(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 185 693 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
23.08.2006 Bulletin 2006/34

(21) Application number: **00918364.1**

(22) Date of filing: **24.03.2000**

(51) Int Cl.:
C12Q 1/68 *(2006.01)*

(86) International application number:
PCT/US2000/007882

(87) International publication number:
WO 2000/058514 (05.10.2000 Gazette 2000/40)

(54) **METHODS FOR IMPROVING SENSITIVITY AND SPECIFICITY OF SCREENING ASSAYS FOR CANCER AND PRECANCER**

METHODEN ZUR VERBESSERUNG DER SENSITIVITÄT UND SPEZIFITÄT VON SCREENINGVERFAHREN FÜR KREBS UND KREBS-VORSTUFEN

METHODES PERMETTANT D'AMELIORER LA SENSIBILITE ET LA SPECIFICITE DES METHODES DE CRIBLAGE POUR CANCER ET PRECANCER

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **26.03.1999 US 277016**

(43) Date of publication of application:
13.03.2002 Bulletin 2002/11

(73) Proprietor: **Exact Sciences Corporation**
Marlborough, Massachusetts 01752 (US)

(72) Inventors:
- **LAPIDUS, Stanley, N.**
  Bedford, NH 03110 (US)
- **SHUBER, Anthony, P.**
  Milford, MA 01757 (US)

(74) Representative: **Crump, Julian Richard John et al**
Mintz Levin Cohn Ferris Glovsky and Popeo
Intellectual Property LLP
The Rectory
9, Ironmonger Lane
London EC2V 8EY (GB)

(56) References cited:
WO-A-95/00669 **WO-A-96/29430**
DE-A- 19 736 691

- **SIDRANSKY DAVID: "Nucleic acid-based methods for the detection of cancer." SCIENCE (WASHINGTON D C), vol. 278, no. 5340, 7 November 1997 (1997-11-07), pages 1054-1058, XP002170462 ISSN: 0036-8075**
- **VET JACQUELINE A M ET AL: "Comparative analysis of p53 mutations in bladder washings and histologic specimens." AMERICAN JOURNAL OF CLINICAL PATHOLOGY, vol. 110, no. 5, November 1998 (1998-11), pages 647-652, XP001009905 ISSN: 0002-9173**
- **CAWKWELL L ET AL: "Frequency of allele loss of DCC, p53, RBI, WT1, NF1, NM23 and APC/MCC in colorectal cancer assayed by fluorescent multiplex polymerase chain reaction." BRITISH JOURNAL OF CANCER, vol. 70, no. 5, 1994, pages 813-818, XP001009897 ISSN: 0007-0920**
- **HOANG JEAN-MARC ET AL: "BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines." CANCER RESEARCH, vol. 57, no. 2, 1997, pages 300-303, XP002162454 ISSN: 0008-5472**

EP 1 185 693 B1

Description

FIELD OF THE INVENTION

[0001] The invention relates to assays to detect nucleic acid markers of cancer with high specificity and high sensitivity for detection.

BACKGROUND OF THE INVENTION

[0002] Cancer is thought to arise from a multi-step process that typically involves multiple genetic mutations leading to uncontrolled cell growth. Many cancers are curable if detected early in their development. For example, colorectal cancers typically originate in the colonic epithelium, and are not extensively vascularized (and therefore not invasive) during early stages of development. The transition to a highly-vascularized, invasive and ultimately metastatic cancer commonly takes ten years or longer. If the presence of cancer is detected prior to extensive vascularization, surgical removal typically is an effective cure. However, colorectal cancer is often detected only upon manifestation of clinical symptoms, such as pain and bloody stool. Generally, such symptoms are present only when the disease is well established, and often after metastasis has occurred. Similarly, with the exception of the Pap smear for detection of pre-malignant cervical lesions, diagnostic screening methods for other types of cancer are best at detecting established disease.

[0003] Most diagnostic assays for cancer are invasive, or at least uncomfortable. Invasive procedures range from performing a tissue biopsy to surgery. Cancer screening procedures frequently result in significant patient discomfort. For example, magnetic resonance imaging requires confinement of the patient, and colonoscopy requires sedation. The discomfort associated with typical invasive screening methods reduces patient compliance with routine screening procedures.

[0004] Moreover, screening for the early detection of cancer (i.e., prior to the onset of symptoms and/or metastasis) often results in an intolerable level of false positive and false negative results. The likelihood of a false positive test result is a function of the specificity of the test. The specificity of a test (expressed as a percentage) is the likelihood that any individual who is negative for disease tests negative for that disease. On the other hand, false negative results are a property of the sensitivity of the test. Sensitivity (also expressed as a percentage) provides the likelihood that a test for a specific disease will identify as positive an individual having that disease. Thus, 5% of the individuals determined to have a disease using an assay with a specificity of 95% will actually not have the disease. Similarly, an assay with a sensitivity of 95% will incorrectly identify a diseased individual as disease-free 5% of the time.

[0005] The problems of sensitivity and specificity are exaggerated in assays for the early detection of cancer because patient samples on which such early detection is performed typically contain relatively small amounts of cancerous cellular material in relation to non-cancerous cellular material. In many cases, patient samples are a heterogeneous mixture of large amounts of normal cells and small amounts of cancerous cells. A good example of such a heterogeneous sample is stool. The typical stool sample contains cells and cellular debris sloughed from the colonic epithelium, byproducts of digestion, and bacteria. In its early stages, colorectal cancer is thought to affect only about 1% of colonic epithelial cells. Any attempt to detect nucleic acids from the 1% of affected cells in the heterogeneous background of the stool sample might give rise to very low sensitivities. Attempts to identify the presence of the indicia of cancer in other heterogeneous samples, such as sputum, pus, urine, nipple aspirate, etc., presents similar problems.

[0006] Recently, a number of genetic mutations have been associated with cancer. For example, alterations in the p53 gene, the Kras oncogene, and the apc tumor suppressor gene are thought to be participants in the multi-step pathway leading to cancer. It has been suggested that mutations in those genes might be a basis for molecular screening assays for the early stages of certain types of cancer. See e.g., Sidransky, et al., Science, 256: 102-105 (1992). Attempts have been made to identify and use nucleic acid markers that are indicative of cancer. However, even when such markers are found, using them to screen patient samples, especially heterogeneous samples, has proven unsuccessful either due to an inability to obtain sufficient sample material, or due to the low sensitivity that results from measuring only a single marker. For example, simply obtaining adequate human DNA from one type of heterogeneous sample (stool) has proven difficult. See Villa, et al., Gastroenterol., 110: 1346-1353 (1996) (reporting that only 44.7% of all stool specimens, and only 32.6% of stools from healthy individuals produced sufficient DNA for mutation analysis). Other reports in which adequate DNA has been obtained have reported low sensitivity in identifying a patient's disease status based upon a single cancer-associated mutation. See Eguchi, et al., Cancer, 77:1707-1710 (1996) (using a p53 mutation as a marker for cancer).

[0007] Sidransky et al (Science 278, 7 Nov 1997, 1054-1058) discloses nucleic acid-based markers for use in detecting cancer, as well as limitations on their clinical use.

[0008] Accordingly, there is a need in the art for high-sensitivity, high-specificity assays for the detection of molecular indicia of cancer, especially in heterogeneous samples.

SUMMARY OF THE INVENTION

[0009] Methods of the invention solve the problem of obtaining accurate (high sensitivity and high specificity) results in an assay for indicia of cancer or precancer in a heterogeneous sample.

[0010] The present invention provides assays conducted on samples obtained from bodily excretions or bodily fluids for cancer or precancer in which the assays have a high sensitivity for detection of cancer or precancer when it is present in a patient sample, and a high specificity against false positive results. In a preferred embodiment, methods of the invention provide the benefits of high sensitivity and high specificity in an assay to detect a small amount of a cancer marker (e.g., a nucleic acid) in a heterogeneous sample having predominantly non-cancerous cells and cellular debris. Accordingly, such methods are especially useful for early detection of cancer or precancer. Methods of the invention greatly increase the accuracy of molecular screening and diagnostic assays for the early detection of cancer or precancer.

[0011] The present invention contemplates that one reason prior art non-invasive methods (i.e., methods conducted on samples obtained non-invasively or minimally invasively) for detecting molecular indicia (especially nucleic acid mutations) of cancer have failed to provide satisfactory results is that such methods have not addressed maintaining high specificity and/or high sensitivity in the assay. Such methods also fail to recognize the benefits of combining high sensitivity and high specificity features in an assay to detect early indicia of cancer, especially when the detection is performed in a heterogeneous sample. The present invention recognizes that screening assays for indicia of cancer, especially early-stage cancer (e.g., when cancer indicia represents about 1% of the cells and cellular debris in an appropriate sample as discussed below), are improved by increasing the specificity and/or the sensitivity of the assay. As described below, the specificity and sensitivity of molecular assays can be improved in several ways.

