US 20110251097A1

(19) **United States**

(12) **Patent Application Publication**

(10) **Pub. No.: US 2011/0251097 A1**

(43) **Pub. Date: Oct. 13, 2011**

Song et al.

(54) **DIAGNOSTIC KIT OF COLON CANCER USING COLON CANCER RELATED MARKER AND DIAGNOSTIC METHOD THEREOF**

(75) Inventors: **Eun Young Song**, Seoul (KR); **Hee Gu Lee**, Daejeon (KR); **Young Il Yeom**, Daejeon (KR); **Jae Wha Kim**, Daejeon (KR); **Na Young Ji**, Gyeonggi-do (KR); **Kyung-Sook Chung**, Daejeon (KR); **Misun Won**, Daejeon (KR); **Seon-Young Kim**, Daejeon (KR); **Joo Heon Kim**, Daejeon (KR); **Young Ho Kim**, Seoul (KR); **Ho Kyung Chun**, Seoul (KR)

(73) Assignee: **KOREA RESEARCH INSTITUTE OF BIOSCIENCE AND BIOTECHNOLOGY**, Daejeon (KR)

(21) Appl. No.: **13/126,500**

(22) PCT Filed: **Apr. 20, 2009**

(86) PCT No.: **PCT/KR2009/002039**

§ 371 (c)(1),
(2), (4) Date: **Apr. 28, 2011**

(30) **Foreign Application Priority Data**

Oct. 22, 2008 (KR) 10-2008-0103387

Publication Classification

(51) **Int. Cl.**

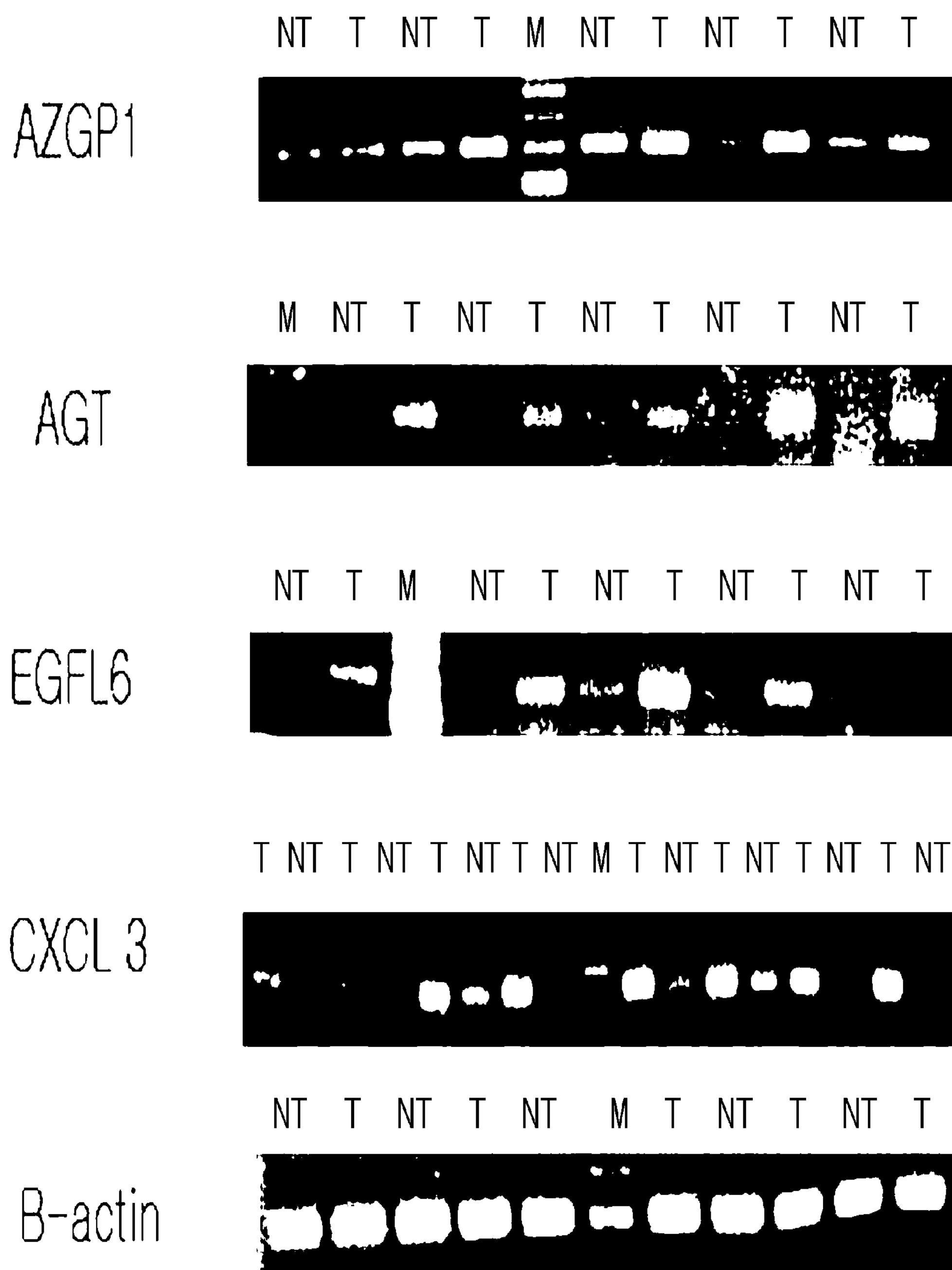
C40B 30/04	(2006.01)
G01N 33/53	(2006.01)
G01N 33/566	(2006.01)
C12Q 1/68	(2006.01)

(52) **U.S. Cl.** **506/9**; 435/6.14; 435/6.12; 436/501; 435/7.92; 435/7.1

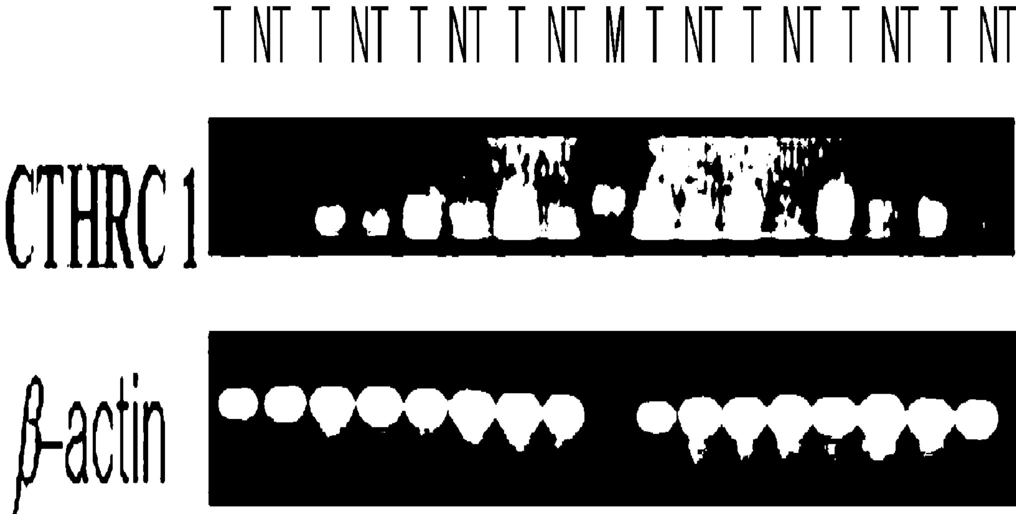
(57) **ABSTRACT**

The present invention relates to a composition for diagnosing colon cancer. The composition comprises at least one marker for measuring an mRNA or protein expression level of at least one gene specific for colon cancer. It can screen the genes which are overexpressed specifically only in colon cancer tissues or blood. The present invention can quantitatively analyze both the mRNA expression levels of the genes and the expression levels of the proteins encoded by the gene at the same time, thereby diagnosing colon caner of an early stage with a high level of reliability.

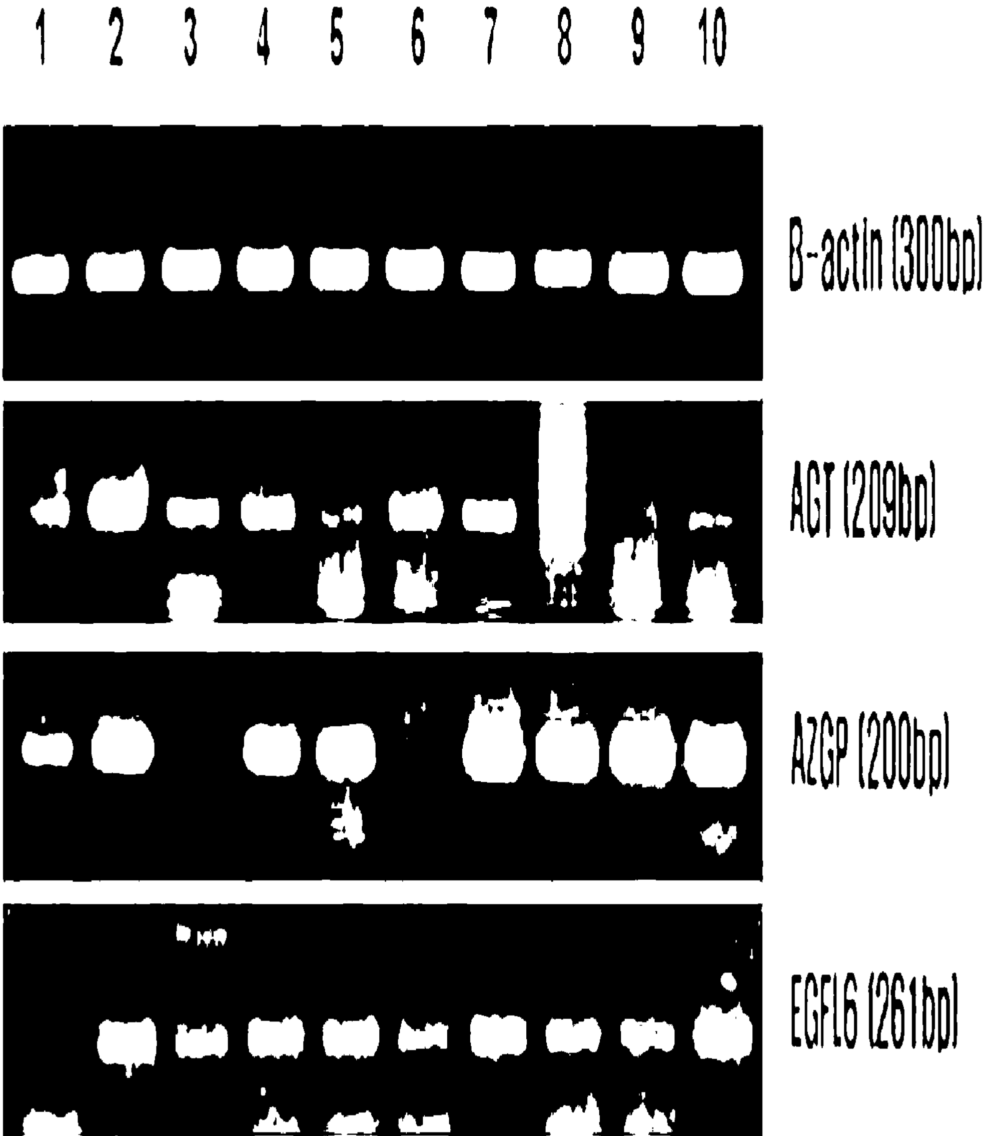
[Fig. 1]



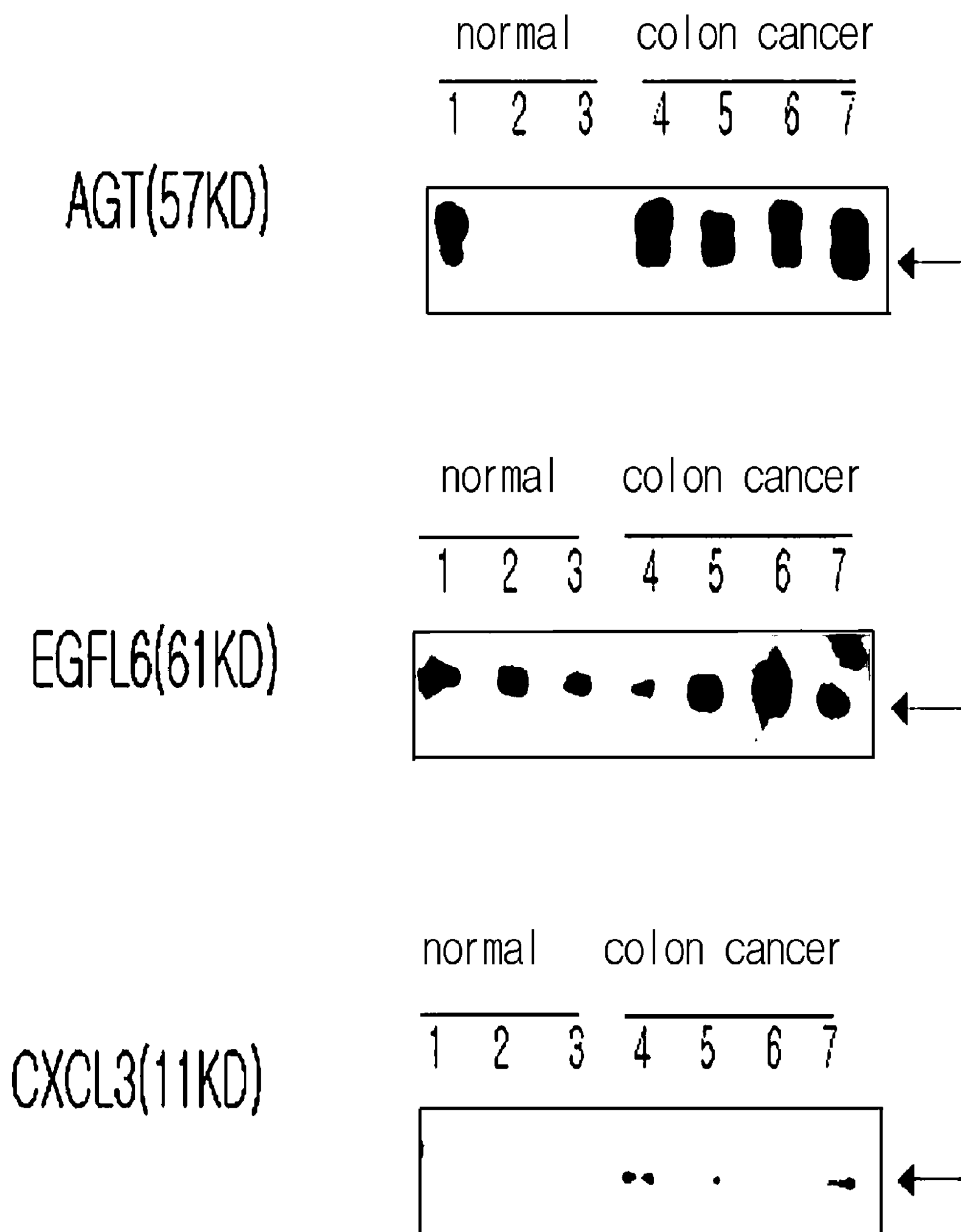
[Fig. 2]



[Fig. 3]

