were added to 5 ng biotinylated mismatch-containing 30-mer. The far right column on the figure represents wells which contained no MutS.

Figure 3 shows the results of binding of mismatch containing DNA to *E. coli* mutS immobilized on 20 nitrocellulose. Mismatch-containing DNA duplexes (2157 and Bio-Het+) show darker (or visible bands) at concentrations where homoduplex DNA shows lighter (or no) bands.

Figure 4 shows the nucleotide sequences of synthetic oligonucleotides (30mers) prepared with a single 25 mismatch at position 15 or 16 or with 1 - 4 unpaired bases between positions 15 and 16. Mismatched or unpaired bases are shown in boldface. Results of studies detecting these mismatches or unpaired bases are shown in Figure 5.

Figure 5 shows the results of binding of DNA 30 duplexes containing the indicated mismatches, or unpaired bases, to *E. coli* mutS immobilized on nitrocellulose.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present inventor conceived of a new, broadly applicable and relatively simple method for detecting a

single base change in a DNA sequence or several such base changes. This method is based upon the formation of a mismatch-containing heteroduplex when a strand of mutant DNA and a "complementary" strand of wild-type DNA
5 hybridize.

The presence of the mismatch is detected in a highly specific manner by first allowing the DNA to bind to an immobilized mismatch-binding protein (MBP), such as the MutS protein of *E. coli*. The presence of DNA bound to the
10 MBP is then detected in any of a number of ways, depending on the label used and whether the assay is a direct assay or a competitive assay. This method stands in stark contrast to methods of the prior art which employ mismatch cutting nuclease enzymes capable of breaking DNA at or near
15 a mispaired base pair.

The methods described herein provide a mutation/polymorphism detection system having the advantages of (a) simplicity, (b) accuracy, (c) ability to be used without radioactivity, (d) ability to detect all
20 single base substitution mutations and addition or deletion mutations of 1-4 bases.

Standard reference works setting forth the general principles of recombinant DNA technology and cell biology, and describing conditions for isolation and
25 handling of nucleic acids, denaturing and annealing nucleic acids, hybridization assays, and the like, include: Sambrook, J. *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989; Albers, B. *et al.*, *MOLECULAR BIOLOGY OF
30 THE CELL*, 2nd Ed., Garland Publishing, Inc., New York, NY, 1989; Watson, J.D., *et al.*, *MOLECULAR BIOLOGY OF THE GENE*, Volumes I and II, Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA, 1987; Darnell, J.E. *et al.*, *MOLECULAR CELL BIOLOGY*, Scientific American Books, Inc., New York, NY,
35 1986; Lewin, B.M., *GENES II*, John Wiley & Sons, New York, NY, 1985, which references are hereby incorporated by reference in their entirety.

MBPs are proteins of around 100 kDa, have been identified in and isolated from both bacteria and higher organisms and selectively bind DNA containing mismatched bases. MBPs have been found in yeast (Valle G *et al.*, 1991 *Yeast* 7:981-988; Miret J.J. *et al.*, 1993, *J. Biol Chem.* *268*:3507-3513), as well as in humans (Stephenson, C. *et al.*, 1989, *J. Biol. Chem. 264*:21177-21782; Karran, P *et al.*, 1990, *Mutat. Res. 236*:269-275; Hughes M.J. *et al.*, 1992, *J. Biol. Chem. 267*:23876-23882; Reenan, A.G. *et al.*, 1993, *Genetics 132*:963-973; Reenan, A.G. *et al.*, 1993, *Genetics 132*:975-985). Mismatch binding proteins from Xenopus and from mouse have been cloned by M. Radman and colleagues.

A preferred MBP is characterized by its ability to bind DNA-DNA (or DNA-RNA or RNA-RNA) duplexes containing mispaired or unpaired bases, to the significant exclusion of single stranded polynucleotides or perfectly matched duplexes. In a preferred embodiment, the intact native MutS protein from *E. coli* is used. However, as used herein, the term "mismatch binding protein" or "MBP" is intended to encompass a functional derivative of the intact, native protein. By "functional derivative" is meant a "fragment," "variant," "analog," or "chemical derivative" of the protein which retains the ability to bind to a mismatch-containing nucleic acid heteroduplex, which permits its utility in accordance with the present invention.

A "fragment" of a MBP refers to any subset of the molecule, that is, a shorter peptide. A "variant" of the protein refers to a molecule substantially similar to either the entire protein or a DNA-hybrid-binding fragment thereof. A variant of a mismatch-binding protein, for example, of MutS, may be prepared by recombinant DNA methods well-known in the art.

A preferred functional derivative of MutS is a homologue of *E. coli* MutS in another species, such as the MutS protein of *Salmonella typhimurium* (Lu, A.L. *et al.*, *supra*; Haber L.T. *et al.*, *supra*; Pang, P.P. *et al.*, *supra*)

or the hexA protein of *Streptococcus pneumoniae* (Priebe S.D. *et al.*, *supra*; Haber *et al.*, *supra*). In addition, possible eukaryotic homologues of MutS or HexA can also be used, such as those encoded by the homologous sequences identified in human, mouse, frog or hamster DNA (Shimada, T. *et al.*, *J. Biol. Chem.* *264*:20171 (1989); Linton, J. *et al.*, *Molec. Cell. Biol.* 7:3058-3072 (1989); Fujii, H. *et al.*, *J. Biol. Chem.* *264*:10057 (1989)).

A "chemical derivative" of the MBP contains additional chemical moieties not normally a part of the protein, including additional stretches of amino acids as in a fusion protein. Covalent modifications of the peptide are included within the scope of this invention. Such modifications may be introduced into the molecule by reacting targeted amino acid residues of the protein with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues.

In selecting a protein as being a useful MBP for the methods of the present invention, assays can be performed by one of ordinary skill in the art using conventional methods. Thus, for example, in evaluating a sample for the presence of a MBP useful in the present invention, one can perform a mismatch binding assay, such as that described for MutS by Jiricny *et al.* (which reference is hereby incorporated by reference in its entirety). Preferably, a filter binding assay is used. To prepare the oligonucleotide heteroduplex, an oligonucleotide, preferably of about 16 bases is labeled with ^{32}P using a kinase reaction and γ-^{32}P-ATP using a kinase such as T4-polynucleotide kinase. The 5'-labeled oligonucleotide (which can be stored at -20°C) is then annealed with a complementary oligonucleotide having a single base pair mismatch under standard conditions. The annealed 16 base pair heteroduplex is mixed with an excess of the protein being tested and kept on ice for 30 minutes. The mixture is then applied to a nitrocellulose filter which has been prewetted in the assay buffer. Gentle suction is applied for several seconds, and the filter is

washed extensively with ice-cold assay buffer. The filter is then dried in air, suspended in scintillation fluid and counted. By virtue of the protein sticking to the filter, any counts on the filter can be attributed to binding to
5 the putative MBP. In the absence of such a protein, the labeled oligonucleotide heteroduplex will pass through the filter. Thus, by using such a simple assay, one can easily detect and select a MBP useful in the methods of the present invention.

10 As used in the present invention, the MBP is immobilized to a solid support or carrier. By "solid support" or "carrier" is intended any support capable of binding a protein. Well-known supports or carriers include natural cellulose, modified cellulose such as
15 nitrocellulose, polystyrene, polypropylene, polyethylene, dextran, nylon, polyacrylamide, and agarose or Sepharose®. Also useful are magnetic beads. The support material may have virtually any possible structural configuration so long as the immobilized MBP is capable of binding to the
20 target nucleic acid molecule. Thus, the support configuration can include microparticles, beads, porous and impermeable strips and membranes, the interior surface of a reaction vessel such as test tubes and microtiter plates, and the like. Preferred supports include nitrocellulose
25 disks or strips. Those skilled in the art will know many other suitable carriers for binding the MBP or will be able to ascertain these by routine experimentation.