[0012] According to the present invention, there is provided a method of detecting cancer or precancer, comprising screening a patient sample obtained non-invasively or minimally invasively for the presence of one or more nucleic acid indicia which are indicative of cancer or pre-cancer, characterised in that said nucleic acid indicia are selected from (a) the length of the nucleic acids in the sample; and (b) the ratio of nucleic acids having a length greater than about 200 bp to nucleic acids having a length less than about 200 bp in the sample; wherein said length and said ratio differ between patients with cancer or precancer and patients who do not have cancer or pre-cancer and wherein said sample is obtained from bodily excretions or from bodily fluids.

[0013] In a preferred embodiment, screening assays of the invention analyze at least two, and preferable between about 3 and about 20 markers (e.g., mutations, loss of heterozygosity, sequence length variations, nucleic acid molecular weight variations, variations in amounts of amplifiable nucleic acid) in patient samples in order to improve the sensitivity and specificity of detection. Such "multiple target" assays allow improved sensitivity because the increased number of markers that are analyzed decreases the likelihood that a patient presenting with indicia of cancer or precancer will be misdiagnosed as negative(i.e., as not having cancer).

[0014] Methods of the invention also comprise combining detection of a mutation known or suspected to be associated with cancer or precancer with detection of loss of heterozygosity at a relevant genomic locus. The combination of analyzing both mutations in cancer-associated nucleic acids, and loss of heterozygosity increases with specificity and sensitivity of any screening assay for cancer or precancer. In a preferred embodiment, the sensitivity of the assay is at least 50%, and the specificity of the assay is at least 85%.

[0015] Methods of the invention comprise combining two or more assays for molecular indicia of cancer or prccancer in order to achieve a desired level of informativeness. For example, in a preferred embodiment, methods of the invention comprise combining two or more assays selected from quantitative PCR, multiple mutation analysis, detection of loss of heterozygosity, and hybrid capture of one or more mutant nucleic acid markers. Accordingly, increased sensitivity and specificity is observed when an assay based upon the amount of amplifiable DNA is combined with an assay for a particular cancer-associated mutation. Examples of these "combination assays", and their resulting sensitivities and specificities, are provided in the detailed description. Such methods are especially useful when applied to a heterogeneous sample in which the nucleic acid to be detected is present in a very small amount relative to other nucleic acids (as well as other molecules) in the sample.

[0016] Methods of the invention also make use of the "informativeness" of markers used therein. The Informativeness of a nucleic acid marker relates to the likelihood of finding the marker in a positive sample. Thus, if a particular mutation, for example a mutation in codon 12 of K-ras, has an informativeness for cancer of 56%, this means that 56% of positive patient samples (i.e., those taken from patients who have cancer) have the K-ras mutation. Methods of the invention combine the use of informative markers (e.g., mutations) and high sensitivity/specificity assays in order to provide reliable screening assays for early diagnosis of cancer or precancer, especially in heterogeneous samples. For purposes of the present invention a mutation is a deletion, addition, substitution, rearrangement, or translocation in a nucleic acid. A loss of heterozygosity is a form of mutation in which all or a portion of one allele is deleted. Also for purposes of the present invention, the terms "markers", "targets", and "mutations" include nucleic acid (especially DNA) mutations (substitutions, additions, rearrangements, translocations, deletions, etc.), as well as other nucleic acid indicia useful in methods of the invention. Such indicia include the amount of amplifiable nucleic acid in a sample, the length of nucleic acids in a sample,

the ratio of long nucleic acids (greater than about 200 bp) to short nucleic acids (less than about 200 bp), and any other nucleic acid variations that differ in patients with cancer and disease-free patients. Also for purposes of the present invention, the terms "healthy" or "disease-free" are intended to mean a patient who does not have cancer or precancer.

[0017] Stool is a good example of a heterogeneous sample in which methods of the invention are especially useful. A typical stool sample contains patient nucleic acids, but also contains heterologous nucleic acids, proteins, and other cellular debris consistent with the lytic function of the various nucleases, proteinases, etc. found in the colon. As stool proceeds from the proximal colon to the distal colon, it (under normal circumstances) solidifies. As the solidifying stool passes through the colon, colonic epithelial cells are sloughed onto the stool. If a patient has a developing tumor or adenoma, cells from the tumor or adenoma will also be sloughed onto stool, and they (or their debris) will contain molecular indicia of disease (e.g., mutations or loss of heterozygosity). In the early stages of development, nucleic acid indicative of an adenoma or tumor comprise only about 1% of the nucleic acid in a voided stool. If a patient is left untreated, proportionately more disease-related nucleic acids are found in stool over time. Methods of the invention are useful for detecting early-stage lesions in heterogeneous samples such as stool. Methods of the invention result in a high degree of sensitivity and specificity for the detection of early-stage disease. Methods of the invention are especially useful in detecting, for example, adenomas in the colon. Adenomas are non-metastatic lesions that frequently have the potential for metastasis. If all adenomas in a patient are detected and removed, the probability of complete cure is virtually certain.

[0018] In a preferred embodiment, nucleic acids or nucleic acid mutations having a high degree of informativeness are chosen. One or more assay(s) is (are) conducted to reliably detect one or more informative nucleic acids, and a diagnosis is made based upon the presence in a patient sample of any one of the informative nucleic acids. In a preferred embodiment, nucleic acids are chosen as targets for analysis in methods of the invention based upon their length and/or sequence characteristics. For example, it has now been discovered that the quantity and/or length of nucleic acids in stool (a prototypical heterogeneous sample) presents a high degree of informativeness regarding a patient's disease status. Patients having, for example, an adenoma produce stool specimens containing more and longer DNA than specimens produced by healthy patients. Moreover, a number of highly-informative DNA mutations are useful in methods of the invention. These include, mutations in the oncogene, Kras (especially mutations at codons 12 and 13); mutations in the cell-cycle regulator, p53; mutations in the apc gene; mutations in the bat-26 segment of the MSH2 mismatch repair gene; and loss of heterozygosity (typically indicated by massive loss of DNA in one allele but not the other). Finally, methods of the invention provide that the informativeness of an assay is increased by screening multiple mutations simultaneously or in sequence, or by combining different assays.

[0019] In another preferred embodiment, methods of the invention provide informative molecular assays for cancer or precancer by providing samples for analysis that have sufficient amplifiable nucleic acid. Thus, in one embodiment, methods of the invention comprise screening samples for amplifiable DNA; classifying samples based upon the amount of DNA capable of being amplified from them; and further screening samples with a predetermined threshold of amplified or amplifiable DNA for the presence of a mutation indicative of cancer or precancer.

[0020] In another preferred embodiment, methods of the invention comprise selecting one or more mutational events that are indicative of cancer or precancer, such that the combined informativeness of the one or more events meets or exceeds a predetermined or desired level of informativeness. The informativeness of any mutation or combination of mutations may be validated by an accepted invasive screening technique. For example, in methods to detect colorectal cancer, the informativeness of a molecular assay may be determined by identification of a lesion using colonoscopy.

[0021] A detailed description of certain preferred embodiments of the invention is provided below. Other embodiments of the invention are apparent upon review of the detailed description that follows.

DETAILED DESCRIPTION OF THE DRAWINGS

[0022] Figure 1 is a polyacrylamide gel showing exemplary "A", "B", "C", and "F" amplifications for use in assays to determine an amount of amplifiable nucleic acid in a sample. Lanes 1, 4, 5, 6, and 7 show an "A" amplification, lane 2 shows a "C" application, lane 3 shows a "B" amplification, and lane 8 shows an "F" amplification (no amplifiable DNA).

DETAILED DESCRIPTION OF THE INVENTION

[0023] Methods of the invention provide non-invasive or minimally-invasive assays for the detection of cancer or precancer at early stages of disease. Methods of the invention are especially useful in detecting cancer or precancer in heterogeneous biological samples. Preferred methods comprise identifying in a patient sample one or more nucleic acid mutations(s) that provide high sensitivity and high specificity for detection of the indicia of cancer or precancer. Methods of the invention may comprise identifying mutations having a known informativeness for cancer or precancer, or may be based upon validating selected mutations or assays to detect them with respect to a standard assay for cancer. By utilizing cancer or precancer markers having a high sensitivity/specificity for detecting the presence of cancer or precancer,

methods of the invention provide improvements in non-invasive or minimally-invasive molecular screening assays. For purposes of the present invention, non-invasive or minimally-invasive means that specimens for analysis are obtained either from bodily excretions (e.g. stool, pus, sputum) or bodily fluids such as blood aspirate, or lymph.

[0024] The invention will be exemplified with experiments to detect the presence of indicia of colorectal cancer or precancer in samples prepared from patient stool specimens. However, the skilled artisan recognizes that methods of the invention can be practiced using a variety of different samples in order to detect a variety of cancers.