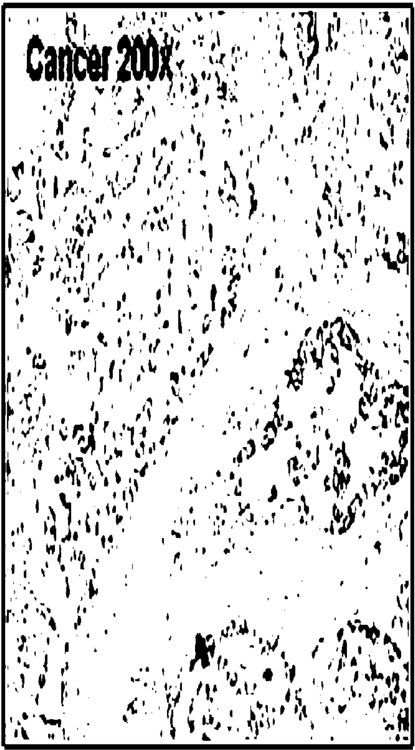


[Fig. 4]



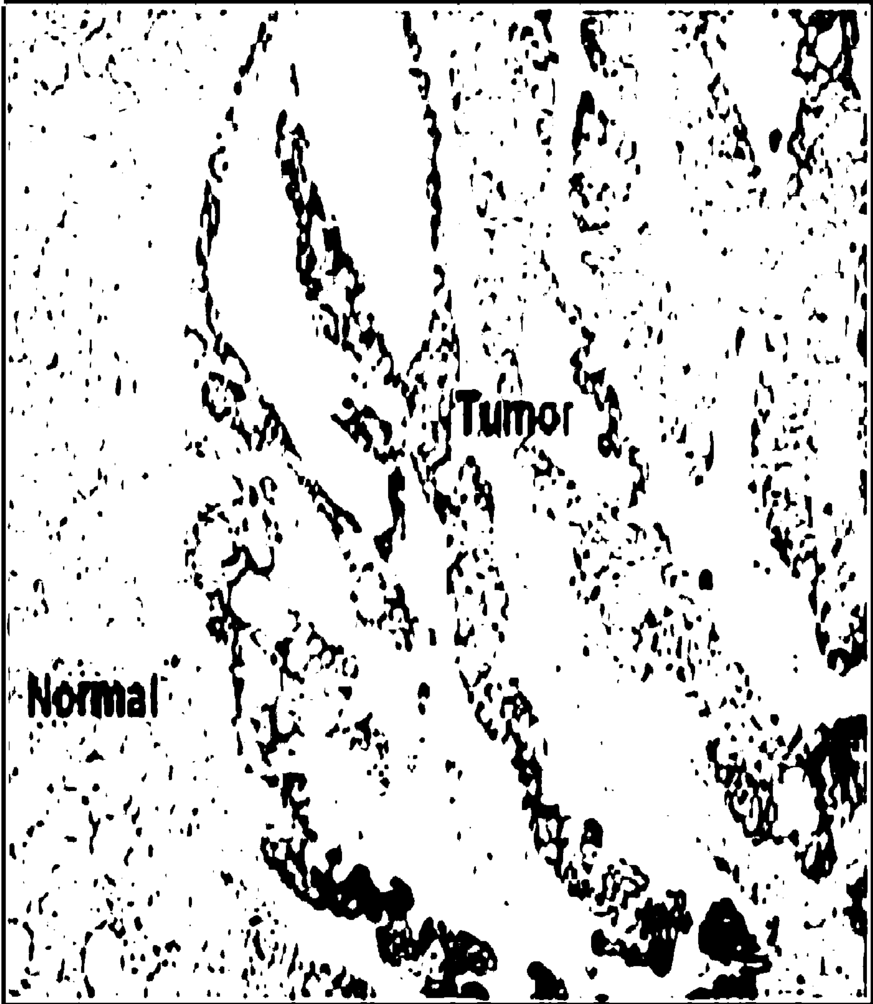
[Fig. 5]

EFGL6



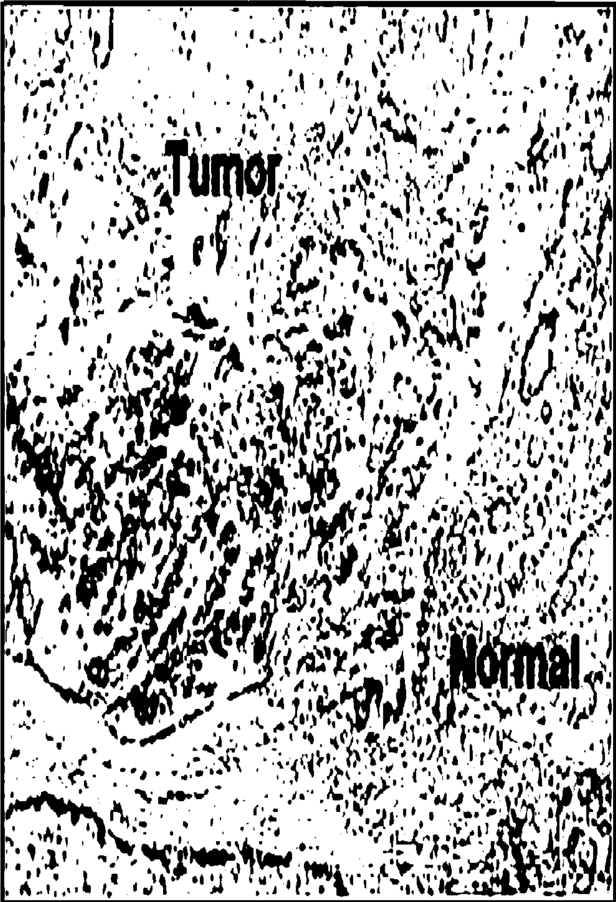
[Fig. 6]

CTHRC1

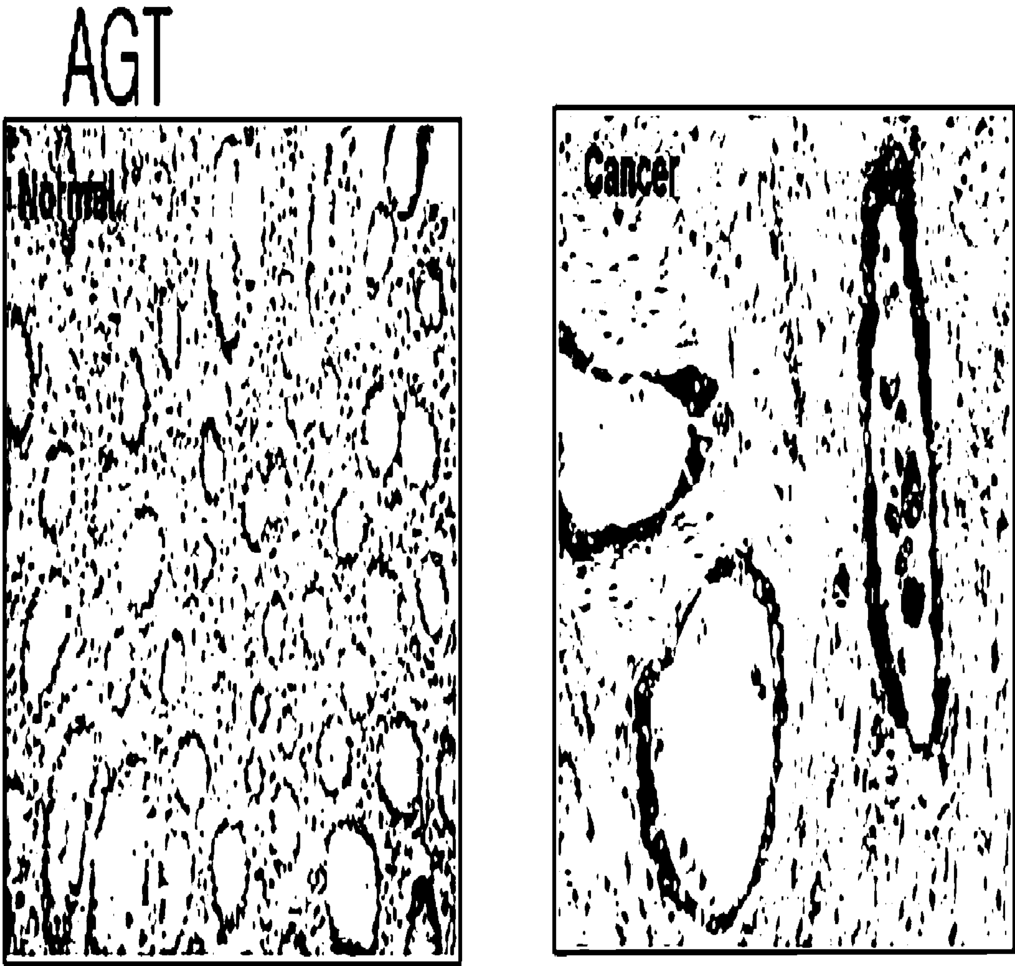


[Fig. 7]

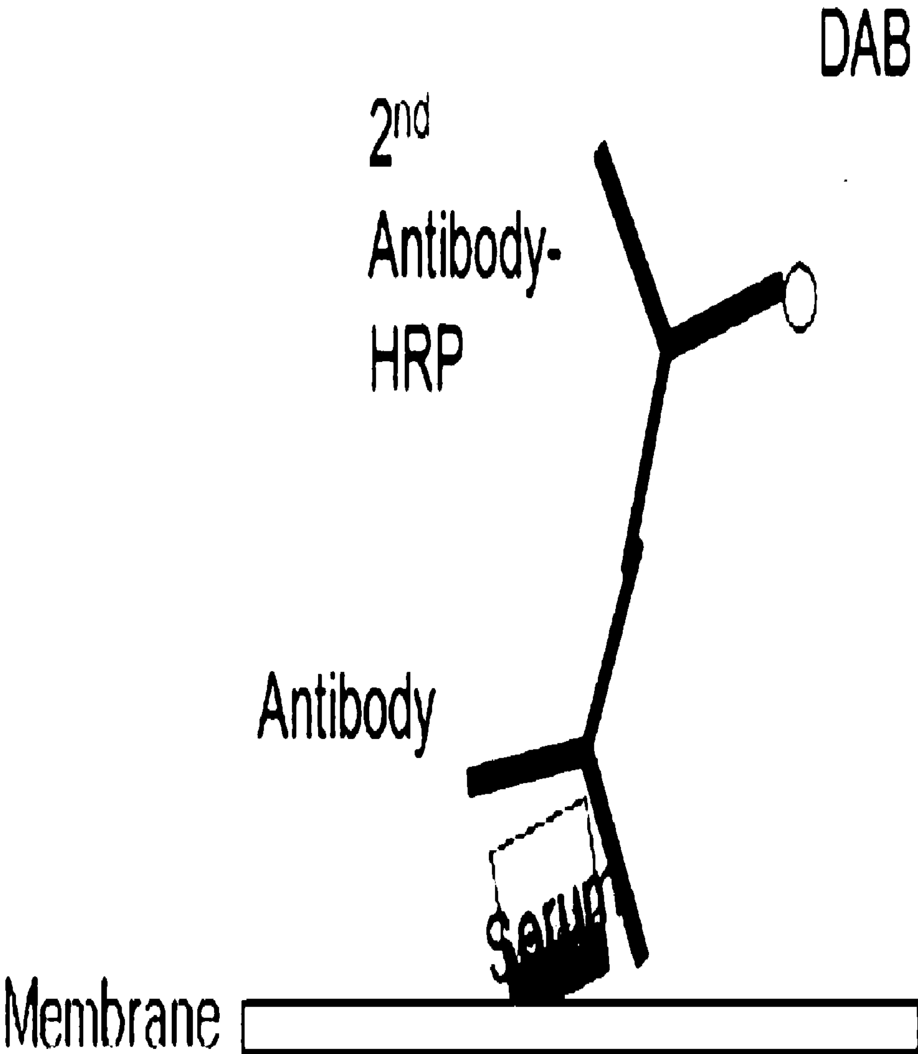
CXCL-3



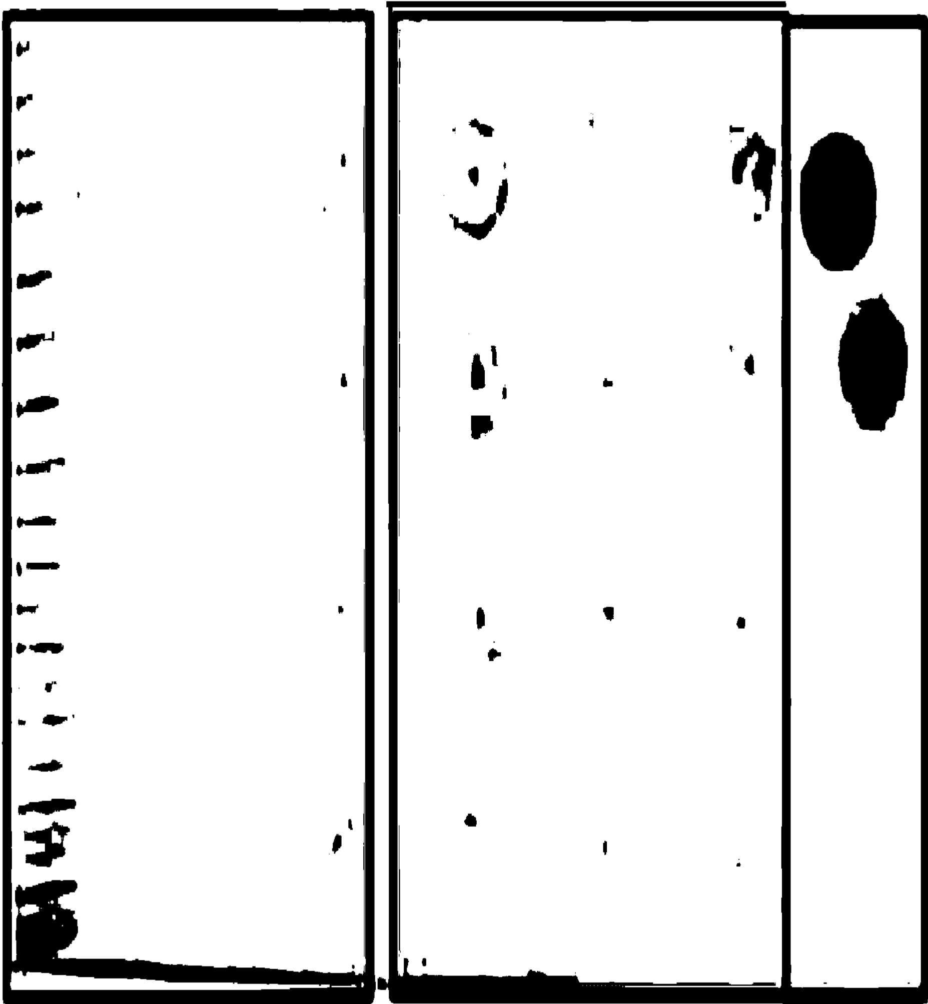
[Fig. 8]