Most preferred is a solid support to which the MBP is attached or fixed by covalent or noncovalent bonds.
30 Preferably, noncovalent attachment is by adsorption using methods that provide for a suitably stable and strong attachment. The MBP is immobilized using methods well-known in the art appropriate to the particular solid support, providing that the ability of the MBP to bind
35 mismatch-containing DNA is not destroyed.

The immobilized MBP is then easily used to detect

heterozygosity (or polymorphism) as well as single base mutations, or to isolate mismatch-containing DNA from a mixture, or to rid a mixture of mismatch-containing DNA.

In one embodiment, the surface of polystyrene or other plastic multiwell plates serves as the solid support. In another embodiment, a solid support to which the MBP is bound is affixed to the bottom of wells of multiwell plates.

In a preferred embodiment, the immobilization and DNA binding can be performed in a 96 well blotting apparatus and the resulting sheet of nitrocellulose (or other support) paper can be removed to evaluate reactions. For example, color development on the nitrocellulose can be used to evaluate binding based on an enzyme as part of the detection system and a chromogenic or chemiluminescent substrate for the enzyme serving as the precursor of the color reactions.

Following attachment of the MBP to the support, the support is treated ("blocked") to prevent further binding of proteins or nucleic acids, using methods and reagents well-known in the art.

The immobilized MBP is contacted with and allowed to bind (to saturation) small oligonucleotide heteroduplex molecules. The oligonucleotides preferably have about 30 base pairs. For testing, a DNA duplex containing a mismatch which is well recognized (*i.e.*, bound) by the MBP is used.

PREPARATION OF OLIGONUCLEOTIDES CONTAINING OR LACKING MISMATCHES

Such oligonucleotides are prepared using a nucleotide modified at the 5′ end with a detectable label such that they can be quantitatively detected by appropriate detection methods, preferably spectrophotometry or chemiluminescence. In a preferred embodiment, the oligonucleotide is biotin-modified, and is detectable using a detection system based on avidin or streptavidin which binds with high affinity to biotin. The streptavidin can

be conjugated to an enzyme, the presence of which is detected using a chromogenic substrate and measuring the color developed.

Examples of useful enzymes in the methods of the present invention are horseradish peroxidase, alkaline phosphatase, glucose-6-phosphate dehydrogenase, malate dehydrogenase, staphylococcal nuclease, delta-V-steroid isomerase, yeast alcohol dehydrogenase, α-glycerophosphate dehydrogenase, triose phosphate isomerase, asparaginase, glucose oxidase, β-galacto-sidase, ribonuclease, urease, catalase, glucoamylase and acetylcholinesterase.

The detectable label may also be a radioisotope which can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

The detectable label may also be a fluorescent compound. When the fluorescently labeled molecule is exposed to light of the proper wave length, its presence can then be detected due to fluorescence using microscopy or fluorometry. Among the most commonly used fluorescent labelling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The detectable label may be a fluorescence emitting metal such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the oligonucleotide using metal chelating groups such as diethylenetriaminepentaacetic acid or ethylenediaminetetra-acetic acid.

The detectable label may be a chemiluminescent compound. The presence of the chemiluminescent-tagged molecule is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the oligonucleotide. Bioluminescence is a type of

chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

A MBP bound to the DNA-DNA, DNA-RNA or RNA-RNA hybrid can be detected either directly or indirectly. For direct detection, the poly- or oligonucleotide duplex is detectably labeled using labels as discussed herein.

For indirect detection, the assay utilizes competition of binding to the MBP of the test DNA with an already bound or a contemporaneously exposed mismatch-containing duplex. Thus, a labeled mismatch-containing oligonucleotide is pre-bound to the MBP or is incubated together with the MBP and test DNA. The more mismatch-containing DNA in the test sample, the less binding of the labeled oligonucleotide to the MBP will occur.

The test sample to be assayed can be in any medium of interest, and will generally be a sample of medical, veterinary, environmental, nutritional, or industrial significance. Human and animal specimens and body fluids particularly can be assayed by the present method, providing that they contain cells from which nucleic acids can be prepared. Preferred sources include blood, sperm, other tissue, milk, urine, cerebrospinal fluid, sputum, fecal matter, lung aspirates, throat swabs, genital swabs and exudates, rectal swabs, and nasopharyngeal aspirates.

DETECTION OF HETEROZYGOSITY OR POLYMORPHISM

To detect heterozygosity or polymorphism in DNA from a diploid organism, test DNA is preferably prepared by denaturing and annealing PCR-amplified DNA from a diploid organism. The test DNA is prepared with labeled primers, annealed and added to a well or other reaction vessel which contains immobilized MBP already bound to mismatched

oligonucleotides. Alternatively, test DNA can be mixed with mismatched oligonucleotides and the mixture added to a well or other vessel containing immobilized MBP

A spectrophotometric reading is made at the wavelength appropriate for quantitative detection of the test DNA. After an incubation period suitable to allow either (1) binding of the test DNA to the immobilized MBP or (2) displacement of the mismatched oligonucleotide from the immobilized MBP, the DNA solution is removed, the well washed, and a spectrophotometric reading made at the wavelength appropriate for quantitative detection of the bound mismatched oligonucleotide.

The ratio of the reading for test DNA to the reading for the mismatched oligonucleotide will be vastly different for mismatch-containing and mismatch-free test DNA over a wide range of DNA concentrations. Standard curves are prepared using known quantities of DNA to allow characterization of test DNA as homozygous or heterozygous without having to quantitate the test DNA prior to the assay. Thus, a single DNA sample is sufficient to determine heterozygosity and a single 96 well microplate will allow the testing of at least about 80 different DNA samples.

DETECTION OF HOMOZYGOUS MUTATIONS

To detect homozygous mutations, known homozygous wild-type DNA must be combined with the test DNA sample before denaturing and annealing. Only test DNA containing a mutation (homozygous) will form mismatch-containing DNA that can compete with the mismatched oligonucleotide for binding to the immobilized MBP.

The method of the present invention can be used with (a) only the mismatched oligonucleotide labeled or (b) only the test DNA labeled. However, both of these methods require that the nucleic acid concentration be determined and that the test be performed at several different test DNA concentrations.

When the mismatched oligonucleotides are labeled, the test is based on competition with several different concentrations of test DNA and comparison of the resulting

curve (with concentration expressed as moles of duplex molecules) with standard curves for mismatched and non-mismatched standards.

When the test DNA is labeled, the test involves measuring the extent of binding to the MBP of several concentrations below saturation and comparison of the resulting curve with standard curves for mismatched and non-mismatched standards.

PURIFICATION OF AMPLIFIED DNA SAMPLES

One of the most revolutionary and widely used technologies currently employed in modern molecular biology is the process of polymerase chain reaction (PCR), which amplifies DNA sequences from starting amounts so minute as to be nearly undetectable. For reviews of PCR, see: Mullis, K.B., 1986, *Cold Spring Harbor Symp. Quant. Biol.* *51*:263-273; Saiki, R.K. *et al.*, 1985, *Bio/Technology* *3*:1008-1012; and Mullis, K.B. *et al.*, 1987, *Meth. Enzymol.* *155*:335-350. In addition, because PCR can amplify specific sequences, it allows the purification of specific sequences, basically in a single step, from genomic DNA. PCR is an essential component of virtually all studies of the human genome, is a central component of gene identification and cloning, is increasingly used in the diagnosis of genetic and infectious diseases and is widely used in forensics.