[0025] A reason that detection of colorectal cancer or precancer (e.g., an adenoma) is exemplified is that a stool specimen is a good example of a heterogeneous environment in which methods of the invention are especially useful (see above). Moreover, colonoscopy (and sigmoidoscopy, a related technique) is a well-known invasive standard that has a high sensitivity and high specificity (although high cost and low patient compliance) with which methods of the invention can be compared and validated.

[0026] Methods of the invention comprise screening a sample, such as one prepared from a stool specimen, for the presence of one or more marker(s) of cancer or precancer (e.g., a colorectal tumor or adenoma), such that the sensitivity of detection is between about 50% and about 100%, and the specificity of detection is between about 85% and about 100%. In a preferred embodiment, methods of the invention combine different types of assays in order to achieve an overall increase in sensitivity and specificity. Thus, methods of the invention comprise conducting an assay for a mutation known to be associated with cancer or precancer, and an assay for a quantity and/or length of DNA expected to occur in cancer or precancer in order to obtain the combined benefits of the sensitivity and specificity of both assays. Moreover, embedded within the concept of utilizing multiple nucleic acid analyses to detect cancer or precancer is the use of multiple genomic targets in each assay in order to provide further increases in sensitivity and specificity. However, as shown below, a single-marker assay is sufficient for practice of the invention if its sensitivity and specificity are within the ranges taught herein.

[0027] The genomic targets and assay methods used according to the invention can vary depending upon the desired level of sensitivity and specificity, as well as the type of cancer or precancer the detection of which is desired. Genomic targets (e.g., mutations) are selected based upon their known sensitivity or specificity or by determining a baseline sensitivity and specificity. In preferred embodiments, methods of the invention comprise the detection of a mutation at a single, informative locus. In other embodiments, assays for informative loci are combined in order to achieve improved sensitivity and specificity of detection relative to invasive techniques. Accordingly, methods of the invention contemplate a combination of assays selected from multiple mutation detection, quantitative polymerase chain reaction (i.e., to determine the amount of amplifiable DNA in a sample), sequence-specific hybrid capture, oligo-ligation, amplification refractory mutation system, single-stranded conformational polymorphism detection, sequencing, mismatch detection, and single base extension. Target loci include chromosomes 1, 5, 8, 17, and 18, particularly chromosome 5q, chromosome 17p, chromosome 8p, chromosome 1q, and chromosome 18q. Preferred loci for use in methods of the invention include p53, apc, bat-26, and others suspected to be predictive of cancer or precancer.

[0028] Other genes are known to be associated with colorectal cancer, and their sensitivity and specificity are determined when not known in the literature by determining the percentage of tumors bearing the mutation, and the percentage of healthy specimens that bear the mutation from a sufficiently large and diverse population. This can be done empirically, or mathematically using algorithms that predict the likelihood of false positive and false negative screening results based upon data relating the presence of a mutation to the presence of cancer or precancer. Confirmation of a patient's clinical status can be accomplished by a standard test such as colonoscopy in the case of colorectal cancer (which has a typical sensitivity of 95% and a typical specificity of 100%).

[0029] For the analysis of stool samples, preferred methods of the invention comprise obtaining at least a cross-section or circumfrential portion of a voided stool as taught in U.S. patent number 5,741,650. While a cross-sectional or circumfrential portion of stool is desirable, methods provided herein are conducted on random samples obtained from voided stool, which include smears or scrapings. Once obtained, the stool specimen is homogenized. A preferable buffer for homogenization is one that contains at least 16mM ethylenediaminetetraacetic acid (EDTA). However, it has been discovered that the use of at least 150mM EDTA greatly improves the yield of nucleic acid from stool. Thus, a preferred buffer for stool homogenization comprises phosphate buffered saline, 20-100 mM NaCl or KCl, at least 150mM EDTA, and optionally a detergent (such as SDS) and a proteinase (e.g., proteinase K).

[0030] After homogenization, nucleic acid is preferably isolated from the stool sample. Isolation or extraction of nucleic acid is not required in all methods of the invention, as certain detection techniques can be adequately performed in homogenized stool without isolation of nucleic acids. In a preferred embodiment, however, homogenized stool is spun to create a supernatant containing nucleic acids, proteins, lipids, and other cellular debris. The supernatant is treated with a detergent and proteinase to degrade protein, and the nucleic acid is phenol-chloroform extracted. The extracted nucleic acids are then precipitated with alcohol. Other techniques can be used to isolate nucleic acid from the sample. Such techniques include hybrid capture, and amplification directly from the homogenized stool. Nucleic acids can be purified and/or isolated to the extent required by the screening assay to be employed.

[0031] Nucleic acids to be analyzed are chosen based upon known or suspected relationships between specific

mutations and cancer or precancer. If desired, sequence-specific hybrid capture is used to isolate specific nucleic acids from the sample. Target nucleic acids may be analyzed by any method of the art. Examples of preferred methods include enumerative analysis of the loss of heterozygosity as taught in U.S. patent number 5,670,325. Enumerative methods do not require knowledge of the sequence of a mutant nucleic acid. Rather such methods determine that there has been an alteration (deletion, substitution, addition, rearrangement, or other mutation) in a wild-type nucleic acid. The investigated loci are chosen based upon the likelihood of an alteration being associated with cancer or precancer. Enumerative methods compare the number in a sample of a wild-type nucleic acid known not to be altered in cancer or precancer with the number of a wild-type nucleic acid known or suspected to be altered in cancer or precancer. A statistically-significant difference in the two numbers indicates a positive screen.

[0032] Mutations in target nucleic acids may also be measured by single base extension techniques to identify a single nucleotide variant indicative of cancer or precancer. Preferably, single base extension assay are cycled. Briefly, cycled single base extension reactions comprise annealing a nucleic acid primer immediately 5' to a region containing a single base to be detected. The single base to be detected represents a marker for mutation. The mutation may be a single point mutation or may be a larger mutation for which the single base is a marker. Two separate reactions are conducted. In the first reaction, primer is annealed to target, and labeled (preferably ^{32}P) nucleic acids complementary to non-wild type (e.g. mutants indicative of disease) variants at the single base to be detected, and unlabeled dideoxy nucleic acids complementary to the wild-type base are combined. Primer extension is stopped the first time a wild-type (dideoxy) base is added to the primer. Presence of label in the extended primer is indicative of the presence of a mutation. A second tube, the positive control contains labeled nucleic acid complementary to the wild-type base in the presence of the primer. A DNA polymerase, such as Sequenase™ (Amersham), is used for primer extension. In a preferred embodiment, a thermostable polymerase, such as Taq or thermal sequenase is used to allow more efficient cycling. Once an extension reaction is completed, the first and second probes bound to target nucleic acids are dissociated by heating the reaction mixture above the melting temperature of the hybrids. The reaction mixture is then cooled below the melting temperature of the hybrids and additional primer is permitted to associate with target nucleic acids for another round of extension reactions. In a preferred embodiment, 10 to 50 cycles of extension reactions are conducted. In a most preferred embodiment, 30 cycles of extension reactions are conducted. After completion of all cycles, extension products are isolated and detected. In alternative embodiments, chain-terminating methods other than dideoxy nucleotides may be used. For example, chain termination occurs when no additional bases are available for incorporation at the next available nucleotide on the primer.

[0033] Methods of the invention are also useful for screening populations of patients in order to identify characteristics in population samples that are indicative of cancer or adenoma. For example, methods of the invention comprise high sensitivity, high specificity screening of populations of patients in order to correlate nucleic acid mutations present in a subset of patient samples with the presence of disease in those patients. Thus, methods of the invention comprise detecting genomic variations in patient samples, correlating those variations with confirmed disease, and using the variations associated with confirmed disease as a diagnostic screen for the disease in subsequent patient samples. Such methods preferably are performed on pooled samples, such as stool samples, from identified populations of patients (e.g., diseased, healthy). Such methods are preferably based upon variations in single nucleotide polymorphic loci. The sensitivity and specificity of detecting variants in those loci as a function of disease is determined. Those loci that predict disease at predefined levels of sensitivity and specificity are selected for use in screening assays for unknown patient samples.

[0034] The following examples provide specific exemplification of the concepts discussed above. The examples utilize a subset of mutational events that are shown to be predictive of disease. Other mutations are contemplated to function as high sensitivity, high specificity diagnostic or screening markers in assays of the invention. Moreover, the assays exemplified below are for purposes of illustration. The invention contemplates a variety of assays useful to screen patient samples for cancer or precancer as long as the assays provide a predetermined level of sensitivity and specificity of at least about 44% and at least about 85%, respectively.

EXAMPLE 1: Multiplex Screening of Stool Samples

[0035] An experiment was conducted to determine the effects of multiple mutation analysis alone on the sensitivity and specificity of cancer or precancer detection in the stool specimens described above. Fifteen mutations, all known or suspected to occur in colorectal cancer or precancer were used to screen the 40 patient samples. The following table catalogs the mutations that were assayed.