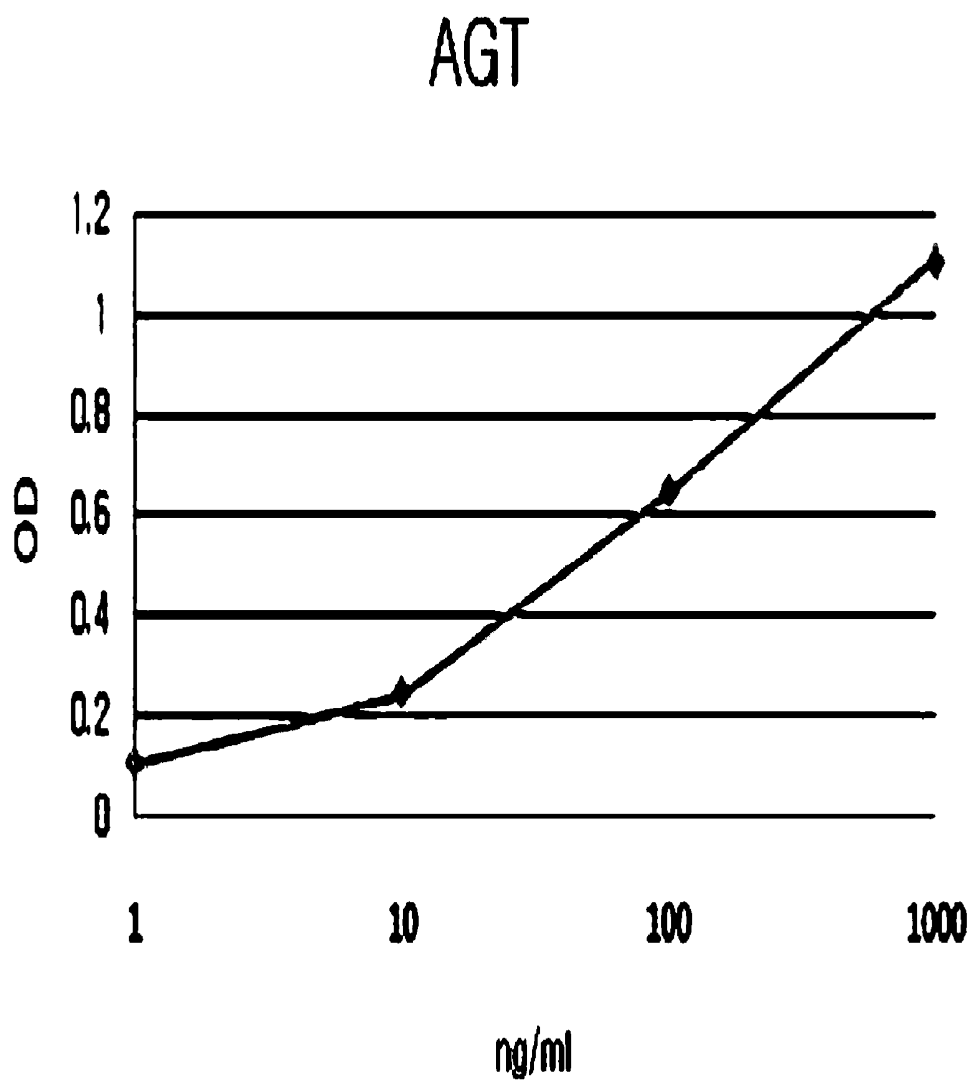
[Fig. 9]



[Fig. 10]

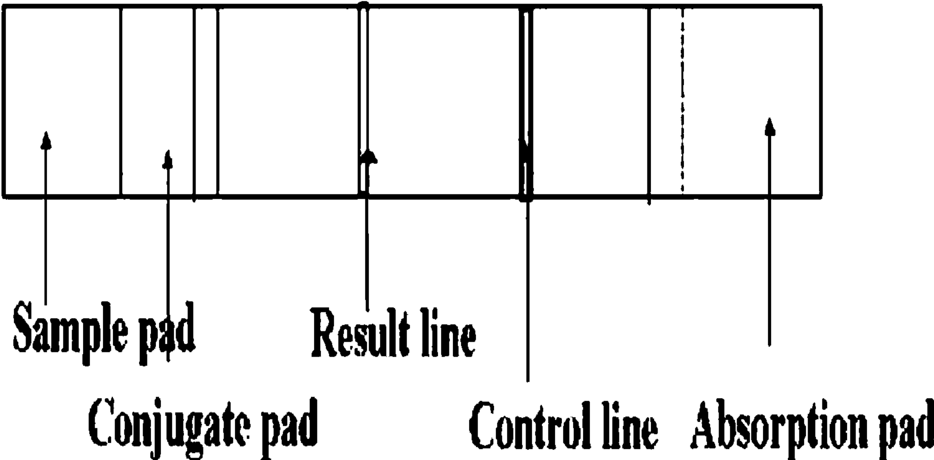


[Fig. 11]

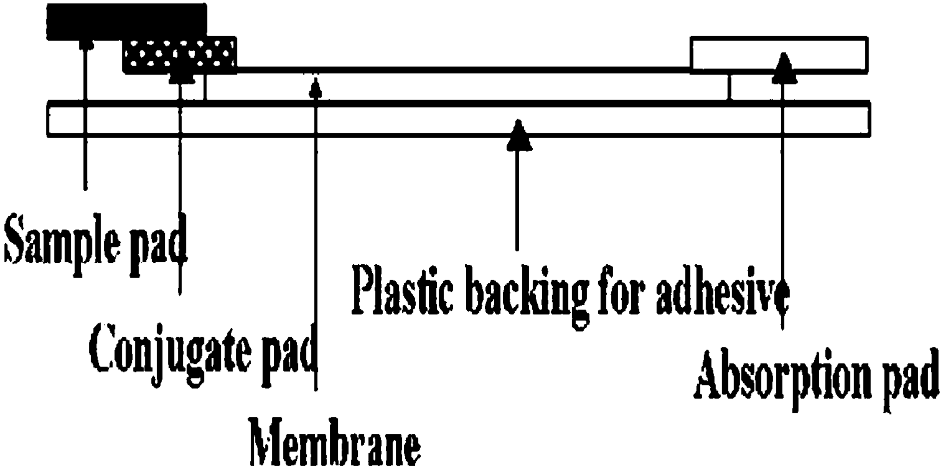


[Fig. 12]

A.



B.



DIAGNOSTIC KIT OF COLON CANCER USING COLON CANCER RELATED MARKER AND DIAGNOSTIC METHOD THEREOF

TECHNICAL FIELD

[0001] The present invention relates to a diagnostic kit of colon cancer using a colon cancer-related marker and a method of yielding information necessary for the diagnosis of colon cancer. More particularly, the present invention relates to a diagnostic composition for colon cancer, comprising at least one marker for measuring an mRNA or protein expression level of at least one gene specific for colon cancer, and a method of yielding information necessary for the diagnosis of colon cancer using the same.

BACKGROUND ART

[0002] The large intestine is the last part of the digestive system in the body in which the food ingested through the mouth is digested and absorbed and even excess food is stayed. The main function of the large intestine is to transport waste out of the body and to absorb water from the waste before it leaves. In addition, the large intestine houses over 700 species of bacteria that perform a variety of functions. The large intestine is about 2 m long and consists of the colon, rectum and the anus. It is said that cancer can occur in the body where mucous membrane exits. However, the sigmoid colon and the rectum are most vulnerable to cancer.

[0003] In Korea, the incidence of colon cancer has been dramatically increasing. Moreover, it is the fourth leading cause of cancer-related death among men in Korea, followed by stomach cancer, lung cancer and liver cancer. It is also shown that similar rates of cancer mortality are found in women and the frequency of colon cancer is higher in men than in women. Most cases occur in patients in their 50s, followed by those in their 60s. Furthermore, the age of the greatest incidence of colon cancer in Korea is likely to be 10 years lower than that in the Western world such as the U.S. and Europe. The incidence frequency of colon cancer accounts for 5%-10% in people in their 30s. In addition, colon cancer is likely to occur in the young generation and it is also found mostly in people who have a family history of colon cancer. In fact, the incidence of colon cancer is caused not by heredity but mostly environmental factors. More specifically, the westernization of the diet and particularly excess intake of animal oil and proteins play a greater role in causing colon cancer. Meanwhile, only 5% of colon cancer cases are attributed to hereditary predisposition. In consequence, people with a high risk of developing colon cancer are those who 1) have been affected by colon polyp, 2) have a family history of colon cancer, 3) suffer from ulcerative colitis for a long period of time, or 4) are attacked by incurable anal fistula.

[0004] Typically, colon cancer can be classified by the Dukes staging system or the UICC staging system. The systems for staging colon cancer are not determined not by the size of tumor, but largely by the extent of local invasion, and the presence of distant metastasis.

[0005] Standards of the Dukes classification and the UICC classification are given in Tables 1 and 2, respectively.

TABLE 1

Description of the Dukes Classification

Stages	Post-operation5-Year Survival Rate	Pathological Conditions
Dukes A	90%	Tumour confined to the intestinal wall
Dukes B	60 80%	Tumour invading through the intestinal wall, but without lymph node involvement
Dukes C	20-50%	With lymph node(s) involvement
Dukes D	Less than 20%	With distant metastasis to the peritoneum, the liver, the lungs, etc.

TABLE 2

Characteristics of UICC Stage Classification

Stages	Pathological Conditions
0	Limited to mucosa
1	Extending into *muscularis propria* but not penetrating through it
2	Penetrating through *muscularis propria*, but not to adjacent organs
3	Penetrating into adjacent organs. Nodes involved
4	Distant metastatic spread into, e.g., the peritoneum, the liver, the lungs, etc.

[0006] Considering that there is a slight difference between these two classifications, it is currently recognized that Dukes A corresponds to UICC stage I, Dukes B to UICC stage II, Dukes C to UICC stage III, and Dukes D to UICC stage IV. Particularly, the Dukes staging system is widely used internationally.

[0007] When detected at the early stage, colon cancer can be completely cured by endoscopic resection or surgical operation. Further, although metastasized to the liver or the lungs (distant metastasis), colon cancer may still be completely cured through surgical therapy in a period in which a surgery could be administered.

[0008] In other words, surgical therapy is the most effective therapy among the currently available therapies. However, if detected too late, cancer spreads to the organs such as the lungs, the liver, the lymph nodes and the peritoneum in which surgical therapy is difficult to apply. For that reason, contrary to the above case, surgical therapy is no use to apply in this case. Consequently, early detection and treatment are indispensable for treating colon cancer effectively.

[0009] Considering that there is a possibility that colon cancer may recur after surgical therapy, the patients should have a regular checkup for the recurrence of colon cancer at intervals of 3 to 4 months after surgical operation. Cancer recurrence is likely to occur in the liver, the lungs and the peritoneum rather than in the other organs. The recurrence is also locally observed in the excised site. The recurrence period of colon cancer is shorter than that of other cancers. The site in which the recurrence occurs is completely cured by resecting. Since more than 80% of the recurrent tumors are diagnosed within 3 years after surgical treatment, no recurrence within five years is defined as a criterion for complete cure.

[0010] If detected at an early stage, nearly 100% of colon cancer can be completely cured. In the meantime, it is very difficult to detect colon cancer in asymptomatic patients since the patients with colon cancer have no subjective symptoms

in the early stage. Accordingly, a periodic checkup should be required to detect colon cancer. An occult blood test is representative of colon cancer screening in detecting colon cancer.

[0011] However, the subject cannot be determined to have colon cancer as he or she shows a positive response in this test, Likewise, the indication of all negative responses does not guarantee the absence of colon cancer.

[0012] In this regard, it is unreasonable to apply the occult blood test as an accurate diagnostic method in detecting colon cancer. The screening methods of colon cancer currently producing useful diagnostic results are summarized in Table 3, below.

TABLE 3

Colon Cancer Examination

Examinations	Methods and Properties
Colonography	After a thorough cleaning out of the bowels, air, together with barium, is injected from the anus into the colon, followed by taking a series of X-ray images which is read by a radiologist.
Colonoscopy	Short colonoscopy for examining S-colon and long colonoscopy for examining the entire colon. Able to examine and remove polyps simultaneously.
Tumor marker	A method for diagnosing concealed cancer through blood test. Tumor markers that guarantee the diagnosis of cancer at an early stage have not yet been found. CEA is representative of tumor markers, but is positively detected only from about half of colon cancer patients. Used as a marker to indicate the progression of colon cancer and the therapeutic effect of a therapy.
Radiologic Diagnosis	Used to examine the progression of primary lesions and the distal metastasis of the cancer to the liver

[0013] A tumor marker characteristic of a specific cancer makes it possible to detect the cancer in an early stage through blood inspection. However, no tumor markers specific for colon cancer have been discovered yet. Although used for colon cancer, the marker CEA is positive only for about half of the patients as seen in Table 3. Thus, this marker is mainly employed to indicate the progression of colon cancer and the therapeutic effect of a therapy, but the marker is not reliable as a diagnostic marker for the early detection of colon cancer.

[0014] AZGP1 (alpha-2-glycoprotein 1, zinc-binding) is a secretary protein which consists of 295 amino acids and has the molecular weight of 33872 Da.

[0015] CXCL6 is a secretary protein which consists of 114 amino acids and has the molecular weight of 11,897 Da. It has a chemotactic function against neutrophils and granulocytes.

[0016] EGFL6 (EGF-like-domain, multiple 6) consists of 553 amino acids and has the molecular weight of 61317 Da, which is largely detected in fetal tissues. The previous study on the above gene is exemplified by U.S. Pat. No. 6,808,890.

[0017] AGT is angiotensinogen (serpin peptidase inhibitor, clade A, member 8) which consists of 485 amino acids and has the molecular weight of 53154 Da. This is also a secretary protein existing as a complex having PRG2 proform comprising disulfide-linked 2:2 heterotetramer, pro-PRG2 and C3 protein during pregnancy.

[0018] This protein was detected in pancreatic ductal cancer tissues (Ohta T, Amaya K, Yi S, Kitagawa H, Kayahara M, Ninomiya I, Fushida S, Fujimura T, Nishimura G, Shimizu K, Miwa K. Angiotensin converting enzyme-independent, local angiotensin II—generation in human pancreatic ductal cancer tissues. Int J Oncol. 2003 September; 23(3):593-8) and human male germ cell tumors (Murty V V, Li R G, Mathew S, Reuter V E, Bronson D L, Bosl G J, Chaganti R S. Replication error-type genetic instability at 1q42-43 in human male germ cell tumors. Cancer Res. 1994 Aug. 1; 54(15):3983-5) in relation to cancer.

[0019] CXCL 3 (C-X-C chemokine ligand 3) is a secretary protein which consists of 107 amino acids and has the molecular weight of 11342 Da. CXCL 3 exists in extracellular space. It has chemotactic activity against neutrophils and plays an important role in case of inflammation response.

[0020] Leading to the present invention, intensive and thorough research through the examination of various genes expected to be involved in colon cancer for expression levels in cancer tissues, such as colon cancer, stomach cancer, breast cancer, prostate cancer, liver cancer, etc. as well as in normal tissues using DNA chips, resulted in the finding that of the genes specifically expressed only in colon cancer tissues, highly putative colon cancer markers were confined to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P, and could be used alone or in combination as diagnostic markers for accurately detecting colon cancer in an early stage.

DISCLOSURE OF INVENTION

Technical Problem

[0021] It is an object of the present invention to provide a diagnosis marker for colon cancer, which can induce a quantitatively analyzable reaction with at least one protein or gene selected from among proteins or genes of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P.

[0022] It is another object of the present invention to provide diagnostic composition of colon cancer, comprising a marker for measuring an mRNA or protein expression level of at least one selected from among AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P.

[0023] It is a further object of the present invention to provide a diagnostic kit for colon cancer, comprising diagnostic composition of colon cancer. It is still a further object of the present invention to provide a method of yielding information necessary for the diagnosis of colon cancer, using diagnostic composition or kit for the colon cancer.

[0024] The present invention also provides a use of the said marker for the production of a composition for the diagnosis of colon cancer.

[0025] The present invention also provides a use of the said marker for the production of a kit for the diagnosis of colon cancer

Technical Solution

[0026] In accordance with an aspect thereof, the present invention provides a diagnostic composition of colon cancer, comprising at least one marker for measuring an mRNA expression level of at least one selected from among genes having base sequences of SEQ ID NOS. 1 to 9.

[0027] In accordance with another aspect thereof, the present invention provides a diagnostic composition of colon cancer, comprising at least one marker for measuring an expression level of a protein encoded by one gene selected from among genes having base sequences of SEQ ID NOS. 1 to 9.