However, for some applications, in particular, gene cloning and mutation detection, PCR suffers from an inherent tendency of the polymerases to make mistakes by inserting incorrect, non-complementary bases during synthesis. Although the fidelity of most replicative polymerases *in vivo* is such that they insert only one incorrect base for every 10^{10} bases replicated, polymerases used in PCR can have error rates as high as one incorrect base for every 10^4 bases replicated. This high an error rate can mean that a significant fraction of the amplified molecules will not be identical in sequence with the starting material.

The present methods are useful for purifying amplified DNA samples using an immobilized MBP to remove

minority sequences and molecules containing sequence alterations introduced by the amplification process. For example, if a DNA segment is amplified through 20 rounds of replication (a common amount of amplification), a significant fraction of the final molecules may contain one or more incorrect bases. In cloning experiments, this greatly increases the risk of cloning a nucleotide sequence different from the starting sequence.

In mutation detection assays involving denaturation and annealing of a PCR-amplified sample, incorrect bases inserted during PCR may be scored as if they were mutations in the original sample. Thus, for accurate mutation detection it is necessary to eliminate all DNA molecules with sequence alterations introduced by PCR copy errors. The method described here accomplishes this purification in a simple and straight-forward manner.

Immobilized MBP can be used to purify amplified DNA samples. MBP is immobilized by binding to solid phase supports, preferably nitrocellulose filters, sepharose beads or magnetic beads. The filters or beads are treated, if necessary, to prevent the binding of double-stranded DNA. The amplified DNA sample is denatured, by heating, and allowed to reanneal. Given the random nature of PCR mistakes, virtually all incorrect bases will be found in mismatched base pairs after annealing.

The immobilized MBP is added to the sample and the solution mixed by gentle shaking. The immobilized MBP, and any bound mismatch-containing DNA, is removed, for example, by removing the filter, by allowing the beads to settle out of solution or by removing the beads magnetically, depending on the nature of the solid support used. This leaves behind precisely matched DNA duplexes.

In addition to purifying amplified DNA samples by removing molecules containing errors introduced during amplification, purification using an immobilized MBP is used to enrich for majority sequences when samples of diverged, repeated DNA sequences, such as of immunoglobulin genes, are being examined.

To remove completely a minority species from a sample amplified from a mixed population of DNA (with

respect to sequence), it may be necessary to perform more than one round of purification as described herein and possibly more than one round of amplification.

Note that the present method can be used to purify sequences from both homozygous and heterozygous amplified sequences, since half of the parental sequences in a heterozygous sample will anneal to the complementary strand of the same parental heritage and thus form a molecule without mismatches. In other words, when the starting material is heterozygous, half of the annealed molecules will be removed from the sample because they contain a mismatch due to differences in the starting sequences. However, half the annealed molecules will not contain such mismatches and so will be removed from the sample only if they contain mismatches which were created as a result of errors during amplification. In any event, mutation detection assays will require a second round of denaturing and annealing.

ALLELE IDENTIFICATION IN MULTI-ALLELIC SYSTEMS

As more alleles of disease-causing genes are identified, and in the quest to develop a polymorphism map of the human genome, it is becoming increasingly important to be able to identify particular alleles of a given gene. Immobilized MBPs provide a simple and straightforward means of allele identification.

Unique, labeled oligonucleotide probes are prepared for each allele of a given gene, such that the probe is perfectly complementary to only one allele, *i.e.*, the probe will form one or more mismatches when paired with an incorrect allele. The probe is mixed with an excess of amplified test DNA such that, after denaturing and annealing, every copy of the probe will be found in duplex. The process is repeated with probes for every allele in question. The annealed DNA mixture is then either:

(1) mixed with immobilized MBP on a support that can be removed from the suspension by centrifugation;

(2) passed through a micro-column of immobilized MBP on an appropriate column support; or

(3) passed through a filter support containing immobilized MBP.

In any case, the immobilized MBP must be in excess such that all mismatch-containing DNA is retained. The supernatant, column flow through or filtrate is analyzed for the presence of label. Only in those cases where the probe is perfectly complementary to an allele in the test DNA will label be detected.

In order to be certain that no single-stranded probe sequences are present, it may be necessary, or at least desirable, to include some single-stranded DNA binding component on the support for the immobilized MBP. This system works equally well for homozygous or heterozygous conditions.

KITS

The present invention is also directed to a kit or reagent system useful for practicing the methods described herein. Such a kit will contain a reagent combination comprising the essential elements required to conduct an assay according to the methods disclosed herein. The reagent system is presented in a commercially packaged form, as a composition or admixture where the compatibility of the reagents will allow, in a test device configuration, or more typically as a test kit, *i.e.*, a packaged combination of one or more containers, devices, or the like holding the necessary reagents, and usually including written instructions for the performance of assays. The kit of the present invention may include any configurations and compositions for performing the various assay formats described herein.

In all cases, the reagent system will comprise (1) an immobilizable or immobilized MBP or functional derivative, preferably mutS, and (2) additional reagents useful in carrying out the assay. The kit may optionally contain labeled mismatch-containing oligonucleotides. For detecting a particular mutation, a kit may also contain labelled primers for carrying out a PCR. A kit according to the present invention can additionally include ancillary

chemicals such as the components of the solution in which binding of duplexes to the immobilized MBP takes place.

Having now generally described the invention, the same will be more readily understood through reference to
5 the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

EXAMPLE I

BINDING OF MISMATCH-CONTAINING DNA BY
10 IMMOBILIZED MISMATCH BINDING PROTEIN

A. Materials and Methods

1. Preparation of immobilized MBP

A nitrocellulose sheet (0.45μm, Schleicher & Schull) was wet with reaction buffer (20mM Tris pH 7.6,
15 0.01 mM EDTA, 5 mM $MgCl_2$, 0.1 mM DTT) and placed in a dot blot apparatus (Bio-Rad).

Purified MBP, *E. coli* MutS, at a concentration of 0.5 μg/10 μl reaction buffer was spotted on the nitrocellulose paper in each well. The wells were
20 incubated at room temperature, and the remaining liquid was pulled through with vacuum. Each well was washed twice with 100 μl reaction buffer by adding the solution to the well and then pouring it out. After the second wash, the remaining solution was pulled through by vacuum.

25 2. Blocking

The nitrocellulose filter was blocked with bovine serum albumin (BSA) to prevent the binding of other proteins or nucleic acids. Reaction buffer (200 μl) containing 1% (w/v) BSA was added to each well. After 1
30 hour at room temperature, the solution was poured out and each well washed with 2 x 100μl reaction buffer by adding the solution to the well and then pouring it out. After the second wash, the remaining liquid was pulled through with vacuum.

3. Oligonucleotides

The sequence of the oligonucleotides used in these studies was taken from the 30 base region surrounding the site of the sickle cell mutation in the human β-globin gene. The mismatch is at the site of the sickle cell mutation, although the mutant sequence used to form the mismatch is not the sickle cell mutation (the actual sickle cell mutation is an A:T → T:A transversion). Biotinylated oligonucleotides were biotinylated on the 5′ end of the mutant strand. Biotinylation is accomplished during synthesis by adding a biotin-modified nucleotide to the 5′ end of the oligonucleotides.