TABLE 1

Gene	Mutation
Kras, codon 13	Position 2, G to A
Kras, codon 12	Position 1, G to A
Kras, codon 12	Position 2, G to A
Apc, codon 1450	Position 1, C to T
Apc, codon 1378	Position 1, C to T
Apc, codon 1367	Position 1, C to T
Apc, codon 1309	Deletion of 5 base pairs
P53, codon 175	Position 2, G to tA
P53, codon 273	Position 1, C to T
P53, codon 273	Position 2, G to A
P53, codon 282	Position 1,C to T
P53, codon 245	Position 1, G to A
P53, codon 245	Position 2, G to A
P53, codon 248	Position 1, C to T
P53, codon 248	Position 2, G to A

[0036] According to methods of the invention, multiple mutation analysis is a preferred means for increasing the sensitivity and specificity of cancer or precancer detection. For example, there are cumulative benefits of combining the informativeness (see above) of a mutation at one allele (e.g., Kras, codon 13, position 2) with the informativeness of a second allele (e.g., Kras codon 12, position 1, or apc codon 1450, position 1) in order to increase the overall sensitivity and specificity of the assay. Accordingly, the benefits of one aspect of the invention are presented below.

[0037] Stool specimens were collected from 40 individuals who presented at the Mayo Clinic (Rochester, MN) with symptoms or history indicating that a colonoscopy should be performed. Each stool sample was frozen. Immediately after providing a stool sample, all individuals were given a colonoscopy in order to determine their disease status. Colonoscopy, an invasive test requiring sedation of the patient, has a sensitivity approaching 95%, and a specificity of nearly 100% for the diagnosis of colonic neoplasia. Based upon the colonoscopy results and subsequent histological analysis of biopsy samples taken during colonoscopy, individuals were placed into one of three groups: normal, cancer, and adenoma. An adenoma, or polyp, is considered clinically relevant if it has a diameter of 1 cm or greater. Thus, all individuals in the adenoma group had a polyp of at least 1 cm in diameter. Patients in the cancer group had tumors diagnosed as cancer, and the disease-free individuals were those for whom colonoscopy showed no sign of cancer or adenoma. Based upon the colonoscopy results, 21 patients were diagnosed with cancer, 9 patients were diagnosed with an adenoma greater than 1 cm, and 10 patients were free of cancer or adenoma.

[0038] Multiple mutation analysis was then performed, on a blinded basis (i.e., scientists performing the assays did not know the results of colonoscopy or histology), on each sample. Each frozen stool specimen, weighing from 7-33 grams, was thawed, homogenized in 500 mM Tris, 16 mM EDTA, and 10 mM NaCl, pH 9.0, at a volume to mass ratio of about 3:1. Samples were then rehomogenized in the same buffer to a final volume-to-mass ratio of 20:1, and spun in glass macro beads at 2356 x g. The supernatant was collected and treated with SDS and proteinase k. The DNA was then phenol-chloroform extracted and precipitated with alcohol. The precipitate was suspended in 10 mM Tris and 1 mM EDTA (1 x TE), pH 7.4. Finally, the DNA was treated with Rnase.

[0039] Human DNA was isolated from the precipitate by sequence-specific hybrid capture. Biotynilated probes against portions of the p53, K-ras, and apc genes were used. The K-ras probe was 5'GTGGAGTATTTGATAGTGTATTAAC-CTTATGTGTGAC 3' (SEQ ID NO: 1). There were two apc probes: apc-1309 was 5'TTCCAGCAGTGTCACAGCAC-CCTAGAACCAAATCCAG 3' (SEQ ID NO: 2), and apc-1378 was 5'CAGATAGCCCTGGACAAACAATGCCACGAAG-CAGAAG 3' (SEQ ID NO: 3). There were four probes against p53, the first (hybridizing to a portion of exon 5) was 5'TACTCCCCTGCCCTCAACAAGATGTTTTGCCAACTGG3' (SEQ ID NO:4), the second (hybridizing to a portion of exon 7) was 5'ATTTCTTCCATACTACTACCCATCGACCTCTCATC3' (SEQ ID NO: 5), the third, also hybridizing to a portion of exon 7 was 5'ATGAGGCCAGTGCGCCTTGGGGAGACCTGTGGCAAGC3' (SEQ ID NO: 6); and finally, a probe against exon 8 had the sequence 5'GAAAGGACAAGGGTGGTTGGGAGTAGATGGAGCCTGG3' (SEQ ID NO:

7). A 10 ul aliquot of each probe (20 pmol/capture) was added to a suspension containing 300 ul DNA in the presence of 310ul 6M GITC buffer for 2 hours at room temperature. Hybrid complexes were isolated using streptavidin-coated beads (Dynal). After washing, probe-bead complexes were suspended at 25° C for 1 hour in 0.1x TE buffer, pH7.4. The suspension was then heated for 4 minutes at 85° C, and the beads were removed.

[0040] Captured DNA was then amplified using PCR, essentially as described in U.S. Patent No. 4,683,202. Seven separate PCRs were run in duplicate using primers directed against Kras, apc, and p53. The primers listed below were used:

PCR-A FOR	Kras	5' GCG GTC CCA AAA GGG TCA GTC CTG CTG AAA ATG ACT GAA 3' (SEQ ID NO: 8)	39
PCR-A-REV		5' (Biotin)GCG GTC CCA AAA GGG TCA GTC ATG AAA ATG GTC AGA GAA A 3' (SEQ ID NO: 9)	40
PCR-B-FOR	APC-1309	5' GCG GTC GCT TTT GGG TCA GTT GTA GTT CAT TAT CAT CTT T 3' (SEQ ID NO: 10)	40
PCR-B-REV		5' (Biotin)GCG GTC GCT TTT GGG TCA GTC TTC GCT CAC AGG ATC TTC A 3' (SEQ ID NO: 11)	40
PCR-C-FOR	APC-1378	5' GCG GTC GCA AAA GGG ACA GTA GGC ACA AAG CTG TTG AAT 3' (SEQ ID NO: 12)	39
PCR-C-REV		5' (Biotin)GCG GTC GCA AAA GGG ACA GTT ATC AAG TGA ACT GAC AGA A 3' (SEQ ID NO: 13)	41
PCR-D-FOR	APC-1450	5' GCG GTC CCA AAA GGG TCA GTC ACC TCC ACC ACC TCC TCA A 3' (SEQ ID NO: 14)	40
PCR-D-REV		5' (Biotin)GCG GTC CCA AAA GGG TCA GTG TAT CAG CAT CTG GAA GAA 3' (SEQ ID NO: 15)	39
PCR-E-FOR	p53 Exon 5	5' GCG GTC CCA AAA GGG TCA GTC CAT CTA CAA GCA GTC A 3' (SEQ ID NO: 16)	37
PCR-E-REV		5' (Biotin)GCG GTC CCA AAA GGG TCA GTC AGA CCT AAG AGC AAT CA 3' (SEQ ID NO: 17)	38
PCR-F-FOR	p53 Exon 7	5' GCG GTC CCA AAA GGG TCA GAT ACC ACC ATC CAC TAC AA 3' (SEQ ID NO: 18)	38
PCR-F-REV		5' (Biotin)GCG GTC CCA AAA GGG TCA GAG TAT GGA AGA AAT CGG TAA 3' (SEQ ID NO: 19)	39
PCR-G-FOR	p53 Exon 8	5' GCG GTC CCT TTT GGG TCA CTC TGC CTC TTG CTT CTC TTT T 3' (SEQ ID NO: 20)	40
PCR-G-REV		5' (Biotin)GCG GTC CCT TTT GGG TCA CTC TTG TCC TGC TTG CTT ACC T 3' (SEQ ID NO: 21)	40

[0041] Samples were heated to 94° C for 5 minutes, and then 40 cycles were conducted between 94° C, 80° C, and 72° C (1 minute each), followed by one cycle at 72° C for 5 minutes.

[0042] The presence of the 15 mutations listed in Table 1 above was determined by cycling single base extension (cycling SBE), essentially as described above. Briefly, two reactions were run. In the first reaction, primer was hybridized adjacent the single base to be detected. ^{32}P-labeled nucleotide complementary to the expected mutant base and unlabeled dideoxy nucleotide complementary to the wild-type base were added. Primer was extended, and the presence of labeled product indicated a mutation was present in the sample. A second reaction was run as a positive control in which only labeled wild-type complement was added to the reaction mixture. Primer extension incorporating the labeled base assured that the reaction was running properly.