[0028] In another preferred embodiment of the present invention, the present invention provides a use of a marker capable of measuring mRNA expression level of a specific

gene selected from the gene group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 or a combined marker comprising at least two markers mentioned above for the production of a diagnostic composition of colon cancer.

[0029] In another preferred embodiment of the present invention, the present invention provides a use of a marker capable of measuring the expression of a protein encoded by a gene selected from the gene group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 or a combined marker comprising at least two markers mentioned above for the production of a composition diagnosis of colon cancer.

[0030] The genes serving as diagnosis markers useful in the present invention are AZGP1 (alpha-2-glycoprotein 1) of SEQ ID NO. 1, CXCL3 (C-X-C chemokine ligand 3) of SEQ ID NO. 2, CXCL6 [chemokine (C-X-C motif) ligand 6, granulocyte chemotactic protein 2] of SEQ ID NO. 3, AGT [angiotensinogen(serpin peptidase inhibitor, clade A, member 8)] of SEQ ID NO. 4, FCGR3A of SEQ ID NO. 5, Col5A2 (collagen, type V, alpha 2) of SEQ ID NO. 6, S100P (S100 calcium binding protein P) of SEQ ID NO. 7, EGFL6 (EGF-like-domain, multiple 6) of SEQ ID NO. 8, and CTHRC1 (collagen triple helix repeat containing 1) of SEQ ID NO. 9.

[0031] The colon cancer diagnostic composition according to the present invention comprises a marker for measuring the mRNA or protein expression level of at least one selected from among the genes of SEQ ID NOS. 1 to 9. Preferably, the composition comprises two or more markers in combination. In this regard, the markers in combination may be composed of markers capable of measuring an mRNA expression level of one of the genes and a protein expression level of the same gene. Alternatively, the markers in combination are composed of markers capable of measuring mRNA expression levels or protein expression levels of two or more of the genes. When comprising the markers in combination, the composition diagnosis of colon cancer in accordance with the present invention can quantitatively analyze both the mRNA expression levels of the genes and the expression levels of the proteins encoded by the gene at the same time, thereby diagnosing colon caner at an early stage with a high level of reliability.

[0032] In an example of the present invention, the expression levels of the genes were found to be two to nine times higher in the biological samples taken from patients with colon cancer than in those taken from normal control.

[0033] It should be understood that base sequences showing sequence homology with those of the genes of SEQ ID NOS. 1 to 9 falls within the scope of the present invention. Likewise, the polypeptide sequences showing sequence homology with those encoded by the gene of SEQ ID NOS. 1 to 9 can be used in the present invention.

[0034] Sequence homology is used to describe the sequence relationships between two or more nucleic acids, polynucleotides, proteins or polypeptides and is understood in the context of the terms including (a) "reference sequence", (b) "comparison window", (c) "sequence identity", (d) "percentage of sequence identity" and (e) "substantial identity" or "homologous".

[0035] (a) A "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence, for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

[0036] (b) A "comparison window" includes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence may be compared to a reference sequence and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions, substitutions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions, substitutions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. It is obvious to those skilled in the art that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence, a gap penalty is typically introduced and is subtracted from the number of matches.

[0037] Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, Adv. Appl. Math. 2: 482 (1981); by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48: 443 (1970); by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. 85: 2444 (1988); by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, Calif., GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis., USA; (Higgins and Sharp, Gene, 73: 237-244, 1988). The BLAST family of programs which can be used for database similarity searches includes: BLASTN for nucleotide query sequences against nucleotide database sequences; BLASTX for nucleotide query sequences against protein database sequences; BLASTP for protein query sequences against protein database sequences; TBLASTN for protein query sequences against nucleotide database sequences; and TBLASTX for nucleotide query sequences against nucleotide database sequences (See, Current Protocols in Molecular Biology, Chapter 19, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). New versions of these or new programs will be obviously available and can be used along with the present invention.

[0038] (c) "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window and which can be mutated typically by addition, deletion or substitution. When percentage of sequence identity is used in reference to proteins, it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e. g. charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences which differ by such conservative substitutions are said to have "sequence similarity" Means for making this adjustment are well-known to those skilled in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conser-

vative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e. g., according to the algorithm of Meyers and Miller, Computer Applic. Biol. Sci., 4: 11-17 (1988) e. g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, Calif., USA).

[0039] (d) "Percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions, substitutions or deletions (gaps) as compared to the reference sequence (which does not comprise additions, substitutions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[0040] (e) i) The term "substantial identity" or "homologous" means that a polynucleotide comprises a sequence that has at least 60% sequence identity, preferably at least 70%, more preferably at least 80%, far more preferably at least 90%, and most preferably 95%, 96%, 97%, 98%, 99% or 100%, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

[0041] Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 60%, more preferably at least 70%, 80%, 90%, and most preferably at least 95%, 96%, 97%, 98%, 99% or 100%. Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. For example, this may occur when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

[0042] (e) ii) The terms "substantial identity" or "homologous" in the context of a peptide indicates that a peptide comprises a sequence with at least 60% sequence identity to a reference sequence, preferably 70%, more preferably 80%, far more preferably 85%, most preferably at least 90% or 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the reference sequence over a specified comparison window.

[0043] The term "diagnosis", as used herein, means the process of identifying a medical condition or disease by its signs and symptoms. For the purpose of the present invention, "diagnosis" is used to mean determining the incidence of colon by examining whether the diagnostic marker of the present invention is expressed.

[0044] As used herein, the term "colon cancer" is intended to refer to cancerous growths on the innermost surface mucous membrane, including colon carcinoma, rectal cancer, and anal cancer.

[0045] The terms "marker for diagnosis", "diagnostic marker" or "diagnosis marker", as used herein, is intended to indicate a substance capable of diagnosing colon cancer by distinguishing colon cancer cells from normal cells, and includes organic biological molecules, quantities of which are increased or decreased in colon cancer cells relative to normal cells, such as polypeptides or nucleic acids (e. g., mRNA, etc.), lipids, glycolipids, glycoproteins and sugars (monosaccharides, disaccharides, oligosaccharides, etc.). Also, primers and antibodies fall within the scope of the markers according to the present invention as long as they can be used to quantitatively measure the change of these biomolecules in expression level in vivo. With respect to the objects of the present invention, examples of the colon cancer diagnostic markers include AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P of respective SEQ ID NOS. 1 to 9, which are genes whose expression is increased in colon cancer cells, related nucleic acids (e.g., mRNAs), organic biomolecules such as lipids, glycolipids, glycoproteins, sugars (monosaccharides, disaccharides, oligosaccharides), primer sets or DNA chips capable of identifying the expression patterns of the mRNAs, and antibodies capable of identifying the expression patterns of the proteins.

[0046] The terms "Marker for diagnosis", "diagnostic marker" or "diagnosis marker", as used herein, is intended to indicate a substance capable of diagnosing colon cancer by distinguishing colon cancer cells from normal cells, and includes organic biological molecules, quantities of which are increased or decreased in colon cancer cells relative to normal cells, such as polypeptides or nucleic acids (e. g., mRNA, etc.), lipids, glycolipids, glycoproteins and sugars (monosaccharides, disaccharides, oligosaccharides, etc.). Also, primers and antibodies fall within the scope of the markers according to the present invention as long as they can be used to quantitatively measure the change of these biomolecules in expression level in vivo. With respect to the objects of the present invention, examples of the colon cancer diagnostic markers include AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P of respective SEQ ID NOS. 1 to 9, which are genes whose expression is increased in colon cancer cells, related nucleic acids (e.g., mRNAs), organic biomolecules such as lipids, glycolipids, glycoproteins, sugars (monosaccharides, disaccharides, oligosaccharides), primer sets or DNA chips capable of identifying the expression patterns of the mRNAs, and antibodies capable of identifying the expression patterns of the proteins.

[0047] The selection and application of significant diagnostic markers determine the reliability of diagnosis results. A significant diagnostic marker means a marker that has high validity, giving accurate diagnostic results, and high reliability, supplying constant results upon repeated measurement. The colon cancer diagnostic markers of the present invention, which are genes whose expression always increases by direct or indirect factors when colon cancer occurs, display the same results upon repeated tests, and have high reliability due to a great difference in expression levels compared to a control, thus having a very low possibility of giving false results. Therefore, based on the results, the diagnosis, obtained by measuring the expression levels of the significant diagnostic markers of the present invention, is valid and reliable.

[0048] At this time, the genes which are expressed on almost the same level between normal colonic epithelial cells and colon cancer cells were excluded. The genes which were expressed at two to nine or more times higher levels specifically in colon cancer cells compared to cells of normal tissues were selected as diagnostic markers of colon cancer.

[0049] As long as it is applied to the quantification of mRNA levels of at least one of the genes, any primer set may be used as a diagnostic marker. Preferable is a primer set binding specifically to one of SEQ ID NOS. 1 to 9. In the present invention, the primer set is selected from among base sequence sets of SEQ ID NOS. 10 to 27.

[0050] As used herein, the term "primer" refers to a short nucleic acid strand having a free 3' hydroxyl group, which forms a base pair with a complementary template so as to serve as a starting point for the production of a new template strand. DNA synthesis or replication requires a suitable buffer, proper temperatures, polymerizing enzyme (DNA polymerase, or reverse transcriptase), and four kinds of nucleotide triphosphates, in addition to primers. The primers useful in the present invention are sense and antisense nucleic acids ranging in length from 7 to 50 nucleotides. As long as its basic property of serving as a starting point is not altered, the primers may incorporate an additional characteristic thereinto.

[0051] The primers useful in the present invention may be chemically synthesized using a phosphoamidite solid support method or other well-known techniques. Its nucleotide sequences may be modified using various means known in the art. Illustrative, non-limiting examples of the modification include methylation, capping, substitution of natural nucleotides with one or more homologues, and alternation between nucleotides, such as uncharged linkers (e.g., methyl phosphonate, phosphotriester, phosphoroamidate, carbamate, etc.) or charged linkers (e.g., phosphorothioate, phosphorodithioate, etc.). Nucleic acids may contain one or more additionally covalent-bonded residues, which are exemplified by proteins (e. g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalating agents (e. g., acridine, psoralene, etc.), chelating agents (e. g., metals, radioactive metals, iron, oxidative metals, etc.), and alkylating agents. The nucleic acid sequences of the present invention may also be altered using a label capable of directly or indirectly supplying a detectable signal. Examples of the label include radioisotopes, fluorescent molecules and biotin.

[0052] In accordance with an embodiment of the present invention, the composition for detecting a diagnostic marker of colon cancer includes a pair of primers specific to one or more genes selected from among AZGP1, CXCL3, CXCL6, AGT, FCGR3A, Col5A2, S100P, EGFL6, and CTHRC1 (Table 4).

TABLE 4

[Table 4] [Table]			
AZGP1		Co15A2	
SEQ ID NO. 10	forward ctctgcggaaat-acctgaaa	SEQ ID NO. 20	forward gacctcgtggt-gacaaaggt
SEQ ID NO. 11	Reverse tgaagaa-catctccccgtaa	SEQ ID NO. 21	Reverse agccgcct-gatcttcagtaa
CXCL3		S100P	
SEQ ID NO. 12	forward ggtgctccccttgttcag	SEQ ID NO. 22	forward agacagc-catgggcatgat
SEQ ID NO. 13	Reverse agggaattcacctcaaga	SEQ ID NO. 23	Reverset catttgagtcct-gccttctc
CXCL6		EGFL6	
SEQ ID NO. 14	forward agatccctggacccagta	SEQ ID NO. 24	forward gcatgaaaaa-gaaggcaaaa
SEQ ID NO. 15	Reverse ttgccaaagggttcaata	SEQ ID NO. 25	Reverse tgtcattcttcagggctttc
AGT		CTHRC1	
SEQ ID NO. 16	forward gctgcaaaacttgacacc	SEQ ID NO. 26	forward tcatcg-cacttcttctgtgga
SEQ ID NO. 17	Reverse attgcctgtagcctgtca	SEQ ID NO. 27	Reverse gccaaccca-gatagcaacatc
FCGR3A		β-actin (Control)	
SEQ ID NO. 18	forward gcttgttgggag-taaaaatg	A	forward gatcattgctc-ctcctgagc
SEQ ID NO. 19	Reverse tccagtcttgttgagctt gcttgctgatccac	B	Reverse actcct-

[0053] In the composition for the diagnosis of colon cancer according to the present invention, any can be used as a marker for measuring expression levels of the proteins as long as it detects a change of proteins in expression level in colon cancer cells. Preferably, the marker is an antibody specific for one of the proteins encoded by the gene of SEQ ID NOS. 1 to 9 (AZGP1, CXCL3, CXCL6, AGT, FCGR3A, Col5A2, S100P, EGFL6, and CTHRC1).

[0054] The term "antibody" as used herein, refers to a specific protein molecule that indicates an antigenic region. With respect to the objects of the present invention, an antibody binds specifically to a marker protein, and includes all of polyclonal antibodies, monoclonal antibodies and recombinant antibodies.

[0055] Since the colon cancer marker protein is identified as described above, it may be used to produce antibodies using techniques widely known in the art.

[0056] Polyclonal antibodies may be produced by a method widely known in the art, which includes injecting the colon cancer marker protein antigen into an animal and collecting blood samples from the animal to obtain serum containing antibodies. Such polyclonal antibodies may be prepared from a certain animal host, such as goats, rabbits, sheep, monkeys, horses, pigs, cows and dogs. The antibodies produced can be isolated and purified using gel electrophoresis, dialysis, salting out, ion exchange chromatography, affinity chromatography, and other techniques.

[0057] Monoclonal antibodies may be prepared by a method widely known in the art, such as a hybridoma method (Kohler and Milstein (1976) European Journal of Immunology 6:511-519), or a phage antibody library technique (Clackson et al., Nature, 352:624-628, 1991; Marks et al, J. Mol. Biol., 222:58, 1-597, 1991). The antibody produced above can be isolated and purified by gel electrophoresis, dialysis, salt precipitation, ion exchange chromatography, affinity chromatography, etc.

[0058] In addition, the antibodies of the present invention include complete forms having two full-length light chains and two full-length heavy chains, as well as functional fragments of antibody molecules. The functional fragments of antibody molecules refer to fragments retaining at least an antigen-binding function, and include Fab, F(ab'), F(ab')2, Fv and the like.

[0059] In the composition for the diagnosis of colon cancer, the antibody is preferably a microparticle-conjugated antibody. The micro particle may be preferably colored latex or colloidal gold particle.

[0060] In the composition for the diagnosis of colon cancer, any antibody may be used as long as it can be applied to the quantitative analysis of the expression level of the proteins encoded by the genes of SEQ ID NOS. 1 to 9. Preferable is an antibody used in an immunochromatographic strip kit, a Luminex assay kit, a protein microarray kit, an ELISA kit or an immunodot kit.

[0061] Preferably, the immunochromatographic strip useful in the composition for the diagnosis of colon cancer comprises (a) a sample pad onto which a sample is absorbed; (b) a conjugate pad in which an antibody binds to proteins encoded by one or more genes selected from among base sequences of SEQ ID NOS. 1 to 9; (c) a test membrane with a test line and a control line, comprising a monoclonal antibody to the proteins encoded by one or more selected from among the genes of SEQ ID NOS. 1 to 9; (d) an absorbent pad into which remaining samples are absorbed; and (e) a support.

[0062] The Luminex assay kit, the microarray kit, or the ELISA kit which may be useful in the composition for the diagnosis of colon cancer preferably comprises a secondary antibody the poly- or monoclonal antibody, whether conjugated with a label, to a protein encoded by the gene selected from among the genes of SEQ ID NOS. 1 to 9.