G:T Mismatch:
Mutant GCACCTGACT CCTG**G**GGAGA AGTCTGCCGT [SEQ ID NO:1]
Wild-type CGTGGACTGA GGAC**T**CCTCT TCAGACGGCA [SEQ ID NO:2]

No Mismatch:
Mutant GCACCTGACT CCTGGGGAGA AGTCTGCCGT [SEQ ID NO:1]
Mutant CGTGGACTGA GGACCCCTCT TCAGACGGCA [SEQ ID NO:3]

4. Binding DNA

Biotinylated oligonucleotides, in 20 μl reaction buffer containing 1% BSA, were added to each well. After 30 minutes at room temperature, remaining liquid was poured out. Each well was washed with 5 x 100 μl reaction buffer by adding the solution to the well and then pouring it out. After the fifth wash, the remaining solution was pulled through by vacuum.

5. Binding Streptavidin-conjugated Horse Radish Peroxidase (HRP)

The presence of biotin was detected by its binding of Streptavidin. A 100 μl volume of Streptavidin-conjugated HRP (Pierce Chemicals) at a concentration of 50 mg/ml in reaction buffer + 1% BSA was added to each well. After 2 hours at room temperature, the solution was poured out and each well washed with 5 x 100 μl reaction buffer by adding the solution and then pouring it out. After the fifth wash, the remaining solution was pulled through with vacuum.

6. Enhanced ChemiLuminescence® (ECL) Development The nitrocellulose sheet was removed from dot blot apparatus and washed 3 times in a petri dish with 10 ml reaction buffer. Five ml ECL development solution (Amersham) was poured over the nitrocellulose. The substrate for HRP in this reagent is a chemiluminescent compound. After 1 minute, the solution was removed. The nitrocellulose was blotted dry and placed between 2 clear plastic sheets. The nitrocellulose thus protected was exposed to X-ray film in the dark for varying periods of time. In the experiments reported here, the exposure time was 1 minute.

7. Competition

In competition studies, the DNA binding was as described above, except that a constant amount of biotinylated mismatch-containing oligonucleotide (5ng) was mixed with varying amounts of unlabeled DNA, with or without a mismatch, and added to the wells.

B. RESULTS

1. Specific binding of mismatch-containing DNA by immobilized MBP

Figure 1 shows the results (in duplicate) of adding increasing amounts of biotinylated mismatch-containing DNA (upper 2 lines) or mismatch-free DNA (lower 2 lines).

The immobilized MBP can detect as little as 0.2 ng of mismatch-containing 30-mer, whereas no detectable binding of mismatch-free 30-mer was observed even with 200 ng of DNA. (The lines around the lower spots were artifacts of incomplete washing.)

2. Competition Assay

Figure 2 shows the results (in duplicate) of adding increasing amounts of unlabeled mismatch-containing 30-mer (upper 2 lines) or mismatch-free 30-mer (lower 2 lines) to 5 ng biotinylated mismatch-containing 30-mer. Although competition was clearly

visible with 50 ng of mismatch-containing DNA, the mismatch-free DNA did not compete until 500 ng, if at all. The far right column on the figure contains no MBP.

5 The results indicate that, at least with the 30mers used above, immobilized MBP discriminates between mismatch-containing and perfectly paired DNA with an efficiency of at least three orders of magnitude. Similar results have been obtained using 10 54mers with a sequence derived from the V3 loop of HIV. Therefore, even if the discrimination decreases as the amount of perfectly paired duplex increases, the discrimination efficiency when using 300mers, considered to be the maximum useful length for 15 polymorphism studies of the human genome, should be on the order of a factor of 100.

EXAMPLE II

HETEROZYGOTE DETECTION IN HUMAN GENOMIC DNA

The methods described above were used to 20 detect heterozygosity at a specific location in human genomic DNA. PCR amplification was performed on a portion of exon 3 of the human glucokinase gene extending from codon 98 (encoding glutamine) to the stop codon (Soffel *et al.*, *Proc. Natl. Acad. Sci. USA* 25 89:7698-7702 (1992)). The wild-type double stranded sequence of 100 bases (SEQ ID NO:4 and SEQ ID NO:5) corresponding to exon 3 human glucokinase is as follows:

```
    5'  gcactaacttcagggtgatgctggtgaaggtgggagaagg
30  3'  cgtgattgaagtcccactacgaccacttccaccctcttcc

        tgaggaggggcagtggagcgtgaagaccaaacaccagatg
        actcctccccgtcacctcgcacttctggtttgtggtctac

        tactccatccccgaggacgcc 3'  (SEQ ID NO:4)
        atgaggtaggggctcctgcgg 5'  (SEQ ID NO:5)
```

35 In the heterozygous DNA (see below) The C of the underlined CAG codon was mutated to T. The DNA

sequences tested were PCR amplified from genomic DNA obtained from: (1) a known heterozygote at that position, (2) a known homozygote at that position and (3) a presumed homozygote at that position.

5 Test DNAs were denatured, by heating, allowed to reanneal and tested for the presence of mismatches (*i.e.*, heterozygotes) by testing their binding in an immobilized mismatch binding protein assay according to the present invention, utilizing *E. coli* MutS.

10 A. Materials and Methods:

1. PCR Amplification

The following templates were used:

(a) 2157 - heterozygote at glucokinase gene exon 3

(b) DGK-101 - Human genomic DNA (homozygote at
15 glucokinase gene exon 3)

(c) Human genomic DNA, male (presumed homozygote at glucokinase gene exon 3) designated Sigma DNA (commercially obtained from Sigma Chemical Co.)

20 The primers (obtained from Operon) were HPLC purified and had the following sequences, corresponding to the 5′ termini of the two DNA strands SEQ ID NO:4 and SEQ ID NO:5:

Primer #1: 5′ (biotin)-GCACTAACTTCAGGGTGATG

25 Primer #2: 5′ -GCGTCCTCGGGGATGGAGTA

PCR Primer #1 contained biotin bound at the 5′ end to allow detection in the ECL detection system described below. Primer #2 was radioactively labeled with a 5-′^{32}P-phosphate to allow quantitation of the
30 different amplification products after removal of unused primers. Primer #2 was ^{32}P labeled using a kinase reaction prior to use in amplification. The kinase reaction mix contained: 70mM Tris HCl, pH 7.6; 10 mM mgCl$_2$; 5mM DTT; 20μCi ^{32}P-ATP; 30 units T4
35 polynucleotide kinase; and 500ng DNA (primer #2). The kinase reaction was performed in a 20μl reaction volume at 37°C for 30 min. Kinase was inactivated by heating at 70°C for 10 min. DNA was stored at -20°C.

The PCR reaction included: 10mM Tris HCl pH 8.3; 50 mM KCl; 1.5mM $MgCl_2$; 0.001% gelatin (w/v); 0.05mM dATP; 0.05mM dTTP; 0.05mM dGTP; 0.05mM dCTP; 0.1μM primer #1; 0.075μM primer #2; 0.025μM ^{32}P primer #2; 200ng template DNA; and 2.5 units AmpliTaq DNA polymerase (Perkin-Elmer). Reaction volume was 100μl. Amplification was carried out for 30 cycles in a Perkin-Elmer thermocycler by denaturing at 90°C for 1 min., annealing at 55°C for 1 min. and extension at 72°C 2 min. Unused primers were removed by centrifugal dialysis using Centricon 30 microconcentrators (Amicon) according to manufacturer's protocol. PCR products were quantitated by running equal amounts (measured as cpm of radioactivity) on non-denaturing 8% polyacrylamide gels, staining with ethidium bromide and comparing them to standard DNA..