[0043] Primers used in SBE were as follows:

Name	Site	Codon/Position	Sequence
SBE-A1	Kras	k12p.1	5' AAC TTG TGG TAG TTG GAG CT 3' (SEQ ID NO: 22)
SBE-A2	Kras	k12p.2	5' ACT TGT GGT AGT TGG AGC TG 3' (SEQ ID NO: 23)
SBE-A3	Kras	k13p.2	5' TGT GGT AGT TGG AGC TGG TG 3' (SEQ ID NO: 24)
SBE-B1	APC-1309	1309 (_ 5)	5' AAA TAG CAG AAA TAA AA 3' (SEQ ID NO: 25)
SBE-C1	APC-1378	1367	5' CTC CCT CCA AAA GTG GTG CT 3' (SEQ ID NO: 26)
SBE-C2	APC-1378	1378	5' GTC CAC CTG TAC ACT ATG TT 3' (SEQ ID NO: 27)

(continued)

Name	Site	Codon/Position	Sequence
SBE-D1	APC-1450	1450	5' CTC AAA CAG CAC AAA CCA AG 3' (SEQ ID NO: 28)
SBE-E1	p53 Exon 5	175p.2	5' CAT GAC GGA GGT TGT GAG GC 3' (SEQ ID NO: 29)
SBE-F1	p53 Exon 7	245p.1	5' GTA ACA GTT CCT GCA TGG GC 3' (SEQ ID NO: 30)
SBE-F2	p53 Exon 7	245p.2	5' TAA CAG TTC CTG CAT GGG CG 3' (SEQ ID NO: 31)
SBE-F3	p53 Exon 7	248p.1	5' CCT GCA TGG GCG GCA TGA AC 3' (SEQ ID NO: 32)
SBE-F4	p53 Exon 7	248p.2	5' CTG CAT GGG CGG CAT GAA CC 3' (SEQ ID NO: 33)
SBE-G1	p53 Exon 8	273p.1	5' GAC GGA ACA GCT TTG AGG TG 3' (SEQ ID NO: 34)
SBE-G2	p53 Exon 8	273p.2	5' ACG GAA CAG CTT TGA GGT GC 3' (SEQ ID NO: 35)
SBE-G3	p53 Exon 8	282p.1	5' GTG CCT ATC CTG GGA GAG AC 3' (SEQ ID NO: 36)

[0044] Reactions were performed under standard denaturation, annealing, extension cycling for 30 cycles and visualized on a 15% denaturing polyacrylamide gel. Counts per minute (CPM) from each cycling reaction were entered into a Packard Instant Imager (wire chamber counter). Percent allele heterogeneity was determined as:

$$\frac{\text{cpm mutant / cpm wild type}}{\text{cmp 1\% mutant ctrl / cpm 1\% wild-type ctrl}}$$

[0045] A positive sample was defined as one in which at least one of the two replicates possessed a mutation with 1% heterogeneity for at least one of the single bases that was analyzed. Any sample in which at least one of the genes analyzed showed a mutation in duplicate was considered positive. The results are summarized below in Table 2. The total numbers of patients in the cancer/adenoma and normal groups are shown under the column "patient status".

TABLE 2

Patient Status	Lesions Detected By Colonoscopy	Lesions Detected By SBE	Sensitivity of SBE	Specificity of SBE
Cancer/Adenoma (30)	21/9	11/4	52%/44%	100%/100%
Normal (10)	0	0		

[0046] As shown in Table 2, multiple mutation analysis using SBE correctly identified 11 out of 21 cancerous lesions identified by colonoscopy for a sensitivity of 52%. Multiple mutation analysis revealed 4 out of 9 adenomas for a sensitivity of 44%. In both cases, multiple mutation analysis using SBE correctly identified all disease-free individuals, resulting in no false positives (specificity of 100%).

[0047] A fecal occult blood test, was run in parallel with the SBE test on all samples from patients diagnosed with an adenoma. Fecal occult blood testing failed to diagnosis any of the 9 adenoma-positive samples, and thus had a sensitivity of 0%. Accordingly, multiple mutation analysis has a far greater sensitivity and specificity than the most common non-invasive technique currently available (fecal occult blood).

Example 2 - Quantitative DNA Analysis

[0048] In this experiment, the same 40 samples described in Example 1 were independently analyzed for their overall content of DNA in order to determine if the amount of amplifiable DNA in stools produced by individuals with cancer or precancer was different than the amount of amplifiable DNA produced in stools from cancer-free individuals. Samples were analyzed "blind", and later correlated to colonoscopy results as described below.

[0049] Aliquots of the DNA obtained from the 40 patients described above in Example 1 were amplified using the primers described above. Each sample was amplified through 7 loci in duplicate (for a total of 14 amplifications for each locus). The products of PCR were placed on a 4% Nusieve (FMC Biochemical) gel (3% Nusieve, 1% agarose), and stained with ethidium bromide (0.5 ug/ml). The resulting amplified DNA was graded based upon the relative intensity of the stained gels. An "A" amplification produced the greatest cumulative intensity (and hence the greatest amount of DNA) after 40 cycles of PCR, "B" and "C" amplifications produced proportionately less gel intensity; and "F" amplifications

produced no or little intensity. Figure 1 shows exemplary A, B, C, and F amplifications. There is sufficient reproducibility in PCR to allow the skilled artisan to classify an amount of amplifiable nucleic acid based upon standards for healthy and cancer populations, or based upon inspection of the gel photograph in Figure 1. The assay only requires that one differentiate "A" amplifications from any of "B", "C", and "F" amplifications. An "A" amplification is one that has a band intensity of that in lane 1 of Figure 1, or to another gel band of similar intensity (e.g., lanes 4, 5, 6, or 7 of Figure 1). The markers to the right of Figure 1 show exemplary amounts of DNA giving rise to A, B, and C amplifications. Thus 200 pg of DNA (lane 12 in the Figure) results in an "A" amplification, and 100 pg (lane 13 in the Figure) results in a "B" amplification, and 50 pg (lane 14 in the Figure) results in a "C" amplification.

[0050] Any DNA sample that produced 9 "A" amplifications out of a possible 14 was graded as positive for cancer or adenoma. The results are shown in Table 3 below:

TABLE 3

Patient Status	Number of Patients Determined By Colonoscopy	Number of Patients Determined By Quantitative DNA Analysis ("A" Amps)	Sensitivity of Quantitative DNA Analysis	Specificity of Quantitative DNA Analysis
Cancer/Adenoma (30)	21/9	14/5	67%/56%	100%/100%
Cancer-free (10)	0	0		

[0051] As shown in Table 3, the amount of amplifiable DNA in stool is highly predictive of a patient's disease status. These results are consistent with the idea that patients with cancer or adenoma in the colon slough more cells (and therefore more DNA) onto the forming stool. Moreover, The DNA derived from cancer or adenoma cells is more intact than DNA derived from normal cells since cancer and adenoma cells have avoided apoptotic degradation of DNA.

Example 3 - Combined Multiple Mutation and Quantitative Analysis

[0052] The results obtained in Examples 1 and 2 above were combined to determine if further increases in sensitivity and specificity would be observed.

[0053] A positive sample in this experiment was one which produced an "A" amplification, and produced a positive multiple mutation result (under the criteria described in Example 1). Samples were prepared and analyzed on a "blind" basis (i.e., without knowing colonoscopy results a priori) as described above. The results are shown in Table 4 below:

TABLE 4

Patient Status	Number of Patients Diagnosed By Colonoscopy	Number of Patients Diagnosed By a Combination of Quant. And Multiple Mutation	Sensitivity	Specificity
Cancer/Adenoma	21/9	16/7	76%/78%	100%/100%

[0054] The results of the combined quantitative and multiple mutation analyses show a sensitivity of detection of 76% for cancer, and 78% for adenomas, each having a specificity of 100%. These results far exceed those of other non-invasive or minimally-invasive techniques (e.g., fecal occult blood testing which has a sensitivity of 0).

Example 4 - Diagnostic Assay Using Bat-26

[0055] The Bat-26 mismatch repair locus (shown in SEQ ID NO: 37) was next used to assess the same 40 samples described above. Deletions in Bat-26 have been associated with colorectal cancer or adenomas. Samples were prepared as described above. A primer was hybridized to the portion of the Bat-26 locus immediately upstream of the poly-A tract. Unlabeled deoxy thymidine, a mixture of labeled and unlabeled deoxycytosine, and unlabeled dideoxy adenine were added along with polymerase. The primer was extended through the poly-A region. The labeled and unlabelled cytosine was extended for the next three bases (nucleotides 222-224, all guanines in the intact sequence) such that label was incorporated into each extended primer. After the poly-A tract and the three guanines, there exist two thymidines in the intact sequence. Thus, the dideoxy adenosine stops primer extension by addition at the end of a primer that has been extended through the poly-A and triguanine regions. Strands were separated, and the length of the strands was observed

on a polyacrylamide gel to detect deletions in the poly-A tract. The results are presented below in Table 5:

TABLE 5

Patient Status	Diagnosis By Colonoscopy	Diagnosis By Bat-26 Detection	Sensitivity of Bat-26 Detection	Specificity of Bat-26 Detection
Cancer/Adenoma	21/9	4/0	19%/0%	100%/0%

[0056] As shown above, Bat-26 alone did not provide the high sensitivity achieved using multiple mutation or quantitation alone, but showed high sensitivity in comparison with other single locus detection assays. Moreover, as shown below, Bat-26 in combination with the other techniques described above produced an overall increase in sensitivity and specificity.