[0063] In accordance with another aspect thereof, the present invention provides a kit for diagnosing colon cancer, comprising the colon cancer diagnostic composition containing one or more markers capable of measuring the expression level of mRNA or protein of the gene selected from among genes of SEQ ID NOS. 1 to 9.

[0064] In another preferred embodiment of the present invention, the present invention provides a use of a marker capable of measuring mRNA expression level of a specific gene selected from the gene group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 or a combined marker comprising at least two markers mentioned above for the production of a kit for the diagnosis of colon cancer.

[0065] In another preferred embodiment of the present invention, the present invention provides a use of a marker capable of measuring the expression of a protein encoded by a gene selected from the gene group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 or a combined marker comprising at least two markers mentioned above for the production of a kit for the diagnosis of colon cancer.

[0066] The term "measurement of mRNA expression levels" or corresponding phrases, as used herein, are intended to refer to a process of assessing the presence and expression levels of mRNA of colon cancer marker genes in biological samples for diagnosing colon cancer, in which the amount of mRNA is measured. Analysis methods for measuring mRNA levels include, but are not limited to, RT-PCR, competitive RT-PCR, real-time RT-PCR, RNase protection assay (RPA), Northern blotting and DNA chip assay.

[0067] The term "measurement of protein expression levels" or corresponding phrases, as used herein, are intended to refer to a process of assessing the presence and expression levels of proteins expressed from colon cancer marker genes in biological samples for diagnosing colon cancer, in which the amount of protein products of the marker genes is measured using antibodies specifically binding to the proteins. Analysis methods for measuring protein levels include, but are not limited to, Western blotting, enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), radioimmunodiffusion, Ouchterlony immunodiffusion, rocket immunoelectrophoresis, immunohistostaining, immunoprecipitation assay, complement fixation assay, FACS, and protein chip assay.

[0068] In a preferable embodiment, the diagnostic kit of the present invention is characterized by including essential elements required for performing RT-PCR. An RT-PCR kit includes a pair of primers specific for each marker gene. The primers are nucleotides having sequences specific to a nucleic acid sequence of each marker gene, and are about 7 by to 50 bp in length, more preferably about 10 by to 30 bp in length. Also, the RT-PCR kit may include primers specific to a nucleic acid sequence of a control gene. The RT-PCR may further include test tubes or other suitable containers, reaction buffers (varying in pH and magnesium concentrations), deoxynucleotides (dNTPs), enzymes such as Taq-poly-

merase and reverse transcriptase, DNAse, RNAse inhibitor, DEPC-treated water, and sterile water.

[0069] As long as it is applied to diagnose colon cancer, any type kit can be used in the present invention. Preferable is a reverse transcription-polymerase chain reaction kit, an immunodot kit, an ELISA kit, an immunochromatography kit, a Luminex assay kit, or a protein microarray kit thanks to their ability to rapidly and accurately measure mRNA or protein expression levels of biological samples. Preferably, a diagnostic kit for the colon cancer may further comprise one or more components, solutions or devices suitable for the analysis of colon cancer.

[0070] The luminex kit useful as a diagnostic kit of the present invention may comprise poly- and monoclonal antibodies to the proteins encoded by the genes of SEQ ID NOS. 1 to 9, and a secondary antibody to the poly- or monoclonal antibodies. The luminex assay according to the present invention is high-throughput quantification method which can analyze as many as 100 analytes at the same time even if the patient samples are present in a small amount (10~20 μl) and are not pretreated. The luminex assay is highly sensitive (pg level) and can perform quantitative analysis within a short time (3~4 hours), so that it is used as an alternative to ELISA or ELISPOT assay. An luminex assay is a multiplexed fluorescent microplate method by which 100 or more biological samples can be analyzed in each well of 96-well plates and employs two laser detectors to progress signal transmission in real time, so that polystyrene beads can be discriminated by 100 or more colors. 100 beads are designed in the following manner. In a 10×10 bead matrix, red fluorescent beads and orange fluorescent beads are divided into 10 or more classes according to intensities on respective sides. Within the matrix, the columns contain beads at different ratios of red and orange colors to form 100 color-coded bead set in total. Also, each bead is coated with an antibody to a target protein and thus can be used for protein quantification through immune responses. In this assay, a sample is analyzed using two laser rays. One laser is used to detect beads to identify the inherent bead number provided while the other laser functions to sense a sample protein reacted with the antibody conjugated to the bead. Therefore, 100 different proteins can be analyzed at the same time in one well. This assay also enjoys the advantage of sensing a sample even if it is present in an amount of as small as 15 μl.

[0071] A luminex kit with which a luminex assay can be performed in accordance with the present invention includes an antibody specific to the marker protein. The antibody may be a monoclonal, polyclonal or recombinant antibody, which has high specificity and affinity to each marker protein and rarely has cross-reactivity to other proteins. Also, the Luminex kit may comprise an antibody specific for a control protein. The Luminex kit may further include reagents capable of detecting bound antibodies, for example, a labeled secondary antibody, chromophores, enzymes (e. g., conjugated with an antibody) and their substrates or other substances capable of binding to the antibodies. Also, the antibody may be an antibody conjugated to microparticles which may be selected from among colored latex particles and colloidal gold particles.

[0072] In another embodiment of the present invention, the diagnostic kit may be characterized by including essential elements required for performing a DNA chip assay. A DNA chip kit may include a substrate plate onto which genes or fragments thereof, cDNA or oligonucleotides, are attached, and reagents, agents and enzymes for preparing fluorescent probes. Also, the substrate plate may include a control gene or fragments thereof, such as cDNA or oligonucleotides.

[0073] Further, preferably, the diagnostic kit is characterized by including essential elements required for performing ELISA. An ELISA kit includes antibodies specific to marker proteins. The antibodies may be monoclonal, polyclonal or recombinant antibodies, which have high specificity and affinity to each marker protein and rarely have cross-reactivity to other proteins. Also, the ELISA kit may include an antibody specific to a control protein. The ELISA kit may further include reagents capable of detecting bound antibodies, for example, a labeled secondary antibody, chromophores, enzymes (e. g., conjugated with an antibody) and their substrates or other substances capable of binding to the antibodies.

[0074] The diagnostic kit for colon cancer comprising an immunochromatographic strip for diagnosing colon cancer is characterized by including essential elements required for performing a rapid diagnostic test which gives an analysis result within 5 min. A rapid diagnostic test kit with an immunochromatographic strip includes antibodies specific to marker proteins. The antibodies may be monoclonal, polyclonal or recombinant antibodies, which have high specificity and affinity to each marker protein and rarely have cross-reactivity to other proteins. Also, the rapid test kit may further include other substances necessary for the diagnosis, for example, a membrane on which specific antibodies and secondary antibodies are immobilized, a membrane with antibody-conjugated beads bound thereto, an absorbent pad, and a sample pad.

[0075] Also, the colon cancer diagnostic kit of the present invention may be characterized by including essential elements required for performing protein microarray for analyzing combined markers simultaneously. The protein microarray kit useful in the present invention includes antibodies specific to marker proteins bound to a solid support. The antibodies may be monoclonal, polyclonal or recombinant antibodies, which have high specificity and affinity to each marker protein and have little cross-reactivity to other proteins. Also, the protein microarray kit may include an antibody specific to a control protein. The protein microarray kit may further include reagents capable of detecting bound antibodies, for example, a labeled secondary antibody, chromophores, enzymes (e. g., conjugated with an antibody) and their substrates or other substances capable of binding to the antibodies. The protein microarray of the present invention may include poly- and/or monoclonal antibodies to the protein bound to the slide and an enzyme-conjugated secondary antibody to the poly- or monoclonal antibodies.

[0076] In another preferred embodiment of the present invention, the present invention provides a method for the diagnosis of colon cancer among patients having high risk of colon cancer, which is composed of the following steps:

[0077] 1) measuring expression levels of one or more genes selected from the gene group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 in biological samples taken from patients; and

[0078] 2) taking the measured expression levels, particularly increased levels, as colon cancer risk index, selecting patients demonstrating higher expression levels than normal people.

[0079] The expression level herein indicates the level of mRNA of one or more genes selected from those genes hav-

ing the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9 or the expression level of a protein expressed from one or more genes selected from the group consisting of those genes having the nucleotide sequences represented by SEQ. ID. NO: 1 NO: 9.

[0080] In accordance with another aspect thereof, the present invention provides a method for yielding information necessary for the diagnosis of colon cancer, comprising measuring mRNA levels in a biological sample from a patient with suspected colon cancer using one or more primer sets, selected from among base sequences of SEQ ID NOS. 10 to 27, specific to one or more genes selected from among genes of SEQ ID NOS. 1 to 9 (AZGP1, CXCL3, CXCL6, AGT, FCGR3A, Col5A2, S100P, EGFL6, and CTHRC1); and comparing mRNA levels of the sample from the patient with those of a normal control sample to determine an increase in mRNA levels.

[0081] The isolation of mRNA from a biological sample may be achieved using a known process, and mRNA levels may be measured by a variety of methods.

[0082] Analysis methods for measuring mRNA levels include RT-PCR, competitive RT-PCR, real-time RT-PCR, RNase protection assay (RPA), Northern blotting and DNA chip assay, but are not limited thereto.

[0083] With the detection methods, a patient with suspected colon cancer is compared with a normal control for mRNA expression levels of a colon cancer marker gene, and the patient's suspected colon cancer is diagnosed by determining whether expression levels of mRNA from the colon cancer marker gene have significantly increased.

[0084] mRNA expression levels are preferably measured by RT-PCR or DNA chip using primers specific to a gene serving as a colon cancer marker.

[0085] After RT-PCT, the products are electrophoresed, and patterns and thicknesses of bands are analyzed to determine the expression and levels of mRNA from a gene used as a diagnostic marker of colon cancer while comparing the mRNA expression and levels with those of a control, thereby simply diagnosing the incidence of colon cancer. Alternatively, mRNA expression levels may be measured using a DNA chip in which the colon cancer marker genes or nucleic acid fragments thereof are anchored at high density to a glass-like base plate. A cDNA probe labeled with a fluorescent substance at its end or internal region is prepared using mRNA isolated from a sample, and is hybridized with the DNA chip. The DNA chip is then read to determine the presence or expression levels of the gene, thereby diagnosing the incidence of colon cancer.

[0086] In accordance with another aspect thereof, the present invention provides a method of diagnosing colon cancer, comprising measuring protein levels by contacting an antibody specific to one or more genes selected from among the genes of SEQ ID NOS. 1 to 9(AZGP1, CXCL3, CXCL6, AGT, FCGR3A, Col5A2, S100P, EGFL6, and CTHRC1) with a biological sample from a patient with suspected colon cancer to form antigen-antibody complexes; and comparing protein levels of the sample from the patient with those of a normal control sample to determine an increase in protein level.

[0087] The isolation of proteins from a biological sample may be achieved using a known process, and protein levels may be measured by a variety of methods.

[0088] Analysis methods for measuring mRNA levels include RT-PCR, competitive RT-PCR, real-time RT-PCR, RNase protection assay (RPA), Northern blotting and DNA chip assay, but are not limited thereto.

[0089] The term "biological sample", as used herein particularly for the measurement of mRNA or protein levels, includes samples displaying a difference in expression levels of a colon cancer marker gene, such as tissues, cells, whole blood, serum, plasma, saliva, sputum, cerebrospinal fluid and urine, but is not limited thereto.

[0090] Analysis methods for measuring protein levels in accordance with the present invention include, but are not limited to, an immunochromatography assay, an immunodot assay, a Luminex assay, an ELISA assay, a protein microarray assay, an immunostaining assay, a Western blotting assay, a radioimmunoassay (RIA), a radioimmunodiffusion assay, an ouchterlony immunodiffusion assay, a rocket immunoelectrophoresis assay, an immunohistostaining assay, an immunoprecipitation assay, a complement fixation assay, FACS, and a protein chip assay.

[0091] The measurement of protein levels by immunodot assay may be carried out by (a) dotting a biological sample on a membrane; (b) reacting the sample with antibodies specific for the proteins encoded by one or more genes selected from among the genes of SEQ ID NOS. 1 to 9; and (c) adding a labeled secondary antibody to the membrane and developing a color. The ELISA assay is preferably a sandwich ELISA assay which can be implemented by (a) immobilizing Antibody 1 to the proteins of one or more genes selected from among the genes of SEQ ID NOS. 1 to 9; (b) reacting the immobilized Antibody 1 with a biological sample from a patient with suspected colon cancer to form an antigen-antibody complex; binding to the complex labeled Antibody 2 specific for the proteins encoded by one or more genes selected from among the genes of SEQ ID NOS. 1 to 9; and detecting the label to determine the protein level. The protein microarray assay preferably comprises (a) immobilizing onto a chip a polyclonal antibody specific for the proteins encoded by one or more genes selected from among the genes of SEQ ID NOS. 1 to 9; (b) reacting the immobilized Antibody 1 with a biological sample from a patient with suspected colon cancer to form an antigen-antibody complex; (c) binding to the complex a labeled monoclonal antibody specific for the proteins encoded by one or more genes selected from among the genes of SEQ ID NOS. 1 to 9; and (d) detecting the label to determine the protein level.

[0092] Through the analysis assays, a quantitative comparison can be made between the antigen-antibody complexes in a normal control and a patient with suspected colon cancer. Based on this comparison, a significant increase in the level of the colon cancer marker gene can be determined, thus giving information necessary for the diagnosis of colon cancer.

[0093] As used herein, the term "antigen-antibody complex" is intended to refer to binding products of a colon cancer marker protein to an antibody specific thereto. The antigen-antibody complex thus formed may be quantitatively determined by measuring the signal size of a detection label.

[0094] Such a detection label may be selected from a group consisting of enzymes, fluorescent substances, ligands, luminescent substances, microparticles, redox molecules and radioactive isotopes, but the present invention is not limited to the examples. Examples of the enzymes available as detection labels include, but are not limited to, β-glucuronidase, β-D-glucosidase, β-D-galactosidase, urase, peroxidase, alkaline phosphatase, acetylcholinesterase, glucose oxidase, hexokinase and GDPase, RNase, glucose oxidase and luciferase,

phosphofructokinase, phosphoenolpyruvate carboxylase, aspartate aminotransferase, phosphenolpyruvate decarboxylase, and β-latamase. Examples of the fluorescent substances include, but are not limited to, fluorescin, isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde, fluorescamin and DAP. As the ligands, bitine derivatives are useful, but are not given as a factor limiting the present invention. Examples of the luminescent substances include acridinium esters, luciferin and luciferase, but are not limited thereto. As for the microparticles, its examples include, but are not limited to, colloidal gold and colored latex. Examples of the redox molecules include, but are not limited to, ferrocene, ruthenium complexes, viologen, quinone, Ti ions, Cs ions, diimide, 1,4-benzoquinone, hydroquinone, $K_4W(CN)_8$, $[Os(bpy)_3]^{2+}$, $[RU(bpy)_3]^{2+}$ and $[MO(CN)_8]^4$. Examples of the radioactive isotopes include, but are not limited to, ^3H, ^{14}C, ^{32}P, ^{35}S, ^{36}Cl, ^{51}Cr, ^{57}Co, ^{58}Co, ^{59}Fe, ^{90}Y, ^{125}I, ^{131}I and ^{186}Re.