2. Immobilized Mismatch Binding Protein Assays:

DNA was denatured and annealed (in Perkin-Elmer thermocycler) according to the following schedule: 100°C for 4 min; 50°C for 1 hour; 75°C for 4 min; 50°C for 30 min, followed by cooling to room temperature.

A nitrocellulose sheet (0.45Mm, Schleicher and Schuell, BA85) was wet by floating in reaction buffer (20mM Tris HCl, pH 7.6; 5mM $MgCl_2$; 0.1mM DTT; 0.01mM EDTA) and placed in a slot blot apparatus (Hoefer Scientific Instruments) over 3 sheets of blotting paper (Schleicher and Schuell GB002). 100μl of reaction buffer was added to each well. After 5 min at room temperature, any remaining buffer was pulled through by vacuum. MutS (500 ng in 20μl reaction buffer) was added to each well. The same volume of reaction buffer was added to "No MutS" wells. The apparatus was kept at room temperature for 20 min before proceeding to the next step.

The nitrocellulose was blocked by adding 200μl of HRP-free BSA to each well. After 1 hr at room

temperature, any remaining solution was pulled through by vacuum.

DNA preparations were added in 20μl reaction buffer containing 3% HRP-free BSA. After 30 min at room temperature, wells were washed 5 times with 100μl reaction buffer by adding the solution to the wells and decanting it.

The presence of biotin-labeled DNA bound on the nitrocellulose sheet was detected by visualizing the binding of streptavidin to biotin. 100 μl streptavidin-HRP in reaction buffer containing 3% HRP-free BSA was added to each well. After 20 min at room temperature, any remaining solution was decanted. Wells were washed 5 times with 100μl reaction buffer by adding solution to each well and then decanting. Any solution remaining after the fifth wash was removed by vacuum.

3. Enhanced Chemiluminescence (ECL) development

The nitrocellulose sheet was removed from the apparatus and washed 4 times for 1 min with 50 ml reaction buffer in a small tray. Nitrocellulose was blotted dry and immersed in 10 ml of ECL development solution (Amersham). After 1 min, the nitrocellulose sheet was removed, blotted dry and placed between two clear plastic sheets. The nitrocellulose thus protected was exposed to X-ray film in the dark for 30 sec.

B. RESULTS

The results are shown in Figure 3. Bio-Het refers to a synthetic 30mer duplex of SEQ ID NO:1 and SEQ ID NO:3. This duplex contains a G:T mismatch at position 15. Bio-Homo refers to a synthetic 30mer duplex of SEQ ID NO:1 and SEQ ID NO:2, which lacks the mismatch at position 15. No binding of Bio-Homo was detected even at 10ng of DNA whereas Bio-Het was detected at amounts as low as 0.1ng. No binding was observed in the absence of MutS ("no MutS" columns)

indicating that all DNA binding observed is MutS dependent.

Binding of heterozygote nucleic acid (2157) to mutS was clearly visible at 0.6ng. Homozygote binding (independent of the source of the DNA) was faintly detectable at 1.25ng and was clearly detected at 2.5ng. Thus, heterozygote DNA was detected at least 2-4 times better than homozygote DNA in this assay. The binding of homozygote DNA at higher concentrations was considered to be the result of errors introduced during amplification by the Taq polymerase. Such a high error rate in nucleotide incorporation by this polymerase is a well-known phenomenon. Thus, such inappropriate binding still represented mismatch binding by the immobilized MutS.

The use of immobilized mismatch binding protein assays for mutation, heterozygosity or polymorphism detection are limited only by the ability to provide substrates free of random mismatches, such as those produced by polymerase errors during PCR amplification.

EXAMPLE III

DETECTION OF SPECIFIC MISMATCHES AND UNPAIRED BASES BY IMMOBILIZED MISMATCH BINDING PROTEIN

Studies were performed to detect single base mismatches and one to four unpaired bases in heteroduplex DNA. These studies utilized synthetic oligonucleotides (30mers, which include SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3) prepared with a single mismatch at position 15 or 16 or with 1 - 4 unpaired bases between positions 15 and 16. The sequences of these oligonucleotides are shown in Figure 4. As described above, these sequences were taken from, or are closely related to, the human β-globin gene in the region of the mutation responsible for sickle-cell anemia. The oligonucleotides were prepared with a 5′ biotin label (to allow detection by the ECL detection

system) and were annealed to unlabeled oligonucleotides such that only one strand of each duplex was labeled. The duplexes were used in a mismatch detection assay using immobilized mismatch binding protein (MutS), essentially as described above in the previous Examples.

Oligonucleotides were diluted in TNE buffer (10mM Tris HCl, pH 8.0; 0.01M NaCl, 1mM EDTA). Biotin labeled oligonucleotides were diluted to 10 mg/μl and unlabeled oligonucleotides to 100 ng/μl. Equal volumes of diluted oligonucleotides were mixed and annealed at 70°C for 10 min, room temperature for 30 min, followed by quenching on ice and were stored at -20°C. The unlabeled oligonucleotides were in 10-fold excess to assure that all biotin-labeled strands were in duplexes.

The results are presented in Figure 5. All mismatches except C:C and G:A were detected in these assays. T:C and C:A mismatches were not teseted in this experiment. Not all detected mismatches are detected equally well. The order of detection sensitivity appeared to be
G:T > G:G > C:T > A:C > T:T > A:A = A:G > C:C.

The fact that A:G mismatches were better detected than G:A mismatches suggested that the sequence of the individual strands may influence the extent of mismatch detection at least in relatively short oligonucleotides such as those used here. However, G:T and T:G mismatches were equally well-detected, suggesting that well-detected mismatches are detected independent of strain orientation.

Heteroduplexes with 1 - 4 unpaired bases were also detected. A heteroduplex with two unpaired bases was more easily detected than a heteroduplex with one unpaired base. Heteroduplexes with three unpaired bases were less well detected than any of the detectable mismatches, and heteroduplexes with four unpaired bases were even less well detected.

These results indicate that immobilize mismatch binding protein assays will detect all

mutations due to a single base change. Thus, although C:C went undetected, the corresponding G:G mismatch, which would of necessity occur in any mutant:wild type pairing that had a C:C mismatch, was easily detected. In addition, the detection system can easily detect mutations arising from the addition of 1 - 3 bases.

The references cited above are all incorporated by reference herein, whether specifically incorporated or not.

Having now fully described this invention, it will be appreciated by those skilled in the art that the same can be performed within a wide range of equivalent parameters, concentrations, and conditions without departing from the spirit and scope of the invention and without undue experimentation.

While this invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. This application is intended to cover any variations, uses, or adaptations of the inventions following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth as follows in the scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method for detecting a mutation from a non-mutated sequence of a target polynucleotide in a sample, comprising:
   (a) incubating a detectably labeled polynucleotide or oligonucleotide from said sample with a mismatch-binding protein immobilized on a solid support, under conditions in which mismatch-containing polynucleotide molecules bind to said immobilized protein; and
   (b detecting the binding of any mismatch-containing polynucleotide from said sample to said mismatch-binding protein,

whereby the presence of detectably labeled polynucleotide or oligonucleotide bound to the mismatch-binding protein is indicative of a mutation in the sequence of the target polynucleotide.

2. A method according to claim 1 wherein said mismatch-binding protein is the MutS protein or a functional derivative thereof.