Example 5 - Cumulative Effects of Kras, Multiple Mutation, Quantitation, and BAT-26

[0057] The results obtained above for Kras, multiple mutation analysis, quantitation, and Bat-26 were combined to determine the cumulative effects of using combinations of those techniques in order to produce increased sensitivity and specificity in a non-invasive assay for cancer or precancer. The results are summarized below in Table 6:

TABLE 6

Assay Combination	Kras and Quantitation and BAT-26	Quantitation and BAT-26	Multiple Mutation and Quantitation and BAT-26
Sensitivity for Detection of Cancer/Adenoma	80%/56%	80%/56%	90%/78%
Specificity for Detection of Cancer/Adenoma	100%	100%	100%

[0058] As shown in the summary above, the combination of multiple mutation analysis, quantitative PCR, and Bat-26 produced a sensitivity approaching that of colonoscopy. A combination of multiple mutation analysis and quantitation alone also produces very high sensitivities. All assays resulted in a specificity of 100% (no false positive results), which is comparable to colonoscopy.

[0059] The foregoing experiments show that even a single high-sensitivity/high specificity non-invasive or minimally-invasive assay produces diagnostic results that are superior to non-invasive/minimally-invasive techniques of the art, and approach results observed with the recognized standard invasive diagnostic procedure (colonoscopy). Moreover, a non-invasive assay utilizing more than one high-sensitivity/high-specificity technique results in diagnostic accuracy approaching 100%. As such methods of the invention provide a significant improvement in the ability to perform accurate non-invasive diagnosis of cancer.

SEQUENCE LISTING

[0060]

<110> Lapidus, Stanley N
Shuber, Anthony P

<120> Methods for improving sensitivity and specificity of screening assays

<130> EXT-030PC

<140>
<141>

<150> US 09/277,016
<151> 1999-03-26

<160> 37

<170> PatentIn Ver. 2.0

<210> 1
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> K-ras probe

<400> 1
gtggagtatt tgatagtgta ttaaccttat gtgtgac 37

<210> 2
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> apc probe apc-1309

<400> 2
ttccagcagt gtcacagcac cctagaacca aatccag 37

<210> 3
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> apc probe apc-1378

<400> 3
cagatagccc tggacaaaca atgccacgaa gcagaag 37

<210> 4
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> p53 probe

<400> 4
tactcccctg ccctcaacaa gatgttttgc caactgg 37

<210> 5
<211> 35

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(35)
<223> p53 probe

<400> 5
atttcttcca tactactacc catcgacctc tcatc 35

<210> 6
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> p53 probe

<400> 6
atgaggccag tgcgccttgg ggagacctgt ggcaagc 37

<210> 7
<211> 37
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(37)
<223> p53 probe

<400> 7
gaaaggacaa gggtggttgg gagtagatgg agcctgg 37

<210> 8
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-A FOR (K-Ras)

<400> 8
gcggtcccaa aagggtcagt cctgctgaaa atgactgaa 39

<210> 9
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-A-REV (K-Ras)

<400> 9
gcggtcccaa aagggtcagt catgaaaatg gtcagagaaa 40

<210> 10
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-B-FOR (APC-1309)

<400> 10
gcggtcgctt ttgggtcagt tgtagttcat tatcatcttt 40

<210> 11
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-B-REV (APC-1309)

<400> 11
gcggtcgctt ttgggtcagt cttcgctcac aggatcttca 40

<210> 12
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-C-FOR (APC-1378)

<400> 12
gcggtcgcaa aagggacagt aggcacaaag ctgttgaat 39

<210> 13
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-C-REV (APC-1378)

<400> 13
gcggtcgcaa aagggacagt tatcaagtga actgacagaa 40

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

<220>
<223> Description of Artificial Sequence:PCR-D-FOR (APC-1450)

<400> 14
gcggtcccaa aagggtcagt cacctccacc acctcctcaa 40

<210> 15
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-D-REV (APC-1450)

<400> 15
gcggtcccaa aagggtcagt gtatcagcat ctggaagaa 39

<210> 16
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-E-FOR (p53 Exon 5)

<400> 16
gcggtcccaa aagggtcagt ccatctacaa gcagtca 37

<210> 17
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-E-REV (p53 Exon 5)

<400> 17
gcggtcccaa aagggtcagt cagacctaag agcaatca 38

<210> 18
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-F-FOR (p53 Exon 7)

<400> 18
gcggtcccaa aagggtcaga taccaccatc cactacaa 38

<210> 19
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-F-REV (p53 Exon 7)

<400> 19
gcggtcccaa aagggtcaga gtatggaaga aatcggtaa 39

<210> 20
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-G-FOR (p53 Exon 8)

<400> 20
gcggtccctt ttgggtcact ctgcctcttg cttctctttt 40

<210> 21
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR-G-REV (p53 Exon 8)

<400> 21
gcggtccctt ttgggtcact cttgtcctgc ttgcttacct 40

<210> 22
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-A1

<400> 22
aacttgtggt agttggagct 20

<210> 23
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-A2

<400> 23
acttgtggta gttggagctg 20

<210> 24
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-A3

<400> 24
tgtggtagtt ggagctggtg 20

<210> 25
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(17)
<223> SBE-B1

<400> 25
aaatagcaga aataaaa 17

<210> 26
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1) .. (20)
<223> SBE-C1

<400> 26
ctccctccaa aagtggtgct 20

<210> 27
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-C2

<400> 27
gtccacctgt acactatgtt 20

<210> 28
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-D1

<400> 28
ctcaaacagc acaaaccaag 20

<210> 29
<211> 20
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-E1

<400> 29
catgacggag gttgtgaggc 20

<210> 30
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-F1

<400> 30
gtaacagttc ctgcatgggc 20

<210> 31
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-F2

<400> 31
taacagttcc tgcatgggcg 20

<210> 32
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-F3

<400> 32
cctgcatggg cggcatgaac 20

<210> 33
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-F4

<400> 33
ctgcatgggc ggcatgaacc 20

<210> 34
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-G1

<400> 34
gacggaacag ctttgaggtg 20

<210> 35
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-G2

<400> 35
acggaacagc tttgaggtgc 20

<210> 36
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(20)
<223> SBE-G3

<400> 36
gtgcctatcc tgggagagac 20

<210> 37
<211> 314
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(314)
<223> BAT-26, wherein each "n" corresponds to a nucleotide of unknown identity

<400> 37

```
ccagtggtat agaaatcttc gatttttaaa ttcttaattt taggttgcag tttcatcact 60
gtctgcggta atcaagtttt tagaactctt atcagatgat tccaactttg gacagtttga 120
actgactact tttgacttca gccagtatat gaaattggat attgcagcag tcagagccct 180
taacctttt caggtaaaaa aaaaaaaaaa aaaaaaaaaa agggttaaaa atgttgattg 240
gttaannnnn nnngacagat agtgaagaag gcttagaaag gagctaaaag agttcgacat 300
caatattaga caag                                                  314
```

Claims

1. A method of detecting cancer or precancer, comprising screening a patient sample obtained non-invasively or minimally invasively for the presence of one or more nucleic acid indicia which are indicative of cancer or pre-cancer, **characterised in that** said nucleic acid indicia are selected from :

   (a) the length of the nucleic acids in the sample; and
   (b) the ratio of nucleic acids having a length greater than about 200 bp to nucleic acids having a length less than about 200 bp in the sample;

   wherein said length and said ratio differ between patients with cancer or pre-cancer and patients who do not have cancer or pre-cancer and wherein said sample is obtained from bodily excretions or from bodily fluids.

2. A method as claimed in claim 1, wherein said sample is obtained from the group consisting of stool, a homogenate of stool, pus, sputum, urine, blood, cerebrospinal fluid, lymph, semen and aspirate.

3. A method as claimed in claim 1 or claim 2, further comprising the step of screening said sample for one or more further nucleic acid indicia that are indicative of cancer or precancer, wherein said further nucleic acid indicia are selected from the group consisting of nucleic acid mutations, loss of heterozygosity, sequence length variations and variations in amounts of amplifiable nucleic acid.

4. A method as claimed in claim 3, comprising screening said sample for between three and twenty nucleic acid indicia, wherein said nucleic acid indicia are selected from the group consisting of nucleic acid mutations, loss of heterozygosity, sequence length variations, and variations in amounts of amplifiable nucleic acid.

5. A method as claimed in claim 3 or claim 4, wherein said screening involves two or more of the following : quantitative PCR, multiple mutation analysis, detection of loss of heterozygosity, hybrid capture of one or more mutant nucleic acid markers, oligo-ligation, amplification refractory mutation system, single-stranded conformational polymorphism detection, sequencing, mismatch detection and single base extension.

6. A method as claimed in claim 5, wherein said screening involves a quantitative polymerase chain reaction assay and an assay for a mutation in bat-26.