[0095] Preferably, the protein expression levels are measured by ELISA. Examples of ELISA include direct ELISA using a labeled antibody recognizing an antigen immobilized on a solid support; indirect ELISA using a labeled antibody recognizing a capture antibody forming complexes with an antigen immobilized on a solid support; direct sandwich ELISA using a labeled antibody recognizing an antigen bound to a antibody immobilized on a solid support; and indirect sandwich ELISA, in which a captured antigen bound to an antibody immobilized on a solid support is detected by first adding an antigen-specific antibody, and then a secondary labeled antibody which binds the antigen-specific antibody. More preferably, the protein expression levels are detected by sandwich ELISA, where a sample reacts with an antibody immobilized on a solid support, and the resulting antigen-antibody complexes are detected by adding a labeled antibody specific for the antigen, followed by enzymatic development, or by first adding an antigen-specific antibody and then a secondary labeled antibody which binds to the antigen-specific antibody, followed by enzymatic development. Information necessary for the diagnosis of colon cancer can be provided by measuring the degree of complex formation of a colon cancer marker protein and an antibody thereto.

[0096] Further, the measurement of protein expression levels is preferably achieved using Western blotting using one or more antibodies to the colon cancer makers. Total proteins are isolated from a sample, separated according to size by electrophoresis, transferred onto a nitrocellulose membrane, and reacted with an antibody. The amount of proteins produced by gene expression is determined by measuring the amount of antigen-antibody complexes produced using a labeled antibody, thereby diagnosing the incidence of colon cancer. The detection method comprises assessing expression levels of maker genes in a control and cells in which colon cancer occurs. mRNA or protein levels may be expressed as an absolute (e.g., μg/ml) or relative (e. g., relative intensity of signals) difference in the amount of marker proteins.

[0097] Also, the measurement of protein expression levels is preferably performed with an immunochromatography diagnostic kit which is characterized by essential elements required for a rapid test which gives a result within 5 min. A rapid test kit using an immunochromatographic strip comprises an antibody specific for a marker protein. The antibody may be a monoclonal, polyclonal or recombinant antibody, which has high specificity and affinity to each marker protein and rarely have cross-reactivity to other proteins.

[0098] In addition, the rapid test kit may further include other reagents capable of detecting bound antibodies, for example, a nitrocellulose membrane onto which specific antibodies and secondary antibodies are immobilized, a membrane with antibody-conjugated beads bound thereto, an absorbent pad, and a sample pad.

[0099] In addition, the measurement of protein expression levels can be carried out with an assay kit which is characterized by including essential elements required for Luminex assay which is typically designed to analyze combined markers at the same time. A Luminex kit includes an antibody specific for a maker protein. The antibody may be a monoclonal, polyclonal or recombinant antibody, which has high specificity and affinity to each marker protein and rarely have cross-reactivity to other proteins. Also, the Luminex kit may comprise an antibody specific for a control protein. The Luminex kit may further include reagents capable of detecting bound antibodies, for example, a labeled secondary antibody, chromophores, enzymes (e. g., conjugated with an antibody) and their substrates or other substances capable of binding to the antibodies.

[0100] The diagnostic kit useful in measuring protein expression levels in accordance with the present invention is characterized by including essential elements required for performing protein microarray so as to analyze combined markers simultaneously. The microarray kit includes antibodies specific to marker proteins bound to a solid support. The antibodies may be monoclonal, polyclonal or recombinant antibodies, which have high specificity and affinity to each marker protein and have little cross-reactivity to other proteins. Also, the protein microarray kit may include an antibody specific to a control protein. The protein microarray kit may further include reagents capable of detecting bound antibodies, for example, a labeled secondary antibody, chromophores, enzymes (e. g., conjugated with an antibody) and their substrates or other substances capable of binding to the antibodies. By a method of analyzing a sample using a protein microassay, proteins are isolated from the sample and hybridized with the protein chip to form antigen-antibody complexes. The protein chip is then read to determine the presence or expression levels of the proteins, thereby providing information necessary for the diagnosis of colon cancer.

[0101] In a preferable embodiment, the protein expression levels may be measured through immunohistostaining using one or more antibodies to the colon cancer marker. Normal colonic epithelial tissues and colon cancer-suspected tissues are taken, immobilized, and embedded in a paraffin block which is then sectioned to slices of micrometers thickness on glass slides, followed by reaction with one of the antibodies. Thereafter, the antibodies which remain unreacted are washed off, and the bound antibodies are labeled with one of the above-mentioned detection labels. Under a microscope, the labeling of the antibodies is read.

Advantageous Effects

[0102] The marker of the present invention for the diagnosis of colon cancer facilitates fast and easy diagnosis of colon cancer by using those genes over-expressed specifically in colon cancer tissues and therefore it can be effectively used for the screening of candidates for colon cancer treatment agents.

BRIEF DESCRIPTION OF DRAWINGS

[0103] FIG. **1** is electrophoresis photographs showing expression levels of AZGP1, AGT, EGFL6, and CXCL3 in normal tissues and colon cancer tissues as identified by reverse transcription PCR.

[0104] FIG. 2 is electrophoresis photographs showing expression level of CTHRC1 in normal tissues and colon cancer tissues as identified by reverse transcription PCR.

[0105] FIG. 3 is an electrophoresis photograph showing expression levels of AZGP1, AGT, and EGFL6 in 10 colon cancer cell lines as identified by RT-PCR.

[0106] FIG. 4 is a view showing expression levels of AGT, EGFL6, and CXCL3 in normal sera and colon cancer sera as identified by Western blotting.

[0107] FIG. 5 is microphotographs showing expression level of EGFL6 in normal mucous membrane and colon cancer tissues as identified by immunohistostaining.

[0108] FIG. 6 is microphotographs showing expression level of CTHRC1 in normal mucous membrane and colon cancer tissues as identified by immunohistostaining.

[0109] FIG. 7 is microphotographs showing expression level of CXCL-3 in normal mucous membrane and colon cancer tissues as identified by immunohistostaining.

[0110] FIG. 8 is microphotographs showing expression level of AGT in normal mucous membrane and colon cancer tissues as identified by immunohistostaining.

[0111] FIG. 9 is a diagram illustrating the principal of immunological dot assay.

[0112] FIG. 10 is a photograph illustrating the comparison of protein expressions between normal serum and colon cancer patient serum, investigated by immunological dot assay.

[0113] FIG. 11 is a standard curve for AGT protein, established by an ELISA assay.

[0114] FIG. 12 is a schematic diagram showing a structure of an immunochromatographic strip according to the present invention.

MODE FOR THE INVENTION

[0115] A better understanding of the present invention may be obtained through the following examples which are set forth to illustrate, but are not to be construed as limiting the present invention.

EXAMPLE 1

Excavation of Genes Overexpressed in Colon Cancer Using DNA Chip

[0116] In order to primarily extract genes which are overexpressed specifically in colon cancer cells compared to normal colonic epithelial cells, 2,230 genes were examined for expression level using DNA chips (48K human microarray, commercially available from Illumina).

[0117] Total mRNA was isolated from normal colonic epithelial cells and colon cancer cells using an RNeasy Mini Kit (QIAGEN) and quantitatively analyzed on a chip (Experion RNA StdSens, Bio-Rad). For use in hybridization, the total mRNA was biotinylated and amplified using Illumina TotalPrep RNA Amplification Kit (Ambion). cDNA was synthesized with T7 oligo-dT primers and biotinylated by in vitro transcription with biotin-UDP.

[0118] The biotin-labeled cDNA thus formed was quantified using NonoDrop. The cDNA prepared from normal colonic epithelial cells and colon cancer cells was hybridized on a chip (Human-6 V2, Illumina). After hybridization, the DNA chip was washed with buffer (Illumina Gene Expression System Wash Buffer, Illumina) to remove non-specific hybridizations and labeled with fluorescent streptavidin-Cy3 conjugate (Amersham).

[0119] The fluorescence-labeled DNA chip was scanned using a confocal laser scanner (Illumina) to give fluorescence data of each spot. The fluorescence data were saved as TIFF images. The TIFF images were quantified with BeadStudio version 3 (Illumina) to quantify the fluorescence intensity at each spot. The quantitative results were normalized using the quantile function supplied by the program Avadis Prophetic version 3.3 (Strand Genomics).

[0120] As a result, 1,601 genes were analyzed for expression level in normal colonic epithelial cells and colon cancer cells, and the genes with overexpression of mRNA in colon cancer cells were finally selected (Table 5).

TABLE 5

	KRIBB Fold change 2^n
I AZGP1	3.00
EGFL6	2.97
S100P	3.25
CTHRC1	2.69
CXCL6	2.67
CXCL3	0.27
FCGR3A	0.26
AGT	2.38
Col5A2	1.14

EXAMPLE 2

mRNA Isolation from Tissues and Cells

[0121] For use in reverse transcription PCR, mRNA was isolated from total 40 tissues consisting of normal colonic epithelial cells and colon cancer cell tissues from 20 patients with colon cancer.

[0122] First, immediately after the surgical resection of tissues, blood was removed from the tissues in sterile phosphate buffered saline and frozen in liquid nitrogen. Thereafter, total mRNA was isolated in a single-step RNA isolation manner using the guanidinium method. The total mRNA thus obtained was quantified with a spectrophotometer and stored in a −70° C. freezer until use.

[0123] 10 colon cancer cell lines (DLD-1, HT29. HCT116, colo205, SW480, SW620, SNU C1, SNU C2A, KM 12C, KM 12SM) were obtained from KCLB (the Korean Cell Line Bank, located at 28, Yeonkun-dong, Jongno, Seoul, Korea).

[0124] Each cell line was cultured for 5~6 days in DMEM (Invitrogen) or RPMI1640 (Invitrogen), supplemented with 10% fetal bovine serum (FBS, Hyclon) and 1 mg/ml penicillin/streptomycin (Sigma), after which total RNA was isolated in a single-step RNA isolation manner using the guanidinium method. The RNA thus obtained was quantified with a spectrophotometer and stored at a −70° C. freezer until use.

EXAMPLE 3

Comparison of Gene Expression Levels by RT-PCR

[0125] The colon cancer-specific, overexpressed genes selected in Example 1 were subjected to RT-PCR.

[0126] An overall DNA sequence of each gene was obtained from the NCBI Core Nucleotide database (Core Nucleotide, http://www.ncbi.nlm.nih.gov/). Based on the DNA sequences, primer sequences for the genes were designed using the Primer3 program. PCR was performed with these designed primers to examine expression levels of the genes. Base sequences of the primers are listed in Table 6, below.

TABLE 6

[Table 6]
[Table]

AZGP1		Co15A2	
SEQ ID NO. 10	forward ctctgcggaaat-acctgaaa	SEQ ID NO. 20	forward gacctcgtggt-gacaaaggt
SEQ ID NO. 11	Reverse tgaagaa-catctccccgtaa	SEQ ID NO. 21	Reverse agccgcct-gatcttcagtaa
CXCL3		**S100P**	
SEQ ID NO. 12	forward ggtgctccccttgttcag	SEQ ID NO. 22	forward agacagc-catgggcatgat
SEQ ID NO. 13	Reverse agggaattcacctcaaga	SEQ ID NO. 23	Reverset catttgagtcct-gccttctc
CXCL6		**EGFL6**	
SEQ ID NO. 14	forward agatccctggacccagta	SEQ ID NO. 24	forward gcatgaaaaa-gaaggcaaaa
SEQ ID NO. 15	Reverse ttgccaaagggttcaata	SEQ ID NO. 25	Reverse tgtcattcttcagggctttc
AGT		**CTHRC1**	
SEQ ID NO. 16	forward gctgcaaaacttgacacc	SEQ ID NO. 26	forward tcatcg-cacttcttctgtgga
SEQ ID NO. 17	Reverse attgcctgtagcctgtca	SEQ ID NO. 27	Reverse gccaaccca-gatagcaacatc
FCGR3A		**β-actin (Control)**	
SEQ ID NO. 18	forward gcttgttgggag-taaaaatg	A	forward gatcattgctc-ctcctgagc
SEQ ID NO. 19	Reverse tccagtcttgttgagctt	B	Reverse actcct-gcttgctgatccac

[0127] Through RT-PCR, the mRNA isolated from the tissues and cell lines of Example 2 were converted into cDNA. In this regard, the cDNA construction was accomplished using a cDNA synthesis kit (AccuScript High Fidelity 1st Stand cDNA Synthesis Kit, STRATAGENE).

[0128] From the cDNA, PCR amplification was carried out in the presence of the designed primers (1st cycle: 94° C., 5 min; 2nd to 35th cycles: 94° C., 40 sec, 56° C., 40 sec, 72° C., 30 sec; final extension: 72° C., 7 min).

[0129] As a result, differences in gene expression level between normal colon cells and colon cancer cells were detected. Coincident with the results of Example 1, the genes of SEQ ID NOS. 1 to 9 was identified to increase their expression levels in the colon cancer cell lines as compared to the normal colon cells (FIGS. **1** to **3**).

EXAMPLE 4

Comparison of Protein Expression Levels in Sera Using Western Blotting

[0130] Protein levels in serum of colon cancer patients and healthy persons were compared using a Western blotting method.

[0131] Sera was isolated from the blood of colon cancer patients and healthy persons and diluted with the same volume of a sample buffer (125 mM Tris pH 6.8, 4% SDS, 10% glycerol, 0.006% bromophenol blue, 1.8% BME) before boiling. 12% SDS-PAGE separated serum proteins. The SDS-PAGE gel in which the serum proteins were separated according to sizes was brought into contact with a nitrocellulose membrane. The application of a current to the gel-membrane associate transferred the proteins onto the membrane which was then blocked for 1 hour in a TBST solution (10 mM Tris, 100 mM NaCl, 0.05% Tween 20) containing 3% FBS albumin, followed by reaction with an AGT antibody (R&D, 1:2000) at 4° C. with shaking overnight. Afterwards, excess antibodies were washed off with PBST, and a horse radish peroxydase-conjugated secondary antibody (ABCAM, Rabbit polyclonal to Mouse IgG) was added and incubated at 4° C. for 1 hour with shaking. The nitrocellulose membrane was immersed in a mixture of 1:1 ECL Solution A (containing Luminol and enhancer):Solution B (containing hydrogen peroxide) and incubated for 1 min with shaking. After being dried suitably, the membrane was attached to a film cassette and developed in a dark room. The same procedure was applied to EGFL6 and CXCL-3.

[0132] The results are shown in FIG. **4**. AGT, EGFL6, and CXCL-3 proteins were not or little detected in healthy persons (lanes 1 to 3) while being overexpressed in patients with colon cancer (lanes 4 to 7), demonstrating the usefulness thereof as colon cancer diagnosis markers (FIG. **4**)

EXAMPLE 5

Comparison of Protein Expression Levels in Tissues Using ImmunoStaining

[0133] Tissue slides were immunostained so as to determine the presence and expression positions of the proteins in normal colonic epithelial tissues and colon cancer tissues.

[0134] To this end, first, normal colonic epithelial cell tissues and colon cancer cell tissues were surgically excised from colon caner patients and embedded in paraffin blocks. Using a microtome, these blocks were cut into slices of 5 μm thickness, followed by the attachment of the slices to glass slides. The tissue slides thus obtained were immunostained, and observed for the presence and positions of the proteins in tissues under a microscope. The antibodies used in this immunostaining were anti-EGFL-6-antibody (Santa Cruz, 1:2000), anti-CTHRC1-antibody (SANTA CRUZ, 1:1000), and anti-CXCL-3-antibody (Aviva, 1:2000), and anti-AGT-antibody (R&D, 1:2000).

[0135] As a result, it was confirmed by immunohistological staining that EGFL6 and CTHRC1 were expressed in cell membrane and cytoplasm and more strongly expressed in colon cancer suspected tissues than in normal mucous membrane. CXCL3 was detected in cytoplasm or nucleus of tumor cells and AGT was expressed in cytoplasm and cell membrane of tumor cells and in endothelial cells as well (FIGS. **5** to **8**).