3. A method according to claim 1 wherein said solid support is selected from the group consisting of natural cellulose, modified cellulose, polystyrene, polypropylene, polyethylene, dextran, nylon, polyacrylamide and agarose.

4. A method according to claim 3, wherein said solid support is a nitrocellulose membrane.

5. A method according to claim 1 wherein said detectably labeled polynucleotide or oligonucleotide is labeled with a chromogenic compound, a chemiluminescent compound, a bioluminescent compound, a fluorescent compound or a radiolabel.

6. A method according to claim 1 wherein said detectably labeled polynucleotide or oligonucleotide is labeled with biotin.

7. A method according to claim 1 wherein
5 said target polynucleotide is DNA.

8. A method for detecting a mutation from a non-mutated sequence of a double stranded target mammalian polynucleotide in a sample, comprising:
(a) denaturing any double stranded
10 polynucleotide in said sample into single strands and allowing said single strands to reanneal into duplexes;
(b) incubating the denatured and reannealed duplexes of step (a) with a mismatch-
15 binding protein immobilized on a solid support, either
(i) in the presence of a detectably labeled mismatch-containing oligonucleotide capable of binding
20 to said mismatch binding protein; or
(ii) wherein said a mismatch-binding protein was preincubated with and allowed to bind a detectably labeled mismatch-containing oligonucleotide;
25 and
(c) detecting the amount of detectably labeled mismatch-containing oligonucleotide bound to said mismatch-binding protein,
30 whereby the presence of a mutation in said double stranded mammalian polynucleotide of the sample results in a decrease in the binding of the detectably labeled oligonucleotide to the mismatch-binding protein.

9. A method according to claim 8 wherein said solid support is selected from the group consisting of natural cellulose, modified cellulose, polystyrene, polypropylene, polyethylene, dextran, 5 nylon, polyacrylamide and agarose.

10. A method according to claim 9, wherein said solid support is a nitrocellulose membrane.

11. A method for the removal from an amplified DNA sample of minority sequences and 10 sequences containing sequence errors introduced during the process of amplification, which method comprises:

(a) subjecting said amplified DNA sample to conditions of denaturation followed by reannealing, such that minority sequence or 15 error-containing sequences form mismatch-containing DNA duplexes, thereby generating a mixture containing mismatched duplexes;

(b) incubating the mixture of step (a) with an immobilized mismatch-binding protein so that 20 mismatched duplexes bind to said mismatch binding protein; and

(c) removing said immobilized mismatch binding protein to which mismatch-containing DNA has bound from said amplified DNA sample,

25 thereby removing said sequences containing sequence errors.

12. A method for identifying a specific allele in a multi-allelic system in a sample of amplified DNA, comprising:

(a) mixing a detectably labeled oligonucleotide probe which is perfectly complementary to the DNA sequence of said specific allele with an excess of amplified test DNA under conditions of denaturation followed by annealing such that, after denaturing and annealing, every copy of the probe is in a duplex DNA;

(b) incubating the mixture formed in step (a) with an excess of immobilized mismatch binding protein, such that all mismatch-containing DNA is retained on said immobilized mismatch binding protein;

(c) removing said immobilized mismatch binding protein to which has bound any mismatch-containing DNA from said amplified test DNA; and

(d) detecting the presence of said detectably labeled probe in said sample from which said immobilized mismatch binding protein has been removed,

wherein, the presence of labeled DNA in said sample indicates that the probe is perfectly complementary to an allele in the test DNA.

13. A kit useful for detecting a mutation from a non-mutated sequence of a target polynucleotide sequence in a sample, said kit being adapted to receive therein one or more containers, said kit comprising:

(a) a first container containing an immobilizable mismatch-binding protein;

(b) a second container containing a solid support capable of immobilizing said mismatch binding protein; and

(c) a third container or a plurality of containers containing a reagent or reagents capable of detecting the binding

of a detectably labeled mismatch-containing nucleic acid duplex to said mismatch-binding protein.

14. A kit useful for detecting a mutation from a non-mutated sequence of a target polynucleotide sequence in a sample, said kit being adapted to receive therein one or more containers, said kit comprising:

(a) a first container containing a mismatch-binding protein immobilized on a solid support;

(b) a second container or a plurality of containers containing a reagent or reagents capable of detecting the binding of a detectably labeled mismatch-containing nucleic acid duplex to said mismatch-binding protein.

15. A kit according to claim 13 wherein said mismatch-binding protein is MutS or a functional derivative thereof.

16. A kit according to claim 14 wherein said mismatch-binding protein is MutS or a functional derivative thereof.

17. A kit according to claim 13 wherein said solid support is selected from the group consisting of natural cellulose, modified cellulose, polystyrene, polypropylene, polyethylene, dextran, nylon, polyacrylamide and agarose.

18. A kit according to claim 17 wherein said solid support is a nitrocellulose membrane.

19. A kit according to claim 14 wherein said solid support is selected from the group consisting of natural cellulose, modified cellulose, polystyrene, polypropylene, polyethylene, dextran, nylon,
5 polyacrylamide and agarose.

20. A kit according to claim 19 wherein said solid support is a nitrocellulose membrane.

21. A mismatch-binding protein immobilized on a solid support, such that said immobilized mismatch
10 binding protein is capable of binding to mismatch-containing polynucleotide molecules.