7. A method as claimed in claim 6, wherein said assay for a mutation in bat-26 comprises a primer extension assay to identify fragments of said bat-26.

8. A method as claimed in claim 5, wherein said screening involves an assay to detect loss of heterozygosity in at least a portion of a chromosomal arm.

9. A method as claimed in claim 8, wherein said assay to detect loss of heterozygosity is selected from the group consisting of enumerative LOH and single base extension.

10. A method as claimed in claim 8, wherein said chromosomal arm is selected from the group consisting of chromosome 17p, chromosome 5p, chromosome 8p, chromosome 1q, and chromosome 18q.

11. A method as claimed in any preceding claim, further comprising one or more of the following steps:

    (a) performing a nucleic acid amplification assay on said sample to detect an amount of amplifiable nucleic acid

in said sample and identifying as a positive screen a sample in which said amount exceeds a predetermined threshold;
(b) performing a multiple mutation assay on said sample and identifying as a positive screen a sample in which at least one mutation is identified.

12. A method as claimed in claim 11, wherein said mutation assay comprises a single base extension assay.

13. A method as claimed in claim 12, wherein said single base extension assay comprises the steps of :

(a) annealing an oligonucleotide primer to said sample under conditions that promote exact complementary hybridisation between said primer and a portion of a nucleic acid in said sample;
(b) extending said primer by a single base;
(c) separating said extended primer by a single base;
(d) identifying the base incorporated into said extended primer, thereby to identify said single base.

14. A method as claimed in claim 12, wherein said single base extension assay comprises the steps of :

(a) annealing an oligonucleotide primer to said sample under conditions that promote exact complementary hybridisation between said primer and a portion of a nucleic acid in said sample;
(b) exposing said sample to two different deoxynucleotides under conditions to promote primer extension;
(c) extending said primer until said primer can no longer be extended;
(d) separating said extended primer from said portion; and
(e) identifying the base incorporated into said extended primer, thereby to identify said single base.

15. The method of claim 14, wherein at least one of said deoxynucleotides is detectably labelled.

16. The method of claim 12, wherein said single base extension assay comprises identifying single nucleotides in each of a plurality of genomic loci.

17. The method of claim 16, wherein said genomic loci are selected from the group consisting of apc, Kras, p53 and bat-26.

18. A method as claimed in any preceding claim, wherein said precancer is an adenoma, or wherein said cancer is colorectal cancer.

Patentansprüche

1. Verfahren zum Nachweis von Krebs oder Krebsvorstufe, aufweisend Untersuchung einer Patientenprobe, welche nicht-invasiv oder minimal invasiv erhalten wurde, bezüglich des Vorhandenseins eines oder mehrerer Nukleinsäure-Indizien, welche auf Krebs oder Krebsvorstufe hinweisen, **dadurch gekennzeichnet, dass** die Nukleinsäure-Indizien gewählt sind von:

(a) der Länge der Nukleinsäuren in der Probe; und
(b) dem Verhältnis von Nukleinsäuren mit einer Länge von mehr als ca. 200 bp zu Nukleinsäuren mit einer Länge von weniger als ca. 200 bp in der Probe;

wobei sich die Länge und das Verhältnis bei Patienten mit Krebs oder Krebsvorstufe und Patienten, welche keinen Krebs oder Krebsvorstufe haben, unterscheiden und wobei die Probe aus Körperausscheidungen oder von Körperflüssigkeiten erhalten wird .

2. Verfahren nach Anspruch 1, bei welchem die Probe erhalten wird von der Gruppe bestehend aus Stuhl, einem Homogenat aus Stuhl, Eiter, Sputum, Urin, Blut, Cerebrospinalflüssigkeit, Lymphe, Samen und Aspirat.

3. Verfahren nach Anspruch 1 oder 2, ferner aufweisend den Schritt des Untersuchens der Probe bezüglich einer oder mehrerer weiterer Nukleinsäure-Indizien, welche auf Krebs oder Krebsvorstufe hinweisen, wobei die weiteren Nukleinsäure-Indizien gewählt sind von der Gruppe bestehend aus Nukleinsäuremutationen, Verlust an Heterozygosität, Sequenzlängenvariationen und Variationen der Mengen von amplifizierbaren Nukleinsäuren.

4. Verfahren nach Anspruch 3, aufweisend Untersuchen der Probe bezüglich zwischen drei und zwanzig Nukleinsäure-Indizien, wobei die Nukleinsäure-Indizien gewählt sind von der Gruppe bestehend aus Nukleinsäuremutationen, Verlust an Heterozygosität, Sequenzlängenvariationen und Variationen der Menge an amplifizierbaren Nukleinsäuren.

5. Verfahren nach Anspruch 3 oder 4, bei welchem das Untersuchen zwei oder mehr der Folgenden umfasst: quantitative PCR, multiple Mutationsanalyse, Nachweis des Verlusts an Heterozygosität, Hybridisierungsbindung eines oder mehrerer mutierter Nukleinsäuremarker, Oligo-Ligation, Amplification Refractory Mutation System, Nachweis von Einzelstrang-Konformationspolymorphismen, Sequenzieren, Nachweis von Mismatch und Einzelbasenextension.

6. Verfahren nach Anspruch 5, wobei das Untersuchen eine quantitative Polymerasekettenreaktion-Untersuchung und eine Untersuchung bezüglich einer Mutation in bat-26 umfasst.

7. Verfahren nach Anspruch 6, bei welchem die Untersuchung bezüglich einer Mutation in bat-26 eine Primer-Extensionsuntersuchung umfasst, um Fragmente von dem bat-26 zu identifizieren.

8. Verfahren nach Anspruch 5, bei welchem das Untersuchen eine Untersuchung zum Nachweis des Verlusts an Heterozygosität in mindestens einem Bereich eines Chromosomenarms umfasst.

9. Verfahren nach Anspruch 8, bei welchem die Untersuchung zum Nachweisen des Verlusts an Heterozygosität gewählt ist aus der Gruppe bestehend aus Zählen von Verlust an Heterozygosität und Einzelbasenextension.

10. Verfahren nach Anspruch 8, bei welchem der chromosomale Arm gewählt ist aus der Gruppe bestehend aus Chromosom 17p, Chromosom 5p, Chromosom 8p, Chromosom 1q und Chromosom 18q.

11. Verfahren nach einem der vorangehenden Ansprüche, ferner aufweisend einen oder mehrere der folgenden Schritte:

(a) Durchführen einer Nukleinsäureamplifikationsuntersuchung mit der Probe zum Nachweis einer Menge an amplifizierbarer Nukleinsäure in der Probe und Identifizieren als positives Ergebnis eine Probe, in welcher die Menge eine vorher bestimmte Schwelle überschreitet;
(b) Durchführen einer multiplen Mutationsuntersuchung mit der Probe und Identifizieren als positives Ergebnis eine Probe, in welcher mindestens eine Mutation identifiziert wird.

12. Verfahren nach Anspruch 11, wobei die Mutationsuntersuchung eine Einzelbasenextensionsuntersuchung umfasst.

13. Verfahren nach Anspruch 12, bei welchem die Einzelbasenextensionsuntersuchung die folgenden Schritte aufweist:

(a) Binden eines Oligonukleotid-Primers an die Probe unter Bedingungen, welche exakte komplementäre Hybridisierung zwischen dem Primer und einem Bereich einer Nukleinsäure in der Probe begünstigen;
(b) Verlängern des Primers um eine einzelne Base;
(c) Separieren des um eine Base verlängerten Primers;
(d) Identifizieren der in den verlängerten Primer inkorporierten Base, um somit die einzelne Base zu identifizieren.

14. Verfahren nach Anspruch 12, bei welchem die Einzelbasenextensionsuntersuchung die folgenden Schritte aufweist:

(a) Binden eines Oligonukleotid-Primers an die Probe unter Bedingungen, welche die exakte komplementäre Hybridisierung zwischen dem Primer und einem Bereich einer Nukleinsäure in der Probe begünstigen;
(b) Aussetzen der Probe gegenüber zwei unterschiedlichen Deoxynukleotiden unter Bedingungen, bei welchen Primerextension begünstigt ist;
(c) Verlängern des Primers, bis der Primer nicht mehr verlängert werden kann;
(d) Separieren des verlängerten Primers von dem Bereich; und
(e) Identifizieren der in den verlängerten Primer inkorporierten Base, um somit die Einzelbase zu identifizieren.

15. Verfahren nach Anspruch 14, bei welchem mindestens eines der Deoxynukleotide nachweisbar markiert ist.

16. Verfahren nach Anspruch 12, bei welchem die Einzelbasenextensionsuntersuchung die Identifizierung einzelner Basen in jedem einer Mehrzahl von genomischen Loci aufweist.

17. Verfahren nach Anspruch 16, bei welchem die genomischen Loci gewählt sind aus der Gruppe bestehend aus apc, Kras, p53 und bat-26.