EXAMPLE 6

Measurement of Protein Levels in Sera by Immunodot Analysis

[0136] Sera from healthy persons and colon cancer patients were compared for secretion levels of the proteins AGT, EFGL6, CXCL3, COL5A2, CTHRC1, and FCGR3A using an immunodot assay with polyclonal antibodies. Each of the serum samples (10 samples per person) of a 5 to 10-fold dilution was dotted in an amount of 2 μl on a nitrocellulose membrane, dried at room temperature, and blocked in 1% BSAT (bovine serum albumin in Tris-buffered saline) solution. They were treated with a polyclonal antibody to AGT (R& D, 1:5000), a polyclonal antibody to EFGL6 (SantaCruz, 1:5000), a polyclonal antibody to CXCL3 (Aviva, 1:5000), a polyclonal antibody to Col5A2 (1:5000), a polyclonal antibody to CTHRC1 (SantaCruz, 1:1000), and a polyclonal antibody to FCGR3A (1:5000) and then with a horse radish peroxidase-conjugated secondary antibody (1:10000), followed by developing in a DAB solution (0.5 mg/ml, diaminobenzidine in PBT). Fluorescence data obtained by scanning were analyzed (FIGS. **9** and **10**).

[0137] It was found to be expressed in larger amounts in colon cancer sera than in normal sera, demonstrating that these genes can be used as effective markers for the diagnosis and prognosis of colon cancer.

EXAMPLE 7

Establishment of ELISA System and Diagnosis of Colon-Cancer Thereby

[0138] 7-1. Establishment of ELISA System

[0139] Monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were diluted to a concentration of 1 μl/ml in 0.1M carbonate buffer (pH 9.6) and plated in an amount of 100 μl/well into 96-well microtiter plates. After incubating overnight at 4° C., the microtiter plates thus coated with the monoclonal antibodies were washed three times with 0.05% Tween-20-containing PBS (PBS-T). Blocking at room temperature for 2 hours with 1% BSA was followed by three rounds of washing with PBS-T. Each dilution of the proteins corresponding to SEQ ID NOS. 1 to 9 was added in an amount of 100 μl to the 96-well microtiter plates and incubated at room temperature for 2 hours, followed by washing three times with PBS-T. Polyclonal antibodies (1:2000 dilution) to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were added in an amount of 100 μl to the 96-well microtiter plates, incubated for 2 hours and washed. 100 μl of a 200-fold diluted, horse radish peroxidase-conjugated secondary antibody was added, incubated at room temperature for 1 hour and washed three times, followed by color development with TMB. Absorbance at 450 nm was read in an ELISA reader (Molecular Device, Sunnyvale, Calif., USA) (FIG. **11**).

[0140] 7-2. Measurement of Protein Levels in sera Using ELISA System.

[0141] Using the ELISA system established in Example 7-1, serum samples were measured for levels of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins. After being diluted five folds, normal and colon cancer sera were calculated for concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

EXAMPLE 8

Kit Construction and Measurement of Protein Level in Serum

[0142] 8-1. Sandwich ELISA Kit

[0143] A kit for measuring concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins was constructed using the following components:

[0144] A. Solid phase antibody: A microtiter plate with an antibody adsorbed thereto. It was constructed by plating polyclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P in an amount of 100 μl per well into a microtiter plate, followed by incubating overnight at 4° C. to adsorb albumin to the solid phase surface.

[0145] B. Detection antibody: Monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P

[0146] C. Enzyme-conjugated antibody: horse radish peroxidase (HRP)-conjugated secondary antibody

[0147] D. Serum dilution buffer

[0148] E. Substrate (TMB)

[0149] F. Washing solution: 0.05% Tween-containing PBS (PBS-T)

[0150] G. Standard solution: Standard solutions of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

[0151] Using the kit, dilutions of sera taken from patients were assayed for levels of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins as follows.

[0152] A suitable dilution of a serum sample in a diluent (D) was added in an amount of 100 μl per well to the solid phase antibody of the component A, and analyzed for con-

centrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins using the sandwich ELISA kit established with the components B, C and E in Example 8-1.

[0153] 8-2. Immunochromatography Kit

[0154] 8-2-1. Construction of Immunochromatographic Strip

[0155] 1) Preparation of Ab-Gold Conjugate

[0156] An antibody was added in a concentration of 15 μg/ml to a colloidal gold particle solution and then incubated at room temperature for 2 hours with agitation. To this solution was added 1/10 volume of 10% BSA, followed by the incubation of the resulting 1% BSA solution for 1 hour. Centrifugation at 12,000 rpm for 40 min precipitated Ab-gold conjugates. The supernatant was discarded and the precipitates were washed with 2 mM borate buffer. This washing process was repeated three times further. Thereafter, 2 mM borate buffer containing 1% BSA was added in an amount of about 1/10 volume of the gold solution to give a suspension. Absorbance at 530 nm was measured using a UV spectrophotometer and dilution was performed to form an O.D. of 3.00.

[0157] 2) Sample Pad

[0158] Provided for absorbing a sample. Made of a cellulose material. As long as it absorbs samples, any can be used as a material for the sample pad.

[0159] 3) Glass Fiber (GF) Membrane

[0160] Pretreated with 20 mM borate buffer containing sucrose.

[0161] 4) Nitrocellulose (NC) Membrane and Line Treatment

[0162] A nitrocellulose membrane (Millipore) was cut into a suitable size (0.7 cm×5 cm). In the cut membranes, goat anti-sheep IgG was applied at a virtual control line about 3.4 m distant from the bottom of the plastic backing while monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were applied to a virtual detection line 2.7 cm distant from the bottom.

[0163] 5) Absorbent Pad

[0164] Made of a cellulose membrane which can absorb materials remaining untreated after the immune response and thus allows the sample solution including analysates to migrate by capillary action.

[0165] 6) Adhesive Plastic Backing

[0166] On an adhesive plastic backing, the sample pad, the GF membrane, the NC membrane and the absorbent pad were laminated, as shown in FIG. **12**, in such a manner as for samples to continuously migrate by capillary action, thus affording an immunochromatographic strip.

[0167] 8-2-2. Result Decision

[0168] 3~5 min after 6~70 μl of a sample (e.g., a mixture of 1:5 (v/v) serum:elution buffer) was loaded on the sample pad, the strip was observed for color development at the control line and the result line and the concentration of the developed color. A positive sample developed red colors at both the control line and the result line. Only the control line was visualized as red for a negative sample.

[0169] 8-3 Luminex Kit

[0170] 8-3-1. Construction of Luminex Kit

[0171] Polyclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were conjugated to beads. A sample dilution was added in an amount of 100 μl and incubated at room temperature for 2 hours, followed by washing three times with PBS-T. Then, they were incubated for 2 hours with 100 μl of each of monoclonal antibodies to the proteins corresponding to SEQ ID NOS. 1 to 9 and washed. An additional one round of incubation was conducted at room temperature for 1 hour with 100 μl of a 2000-fold diluted, PE (phycoerythrin)-conjugated secondary antibody. They were washed three times before measurement in a luminex device. The fluorescence intensities were plotted against concentrations to give a standard curve.

[0172] 8-3-2. Sandwich Luminex Kit

[0173] A luminex kit for measuring concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were constructed using the following components.

[0174] A. Solid phase antibody: fluorescent beads with polyclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins adsorbed thereto.

[0175] B. Detection antibody: monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P

[0176] C. Enzyme-conjugated antibody: peroxidase-conjugated secondary antibody

[0177] D. Serum dilution buffer

[0178] F. Washing solution: 0.05% Tween-containing PBS (PBS-T)

[0179] G. Standard solution: Standard solutions of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

[0180] Using the kit, dilutions of sera taken from patients were assayed for proteins as follows. A suitable dilution of a serum sample in a diluent (D) was added in an amount of 100 μl per well to the solid phase antibody of the component A, and analyzed for concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins using the components B, C and E.

[0181] 8-4. Protein Microarray Kit

[0182] 8-4-1. Protein Microarray System

[0183] Well chips from Proteagen were coated with monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins. The chips were blocked with BSA buffer, incubated at room temperature for 1 hour with 100 μl of a serum dilution, and washed three times with PBS-T. Again, the chips were incubated at 37° C. for 1 hour with 100 μl of each of diluted monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins and washed. An additional one round of incubation was also conducted at room temperature for 0.5 hours with 100 μl of a 2000-fold diluted, Cy3-conjugated secondary antibody. The chips were washed three times before the measurement of fluorescent intensity at 532 nm. The fluorescent intensities were plotted against concentrations to give a standard curve. The protein microarray system thus established was used to determine the serum levels of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

[0184] 8-4-2. Sandwich Protein Microarray Kit

[0185] A sandwich protein microarray kit for measuring concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins were constructed using the following components.

[0186] A. Solid phase antibody: a slide coated with polyclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

[0187] B. Detection antibody: monoclonal antibodies to AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P

[0188] C. Enzyme-conjugated antibody: Cy3-conjugated secondary antibody

[0189] D. Serum dilution buffer

[0190] F. Washing solution: 0.05% Tween-containing PBS (PBS-T)

[0191] G. Standard solution: Standard solutions of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins.

[0192] Using the kit, dilutions of sera taken from colon cancer patients were assayed for proteins as follows. A suitable dilution of a serum sample in a diluent (D) was added in an amount of 100 μl per well to the solid phase antibody of the component A, and analyzed for concentrations of AZGP1, EGFL6, CTHRC1, CXCL3, CXCL-6, AGT, FCGR3A, Col5A2, and S100P proteins using the components B, C and E in the same manner as in the sandwich method of Example 8-4-1.

INDUSTRIAL APPLICABILITY

[0193] As described hitherto, the present invention provides diagnostic markers for accurately diagnosing colon cancer at an early stage and determining the metastasis and prognosis of colon cancer, thus affording data useful in the treatment and monitoring of colon cancer.

[0194] With ability to determine mRNA or protein expression levels of genes specific to colon cancer readily and rapidly, the colon cancer diagnosis markers of the present invention can also be used in research for developing anticancer agents against colon cancer.

[0195] Although the preferred embodiment(s) of the present invention have(has) been disclosed for illustrative purposes, those skilled in the art will appreciate that various modifications, additions and substitutions are possible, without departing from the scope and spirit of the invention as disclosed in the accompanying claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 29

<210> SEQ ID NO 1
<211> LENGTH: 1247
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

gataatatct gtgcctcctg cccagaaccc tccaagcaga cacaatggta agaatggtgc 60

ctgtcctgct gtctctgctg ctgcttctgg gtcctgctgt cccccaggag aaccaagatg 120

gtcgttactc tctgacctat atctacactg ggctgtccaa gcatgttgaa gacgtccccg 180

cgtttcaggc ccttggctca ctcaatgacc tccagttctt tagatacaac agtaaagaca 240

ggaagtctca gcccatggga ctctggagac aggtggaagg aatggaggat tggaagcagg 300

acagccaact tcagaaggcc agggaggaca tctttatgga gaccctgaaa gacatcgtgg 360

agtattacaa cgacagtaac gggtctcacg tattgcaggg aaggtttggt tgtgagatcg 420

agaataacag aagcagcgga gcattctgga aatattacta tgatggaaag gactacattg 480

aattcaacaa agaaatccca gcctgggtcc ccttcgaccc agcagcccag ataaccaagc 540

agaagtggga ggcagaacca gtctacgtgc agcgggccaa ggcttacctg gaggaggagt 600

gccctgcgac tctgcggaaa tacctgaaat acagcaaaaa tatcctggac cggcaagatc 660

ctccctctgt ggtggtcacc agccaccagg ccccaggaga aaagaagaaa ctgaagtgcc 720

tggcctacga cttctaccca gggaaaattg atgtgcactg gactcgggcc ggcgaggtgc 780

aggagcctga gttacgggga gatgttcttc acaatggaaa tggcacttac cagtcctggg 840

tggtggtggc agtgccccccg caggacacag ccccctactc ctgccacgtg cagcacagca 900

gcctggccca gcccctcgtg gtgccctggg aggccagcta ggaagcaagg gttggaggca 960

atgtgggatc tcagacccag tagctgccct tcctgcctga tgtgggagct gaaccacaga 1020

aatcacagtc aatggatcca caaggcctga ggagcagtgt ggggggacag acaggaggtg 1080

gatttggaga ccgaagactg ggatgcctgt cttgagtaga cttggaccca aaaaatcatc 1140

-continued

```
tcaccttgag cccaccccca ccccattgtc taatctgtag aagctaataa ataatcatcc   1200

ctccttgcct agcataaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa                 1247


<210> SEQ ID NO 2
<211> LENGTH: 1166
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

gctccgggaa tttccctggc ccggccgctc cgggctttcc agtctcaacc atgcataaaa     60

agggttcgcc gatcttgggg agccacacag cccgggtcgc aggcacctcc ccgccagctc    120

tcccgcttct cgcacagctt cccgacgcgt ctgctgagcc ccatggccca cgccacgctc    180

tccgccgccc ccagcaatcc ccggctcctg cgggtggcgc tgctgctcct gctcctggtg    240

gccgccagcc ggcgcgcagc aggagcgtcc gtggtcactg aactgcgctg ccagtgcttg    300

cagacactgc agggaattca cctcaagaac atccaaagtg tgaatgtaag gtcccccgga    360

ccccactgcg cccaaaccga agtcatagcc acactcaaga atgggaagaa agcttgtctc    420

aaccccgcat cccccatggt tcagaaaatc atcgaaaaga tactgaacaa ggggagcacc    480

aactgacagg agagaagtaa gaagcttatc agcgtatcat tgacacttcc tgcagggtgg    540

tccctgccct taccagagct gaaaatgaaa aagagaacag cagctttcta gggacagctg    600

gaaaggactt aatgtgtttg actatttctt acgagggttc tacttattta tgtatttatt    660

tttgaaagct tgtattttaa tattttacat gctgttattt aaagatgtga gtgtgtttca    720

tcaaacatag ctcagtcctg attatttaat tggaatatga tgggttttaa atgtgtcatt    780

aaactaatat ttagtgggag accataatgt gtcagccacc ttgataaatg acagggtggg    840

gaactggagg gtggggggat tgaaatgcaa gcaattagtg gatcactgtt agggtaaggg    900

aatgtatgta cacatctatt ttttatactt tttttttaaa aaaagaatgt cagttgttat    960

ttattcaaat tatctcacat tatgtgttca acatttttat gctgaagttt cccttagaca   1020

ttttatgtct tgcttgtagg gcataatgcc ttgtttaatg tccattctgc agcgtttctc   1080

tttcccttgg aaaagagaat ttatcattac tgttacattt gtacaaatga catgataata   1140

aaagttttat gaaaaaaaaa aaaaaa                                        1166


<210> SEQ ID NO 3
<211> LENGTH: 1677
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

accccttctt tccacactgc cccctgagtt cagggaattt ccccagcatc ccaaagcttg     60

agtttcctgc cagtcgggag ggatgaatgc agataaaggg agtgcagaag gcacgaggaa    120

accaaagtgc tctgtatcct ccagtctccg cgcctccacc cagctcagga acccgcgaac    180

cctctcttga ccactatgag cctcccgtcc agccgcgcgg cccgtgtccc gggtccttcg    240

ggctccttgt gcgcgctgct cgcgctgctg ctcctgctga cgccgccggg gcccctcgcc    300

agcgctggtc ctgtctctgc tgtgctgaca gagctgcgtt gcacttgttt acgcgttacg    360

ctgagagtaa accccaaaac gattggtaaa ctgcaggtgt tccccgcagg cccgcagtgc    420

tccaaggtgg aagtggtagc ctccctgaag aacgggaagc aagtttgtct ggacccggaa    480
```