22. The immobilized mismatch-binding protein of claim 21 which is the MutS protein or a functional derivative thereof.

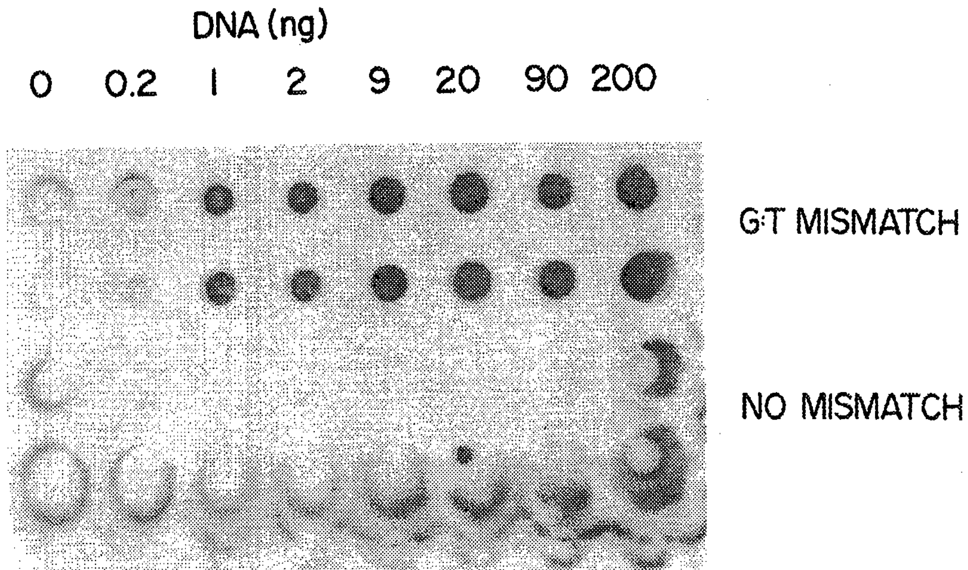


FIG. 1

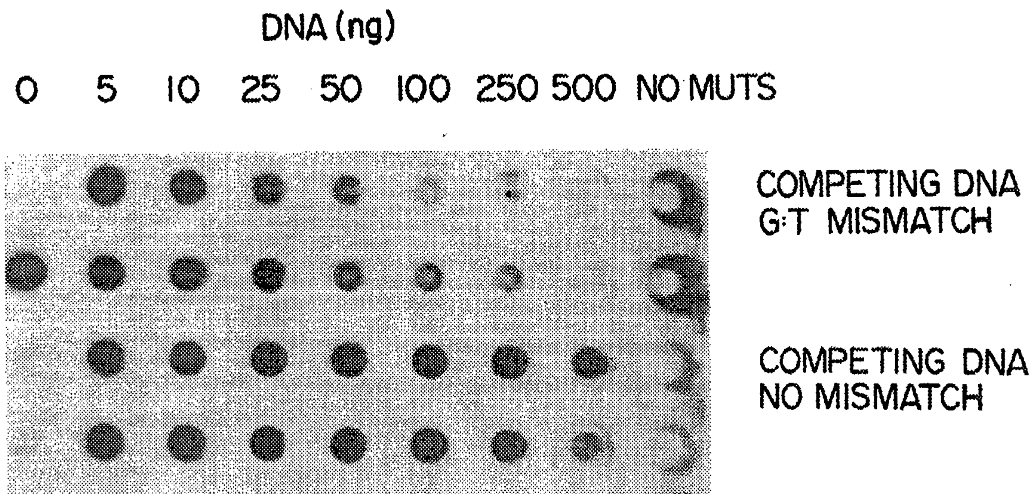


FIG. 2

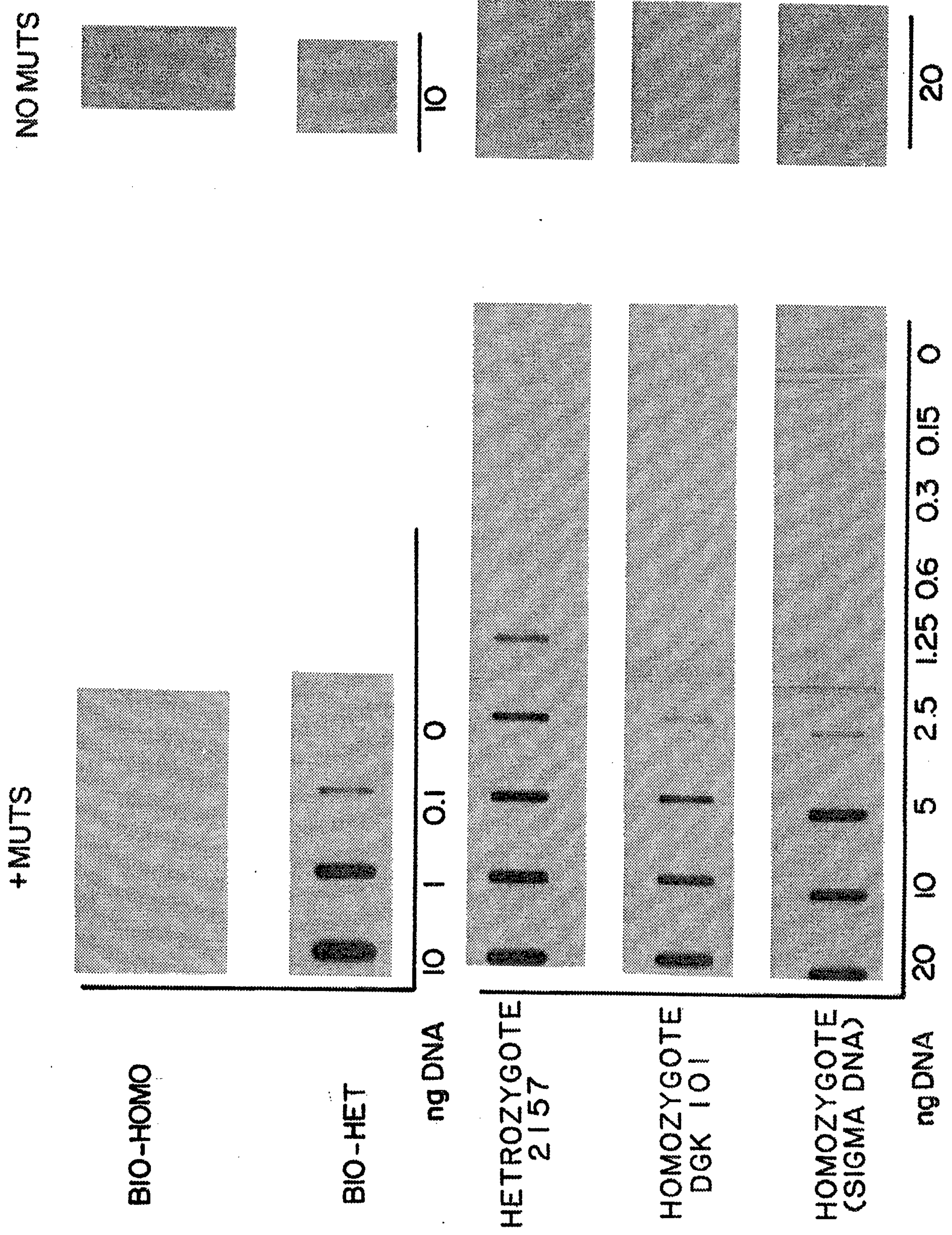


FIG. 3

Type	Nucleotide Sequence	SEQ ID NO:
Homoduplex:	GCACCTGACTCCTGGGGAGAAGTCTGCCGT	1
	CGTGGACTGAGGACCCCTCTTCAGACGGCA	3
G:T Mismatch:	GCACCTGACTCCTG**G**GGAGAAGTCTGCCGT	1
	CGTGGACTGAGGAC**T**CCTCTTCAGACGGCA	2
T:G Mismatch	GCACCTGACTCCTG**T**GGAGAAGTCTGCCGT	6
	CGTGGACTGAGGAC**G**CCTCTTCAGACGGCA	7
C:T Mismatch	GCACCTGACTCCTG**C**GGAGAAGTCTGCCGT	8
	CGTGGACTGAGGAC**T**CCTCTTCAGACGGCA	2
G:G Mismatch	GCACCTGACTCCTG**G**GGAGAAGTCTGCCGT	9
	CGTGGACTGAGGAC**G**CCTCTTCAGACGGCA	10
A:G Mismatch	GCACCTGACTCCTG**A**GGAGAAGTCTGCCGT	11
	CGTGGACTGAGGAC**G**CCTCTTCAGACGGCA	12
G:A Mismatch	GCACCTGACTCCTG**G**GGAGAAGTCTGCCGT	1
	CGTGGACTGAGGAC**A**CCTCTTCAGACGGCA	13
A:C Mismatch	GCACCTGACTCCTG**A**GGAGAAGTCTGCCGT	14
	CGTGGACTGAGGAC**C**CCTCTTCAGACGGCA	3
A:A Mismatch	GCACCTGACTCCTG**A**GGAGAAGTCTGCCGT	15
	CGTGGACTGA GGAC**A**CCTCTTCAGACGGCA	16
T:T Mismatch	GCACCTGACTCCTG**T**GGAGAAGTCTGCCGT	17
	CGTGGACTGAGGAC**T**CCTCTTCAGACGGCA	2
C:C Mismatch	GCACCTGACTCCTGG**C**GAGAAGTCTGCCGT	18
	CGTGGACTGAGGACC**C**CTCTTCAGACGGCA	3
Unpaired C	GCACCTGACTCCTGG**C**GGAGAAGTCTGCCGT	19
	CGTGGACTGAGGACC CCTCTTCAGACGGCA	3
Unpaired CA	GCACCTGACTCCTGG**CA**GGAGAAGTCTGCCGT	20
	CGTGGACTGAGGACC CCTCTTCAGACGGCA	3
Unpaired CAG	GCACCTGACTCCTGG**CAG**GGAGAAGTCTGCCGT	21
	CGTGGACTGAGGACC CCTCTTCAGACGGCA	3
Unpaired CAGG	GCACCTGACTCCTGG**CAGG**GGAGAAGTCTGCCGT	22
	CGTGGACTGAGGACC CCTCTTCAGACGGCA	3