18. Verfahren nach einem der vorangehenden Ansprüche, bei welchem die Krebsvorstufe ein Adenom ist, oder bei welchem der Krebs Colorectal-Krebs ist.

Revendications

1. Procédé pour détecter un cancer ou un précancer, comprenant le criblage d'un échantillon de patient, obtenu de manière non invasive ou minimalement invasive, pour détecter la présence d'un ou plusieurs indices d'acide nucléique qui sont indicatifs d'un cancer ou d'un précancer, **caractérisé en ce que** lesdits indices d'acide nucléique sont choisis parmi :

    (a) la longueur des acides nucléiques dans l'échantillon ; et
    (b) le rapport des acides nucléiques ayant une longueur supérieure à environ 200 pb sur les acides nucléiques ayant une longueur inférieure à environ 200 pb dans l'échantillon ;

    dans lequel ladite longueur et ledit rapport diffèrent entre des patients cancéreux ou précancéreux et des patients non cancéreux ou non précancéreux, et dans lequel ledit échantillon est obtenu à partir d'excrétions corporelles ou de fluides corporels.

2. Procédé selon la revendication 1, dans lequel ledit échantillon est obtenu à partir du groupe constitué par les selles, un homogénat de selles, le pus, le crachat, l'urine, le sang, le liquide cérébrospinal, la lymphe, le sperme et un aspirat.

3. Procédé selon la revendication 1 ou la revendication 2, comprenant en outre l'étape consistant à cribler ledit échantillon pour détecter un ou plusieurs autres indices d'acide nucléique qui sont indicatifs d'un cancer ou d'un précancer, dans lequel lesdits autres indices d'acide nucléique sont choisis dans le groupe constitué par les mutations d'acide nucléique, la perte de caractère hétérozygote, les variations de longueur de séquence et les variations de quantités d'acide nucléique amplifiable.

4. Procédé selon la revendication 3, comprenant le criblage dudit échantillon pour détecter trois à vingt indices d'acide nucléique, dans lequel lesdits indices d'acide nucléique sont choisis dans le groupe constitué par les mutations d'acide nucléique, la perte de caractère hétérozygote, les variations de longueur de séquence, et les variations de quantités d'acide nucléique amplifiable.

5. Procédé selon la revendication 3 ou la revendication 4, dans lequel ledit criblage met en jeu au moins deux des suivants : PCR quantitative, analyse de mutations multiples, détection de perte de caractère hétérozygote, capture par hybridation d'un ou plusieurs marqueurs d'acides nucléiques mutants, oligoligature, système de mutation réfractaire à l'amplification, détection de polymorphisme de conformation monocaténaire, séquençage, détection de mésappariement, et extension de base unique.

6. Procédé selon la revendication 5, dans lequel ledit criblage met en jeu un essai quantitatif d'amplification en chaîne par polymérase et un essai pour une mutation dans bat-26.

7. Procédé selon la revendication 6, dans lequel ledit essai pour une mutation dans bat-26 comprend un essai d'extension d'amorce pour identifier des fragments dudit bat-26.

8. Procédé selon la revendication 5, dans lequel ledit criblage met en jeu un essai pour détecter une perte de caractère hétérozygote dans au moins une partie d'un bras chromosomique.

9. Procédé selon la revendication 8, dans lequel ledit essai pour détecter une perte de caractère hétérozygote est choisi dans le groupe constitué par une LOH énumérative et une extension de base unique.

10. Procédé selon la revendication 8, dans lequel ledit bras chromosomique est choisi dans le groupe constitué par le chromosome 17p, le chromosome 5p, le chromosome 8p, le chromosome 1q, et le chromosome 18q.

11. Procédé selon l'une quelconque des revendications précédentes, comprenant en outre une ou plusieurs des étapes

suivantes :

(a) réalisation d'un essai d'amplification d'acide nucléique sur ledit échantillon pour détecter une quantité d'acide nucléique amplifiable dans ledit échantillon et identification, en tant que criblage positif, d'un échantillon dans lequel ladite quantité dépasse un seuil prédéterminé ;
(b) réalisation d'un essai de mutations multiples sur ledit échantillon et identification, en tant que criblage positif, d'un échantillon dans lequel au moins une mutation est identifiée.

12. Procédé selon la revendication 11, dans lequel ledit essai de mutation comprend un essai d'extension de base unique.

13. Procédé selon la revendication 12, dans lequel ledit essai d'extension de base unique comprend les étapes consistant à :

(a) anneler une amorce oligonucléotidique audit échantillon dans des conditions qui favorisent une hybridation complémentaire exacte entre ladite amorce et une partie d'un acide nucléique dans ledit échantillon ;
(b) étendre ladite amorce d'une seule base ;
(c) séparer ladite amorce étendue d'une seule base ;
(d) identifier la base incorporée dans ladite amorce étendue, de façon à identifier ainsi ladite base unique.

14. Procédé selon la revendication 12, dans lequel ledit essai d'extension de base unique comprend les étapes consistant à :

(a) anneler une amorce oligonucléotidique audit échantillon dans des conditions qui favorisent une hybridation complémentaire exacte entre ladite amorce et une partie d'un acide nucléique dans ledit échantillon ;
(b) exposer ledit échantillon à deux désoxynucléotides différents dans des conditions favorisant l'extension d'amorce ;
(c) étendre ladite amorce jusqu'à ce que ladite amorce ne puisse plus être étendue ;
(d) séparer ladite amorce étendue d'avec ladite partie ; et
(e) identifier la base incorporée dans ladite amorce étendue, de façon à identifier ainsi ladite base unique.

15. Procédé selon la revendication 14, dans lequel au moins l'un desdits désoxynucléotides est marqué de façon détectable.

16. Procédé selon la revendication 12, dans lequel ledit essai d'extension de base unique comprend l'identification de nucléotides uniques dans chacun d'une pluralité de sites génomiques.

17. Procédé selon la revendication 16, dans lequel lesdits sites génomiques sont choisis dans le groupe constitué par apc, Kras, p53 et bat-26.

18. Procédé selon l'une quelconque des revendications précédentes, dans lequel ledit précancer est un adénome, ou dans lequel ledit cancer est un cancer colorectal.

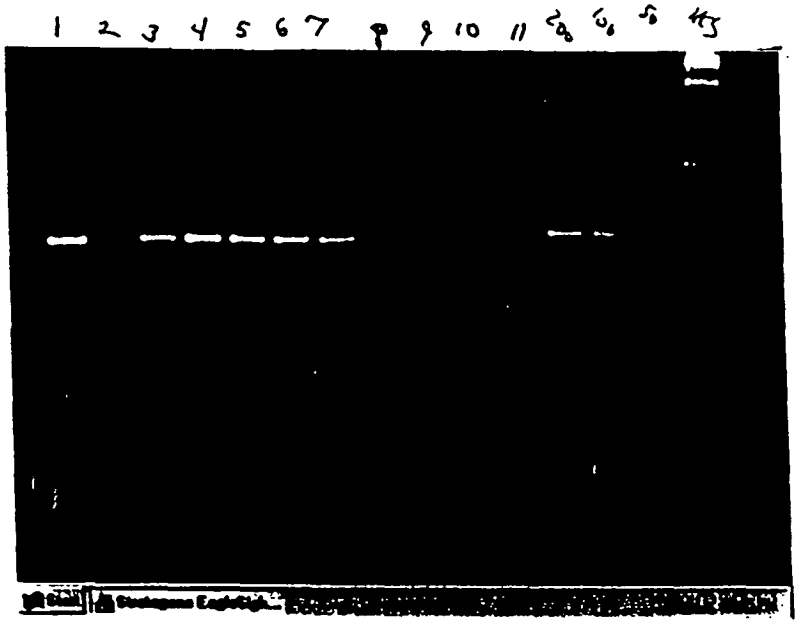

Fig. 1

专利名称(译)	提高癌症和癌前病变筛选试验的灵敏度和特异性的方法		
公开(公告)号	EP1185693B1	公开(公告)日	2006-08-23
申请号	EP2000918364	申请日	2000-03-24
[标]申请(专利权)人(译)	精密科学公司		
申请(专利权)人(译)	EXACT科学公司		
当前申请(专利权)人(译)	EXACT科学公司		
[标]发明人	LAPIDUS STANLEY N SHUBER ANTHONY P		
发明人	LAPIDUS, STANLEY, N. SHUBER, ANTHONY, P.		
IPC分类号	C12Q1/68 G01N33/53 C12N15/09 C12Q1/6886 G01N27/447 G01N33/574		
CPC分类号	C12Q1/6886 C12Q2600/156 Y10T436/143333		
优先权	09/277016 1999-03-26 US		
其他公开文献	EP1185693A2		
外部链接	Espacenet		

摘要(译)

本发明的方法包括指示癌症或癌前病变的标志物的测定。通过非侵入性或微创方法对从患者获得的样品进行本发明的测定。本发明提供具有高灵敏度和高检测特异性的癌症或癌前病变的核酸标记。