FIG. 4

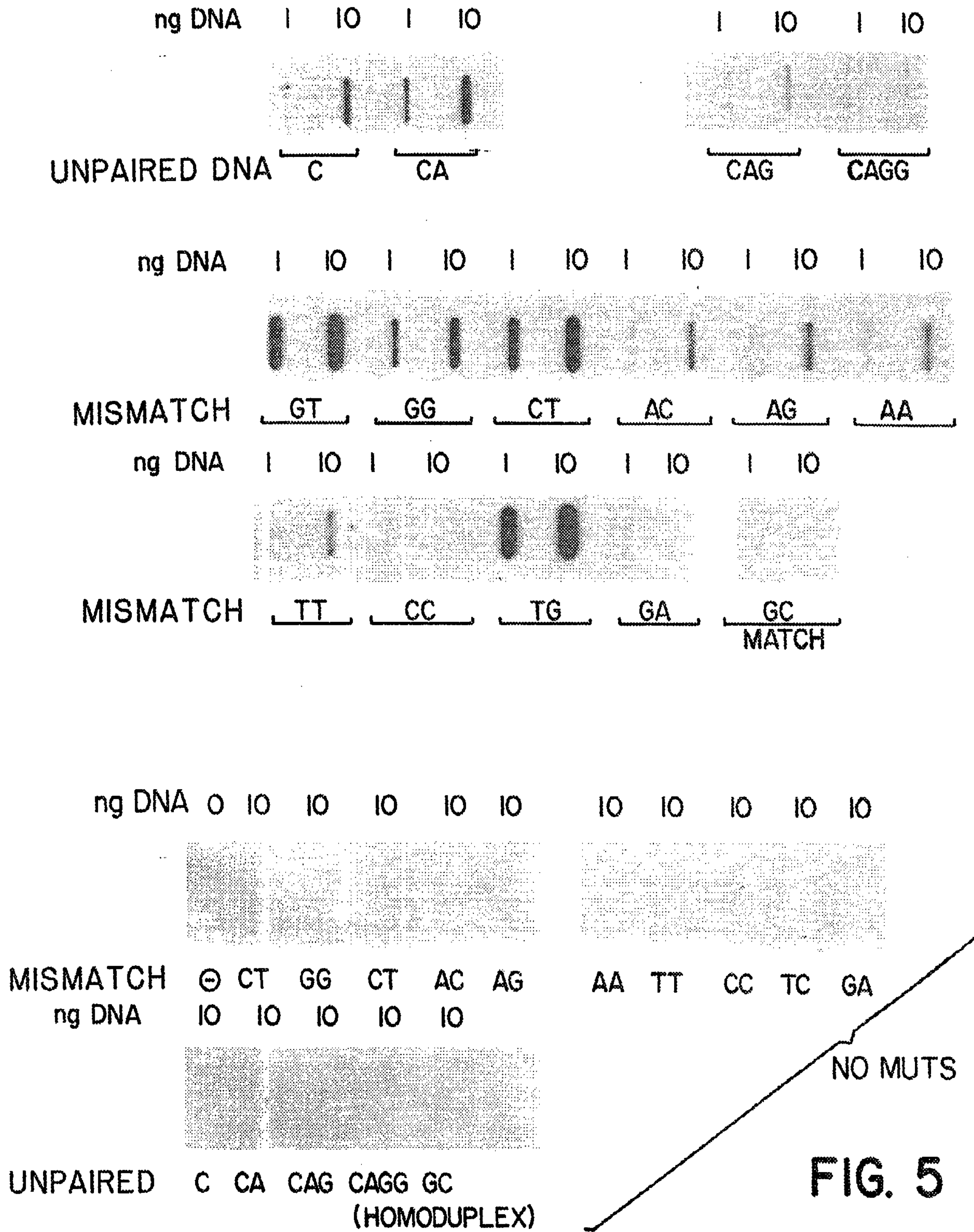


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/12768

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; G01N 33/543, 33/544, 33/548, 33/566
US CL :435/6, 7.5; 436/501, 518, 528, 529, 530, 531, 536
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 7.5, 975; 436/501, 518, 528, 529, 530, 531, 536

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG
search terms: mismatch binding protein?, MUTS, MUT(w)S, assay, detect?, remov?, test

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	NUCLEIC ACIDS RESEARCH, Volume 16, Number 16, issued 1988, J. Jiricny et al, "Mismatch-Containing Oligonucleotide Duplexes Bound by the *E. coli mutS*-Encoded Protein", pages 7843-7853, see entire document.	13, 15, 17, 18 -------- 1-12, 14, 16, 19-22
Y	WO, A, 93/02216 (WAGNER ET AL.) 04 February 1993, see entire document .	1-12, 14, 16, 19-22
Y	AMERICAN JOURNAL OF HUMAN GENETICS, Volume 51, Number 4, issued October 1992, A. I. Lishanskaya et al, "Mutation Detection in the Cystic Fibrosis Gene Using an E. coli Mismatch Binding Protein, mutS", page A385, Abstract Number 1517, see entire abstract.	1-10, 12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

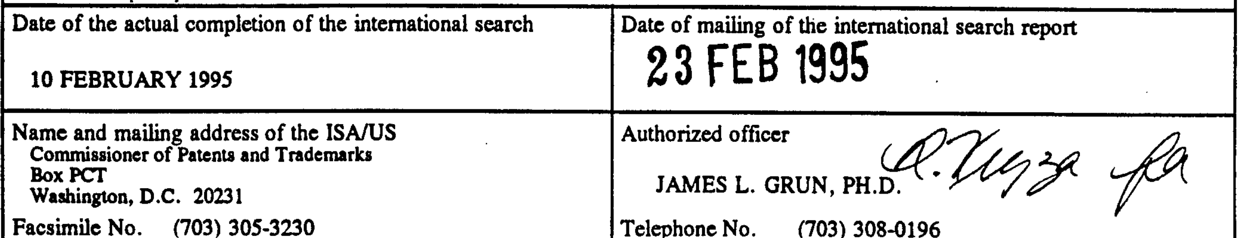
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family



Date of the actual completion of the international search	Date of mailing of the international search report
10 FEBRUARY 1995	23 FEB 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer JAMES L. GRUN, PH.D.
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/12768

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 91, issued March 1994, A. Lishanski et al, "Mutation Detection by Mismatch Binding Protein, MutS, in Amplified DNA: Application to the Cystic Fibrosis Gene", pages 2674-2678, see entire document.	1-22
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 86, issued August 1989, R.K. Saiki et al, "Genetic Analysis of Amplified DNA with Immobilized Sequence-Specific Oligonucleotide Probes", pages 6230-6234, see especially page 6230.	12
A	JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 263, Number 14, issued 15 May 1988, S. Su et al, "Mispair Specificity of Methyl-Directed DNA Mismatch Correction *In Vitro*", pages 6829-6835, see entire document, especially pages 6833-6834.	1-22