-continued

```
gccccttttc taaagaaagt catccagaaa attttggaca gtggaaacaa gaaaaactga    540

gtaacaaaaa agaccatgca tcataaaatt gcccagtctt cagcggagca gttttctgga    600

gatccctgga cccagtaaga ataagaagga agggttggtt tttttccatt ttctacatgg    660

attccctact ttgaagagtg tgggggaaag cctacgcttc tccctgaagt ttacagctca    720

gctaatgaag tactaatata gtatttccac tatttactgt tattttacct gataagttat    780

tgaacccttt ggcaattgac catattgtga gcaaagaatc actggttatt agtctttcaa    840

tgaatattga attgaagata actattgtat ttctatcata cattccttaa agtcttaccg    900

aaaaggctgt ggatttcgta tggaaataat gttttattag tgtgctgttg agggaggtat    960

cctgttgttc ttactcactc ttctcataaa ataggaaata ttttagttct gtttcttggg   1020

gaatatgtta ctctttaccc taggatgcta tttaagttgt actgtattag aacactgggt   1080

gtgtcatacc gttatctgtg cagaatatat ttccttattc agaatttcta aaaatttaag   1140

ttctgtaagg gctaatatat tctcttccta tggttttaga cgtttgatgt cttcttagta   1200

tggcataatg tcatgattta ctcattaaac tttgattttg tatgctattt tttcactata   1260

ggatgactat aattctggtc actaaatata cactttagat agatgaagaa gcccaaaaac   1320

agataaattc ctgattgcta atttacatag aaatgtattc tcttggtttt ttaaataaaa   1380

gcaaaattaa caatgatctg tgctctgaaa gttttgaaaa tatatttgaa caatttgaat   1440

ataaattcat catttagtcc tcaaaatata tatagcattg ctaagatttt cagatatcta   1500

ttgtggatct tttaaaggtt ttgaccattt tgttatgagg aattatacat gtatcacatt   1560

cactatatta aaattgcact tttatttttt cctgtgtgtc atgttggttt ttggtacttg   1620

tattgtcatt tggagaaaca ataaaagatt tctaaaccaa aaaaaaaaaa aaaaaaa      1677
```

```
&lt;210&gt; SEQ ID NO 4
&lt;211&gt; LENGTH: 15628
&lt;212&gt; TYPE: DNA
&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 4
```

```
tttaaagcct tacgtagaag atcccccagc tgatagtcag ccttgggcat ggattaaggg     60

cttttaacca atcttgcaac aagtttaagc agatattctt tattgggtcc aatctaacca    120

aaattatttt cttatgttct ccccagtaac gtgtcattat taagagaagt ttggcttgct    180

tagaggccaa atttagaggg tcctgaaatt ttattttctt ttacaccact ttccagcatg    240

ttacctgatc agttgtttat tatctttgct gttgaatgga gtgatcattc caagggcccg    300

aggcaggagg cccaggcaca gtggaaactc tcccaaagac caggatcttt gttttgttcc    360

ctgacatatg ctgagcacca ggaatagtga atgaatgaaa caaattgtga ggctttaaag    420

agccgaaata tttaaacact gggcacaagg ttgttgctta atcagtgcta gatccttacc    480

tccccccttgt gtccaggtcg acttgttact gcagttaaac cacttgctga tcctcaaaca    540

actagttagt ggcacagcca ggcctaggac cccagtctct actgttccaa ctaacccatt    600

cgcaggcagg agcactttga atggtctctt attttaaaaa aattaaatta aaattgtcta    660

tttatttaga gacagagtct tactctgtag cccaggctcg agtgcagtgg tgcaatcata    720

gctcactgta acctccatct cctggcctca aaaagtgttt gaattacaga tgcgaggcac    780

tgtacctggc ccgaatgttc tgttcagaca aagccacctc taagtcgctg tggggcccca    840
```

-continued

```
gacaagtgat ttttgaggag tccctatcta taggaacaaa gtaattaaaa aaatgtattt    900
cagaatttac aggcccatgt gagatatgat ttttttaaat gaagatttag agtaatgggt    960
aaaaaagagg tatttgtgtg tttgttgatt gttcagtcag tgaatgtaca gcttctgcct   1020
catatccagg caccatctct tcctgctctt tgttgttaaa tgttccattc ctgggtaatt   1080
tcatgtctgc catcgtggat atgccgtggc tccttgaacc tgcttgtgtt gaagcaggat   1140
cttccttcct gtcccttcag tgccctaata ccatgtattt aaggctggac acatcaccac   1200
tcccaacctg cctcacccac tgcgtcactt gtgatcactg gcttctggcg actctcacca   1260
aggtctctgt catgccctgt tataatgact acaaaagcaa gtcttaccta taggaaaata   1320
agaattataa cccttttact ggtcatgtga aacttaccat ttgcaatttg tacagcataa   1380
acacagaaca gcacatcttt caatgcctgc atcctgaagg cattttgttt gtgtctttca   1440
atctggctgt gctattgttg gtgtttaaca gtctccccag ctacactgga aacttccaga   1500
aggcactttt cacttgcttg tgtgttttcc ccagtgtcta ttagaggcct ttgcacaggg   1560
taggctcttt ggagcagctg aaggtcacac atcccatgag cgggcagcag ggtcagaagt   1620
ggcccccgtg ttgcctaagc aagactctcc cctgccctct gccctctgca cctccggcct   1680
gcatgtccct gtggcctctt gggggtacat ctcccggggc tgggtcagaa ggcctgggtg   1740
gttggcctca ggctgtcaca cacctaggga gatgctcccg tttctgggaa ccttggcccc   1800
gactcctgca aacttcggta aatgtgtaac tcgaccctgc accggctcac tctgttcagc   1860
agtgaaactc tgcatcgatc actaagactt cctggaagag gtcccagcgt gagtgtcgct   1920
tctggcatct gtccttctgg ccagcctgtg gtctggccaa gtgatgtaac cctcctctcc   1980
agcctgtgca caggcagcct gggaacagct ccatccccac ccctcagcta taaatagggc   2040
atcgtgaccc ggccagggga agaagctgcc gttgttctgg gtactacagc agaaggtaag   2100
ccgggggccc cctcagctcc ttctcggcct tgtctctctc agatgtaact gagctgtggg   2160
ctaggaggaa aaggccggga ggaggcacgg tgatgactga aaaacctctc ccctctcata   2220
agaccagtca tccggacgcg ggctttcccc cactcggtgc ccacctgggg tcttacagga   2280
ggagctgctc ctcctcagca ataggacaag atggtcaggt cttcctgctt ccgctgagaa   2340
aagttagggt cctcaggaac ggagcagact ggtacaggaa cagagtcatc atggccaaga   2400
gtccaccggg tcctcttgcc atcaggagga atagcagggc ttgtgcagga attggggctg   2460
gagggaaggg ccgggctcgg tcagtctcca gctgggatcc ccagagtggt caccctaccc   2520
ctccctcgag acagactgcc tgactgtgtg tcatcaggct ggtcaccgtc tccctgaacc   2580
tcgatttgct cacctataaa atggaactaa taacgatgcc tgggctccct gtctcagggg   2640
ctctggtata gctgaagaga actaatataa catgaaagtg ctttctaagc tttgggataa   2700
gctaaaaggc agattccaat tttattcgag ggcagcgtag attggtgctt cagctcgtgg   2760
atgacagagt caggggggcct ggttctgagt cctagttctg tctcttccca gctgtgtgac   2820
gttgaacaag tcactggacc tctctgttcc tctgcaaaac agcatgaacc aattcattaa   2880
ctacttctcc aggatgcagt aggtcccagg gactatccta ggaatgtggg ctgtattagt   2940
aaacacaaca gcgggaaccc tgttccgggg ctcacattca catcagagca aacagacaaa   3000
gacgctggac agaataagtg cataactaca tggtacagag ggttataagg agggaaaagg   3060
ggagctggat gagagagttg agagtgcccg gtgtggtggg gaaagctgca gggtgaaata   3120
```

-continued

```
ctgcatcagg gaaacctcag ggaaggtgag gactatggtg aggtcagagg ggttgatatg   3180

agaacagtgc cctgcaaatg gcaggcacca caggagcatg agccgtcatc ttcaccttta   3240

gcattcagcc cgggagaagt agggagacat agaaggggca ggtgctggcc aagaggcagg   3300

ggcaggagag gagaaggcgg aggggcactc agggcgaggg tgtcaggccc gccaccccag   3360

agcaccatta ctcccaggac gcggctgcgt gcagacctgg aaccagccta gggagcagcc   3420

gcagatcaca actgagaaca aacgacagtc tctgcctcaa aaatggccca tggaattgcg   3480

tctctggaga cgctgcctga gcaggagcag cacagtgagc gggctgcatc gaccagcgcc   3540

atccaaaccc cgaacagttg gcgcttgtca ggcaggactt cccagcagtc ggttcccaca   3600

ggtttcccct gttgacctga tttgatgtga ctgtctagat taggtgtgaa ctggtggctt   3660

aggcttctct gcacagaaag gcctgcaagc agcagagaga gttttctgtt ccatttttcc   3720

atgtcatgtg gctcttcctg agaacagcgg atggagtcaa atgcatgggg agtggggtga   3780

gatggtagct gaggtcagaa tttggcattt gaatgactga agcagaacaa aacacaccag   3840

gtacttcagc agctgcaccg tgttgagggc aggtgctggt tacgggtctg ggtgagggaa   3900

gccagctgcc aatgtaagaa gaatgactgg gtatgcttag atgaagcaga aaaatctagg   3960

catcaaggtg gccttgagtc agtgatgaca cgctacagct ccaaggaagc ctggcctagc   4020

cctgggggga cagaaaaggc caagaagtga cgatattgca gtacaccccc ctccacaaga   4080

aatgagtgag atgtggtaca aaatgttaga attgaatgaa tcaatagaat aaacgttcat   4140

cccttcaatc aagaagagtc agatgaaatg aattagcagg gccagcccaa gaacctcttc   4200

tgggggtctc agggtagctt tcatttgtag cagctgaggc tgaagcccag ctgcaaggcc   4260

tttgagagaa cgtggtgctg gacccgtgtc tagggcaggg gttctaaacc ctgcttacat   4320

atcagagtca cctgagaatt ttctattttt tttttttttt ttttatacgt ggtcccagca   4380

cagactaagg aatccaacta tcattgggca agccatgcta ggtatgcatg cctttggggc   4440

tctgcagggg atagcgctat gcagggatgg ttgagagctg gttttggggt tgagacacgt   4500

gggaaatact tggactttgg gctgagcctg tggtgctcaa tcccggctgc atgttgggac   4560

cacagggaga tgacaaaacc atccccagcc ctcaccctag ggccctcgaa tgagcatctc   4620

aggggtctag gaggcctcca caaagaccta ctgattggca cacacttgtt tctctaggaa   4680

gagaacttac agctgcaggc aggagcatgt cttaatctgc ttgggctgcc ataagtacca   4740

cagactggga gggtttaaca acagaaatgt gttatctcac agttctggaa gctagaagcc   4800

tgggagccag ccatcagcag agttggtttc ctctgggtcc tctatccttg gcttgtagat   4860

ggccgtcttc tctctgtgtc cccacatggt cttccctctg tgtccccaca tggtcttccc   4920

tctgtgtgtg tccatgtcct catctcctct tctcataagg acacaggtca tattagatca   4980

gggctcaccc tcatggcctc attttaactt aatcatctct ttaaagatcc tgtctccaaa   5040

taatggtcac attctgaggt cctggggttg aggacttcaa cacgggcatt atggccgttg   5100

ggggaggtag gacataattc agctgatatt ggtgcatttt gcacttggat catgtagata   5160

ttttccatgg agctttgaat ccatttcttc tttttttgt agacatgaat ggatttattc   5220

tgggctaaat ggtgacaggg aatattgaga caatgaaaga tctggttaga tggcacttaa   5280

aggtcagtta ataaccacct ttcacccttt gcaaaatgat atttcagggt atgcggaagc   5340

gagcacccca gtctgagatg gctcctgccg gtgtgagcct gagggccacc atcctctgcc   5400
```

-continued

```
tcctggcctg ggctggcctg gctgcaggtg accgggtgta catacacccc ttccacctcg   5460
tcatccacaa tgagagtacc tgtgagcagc tggcaaaggc caatgccggg aagcccaaag   5520
accccacctt catacctgct ccaattcagg ccaagacatc ccctgtggat gaaaaggccc   5580
tacaggacca gctggtgcta gtcgctgcaa aacttgacac cgaagacaag ttgagggccg   5640
caatggtcgg gatgctggcc aacttcttgg gcttccgtat atatggcatg cacagtgagc   5700
tatggggcgt ggtccatggg gccaccgtcc tctccccaac ggctgtcttt ggcaccctgg   5760
cctctctcta tctgggagcc ttggaccaca cagctgacag gctacaggca atcctgggtg   5820
ttccttggaa ggacaagaac tgcacctccc ggctggatgc gcacaaggtc ctgtctgccc   5880
tgcaggctgt acagggcctg ctagtggccc agggcagggc tgatagccag gcccagctgc   5940
tgctgtccac ggtggtgggc gtgttcacag ccccaggcct gcacctgaag cagccgtttg   6000
tgcagggcct ggctctctat accccgtgtgg tcctcccacg ctctctggac ttcacagaac   6060
tggatgttgc tgctgagaag attgacaggt tcatgcaggc tgtgacagga tggaagactg   6120
gctgctccct gacgggagcc agtgtggaca gcaccctggc tttcaacacc tacgtccact   6180
tccaaggtaa ggcaaacctc tctgctggct ctggccctag gacttagtat ccaatgtgta   6240
gctgagatca gccagtcagg ccttggagat gggcaggggg cagccctgcg gacatacctg   6300
gtgaccaccc ttgagaagtg gggaagtggc tgctccgctg ggtccctgga tgggccgtcc   6360
acctcctgga cctgctgccc tactatgtgc acgactatac aacatccttt ttcttacatc   6420
atttaatccc cttatgatgt ggtgaagagg tatttgtgcc tttgtttacc agtgaagaaa   6480
tagagactcg gagaaacaaa gtgccttgct caagatggca cagccaccag tgggggtcct   6540
gggattgaaa cccacatctc ctggccccac agcccagttc tacactcaga agggtcaggt   6600
tcatatctct tgagaaggtc aggaactggg gtccctggcc catgcagaaa taagcaattg   6660
gcttgcttaa atccctttca tgttaggagg ggcattactg aaaaccctct actacaaaga   6720
ttgttgattt tttttttttt ttttattgag acagggtctt gttctgtcac ccaggctgca   6780
gtgtagtggt gccatcattg ctcactgtag ccttgaactc ctggcctcaa gcgatcctcc   6840
cacctctgcc ttccaaagtg ttgggattaa aggtgtgagc cactgcaccc agccacagat   6900
tgcttaaagc attcatttaa caaatacttg ttgaggattt gctacttgta agactttaag   6960
cctggcatct cagaggaggc cagaggaggg ctgtataggc cctgcctcca ggcttttaaa   7020
ggtcaatggg caaatgccta ggatttggag ctgcagggaa acgtgctcca caaggtaact   7080
cagggaagcc tcggggctct cagaggacag aggtcactgg ggagcggaga gcaggccttg   7140
cctggcagtg agggcaacag ggctggtgaa gctaggagca agcatgatga gcccagcctg   7200
cagagtttgg ggcaaggaac gaggatgggg cggttggctt ggcatgagtg ttgaaccaga   7260
aaatgggcct ggggagggca gagctggaga cactttgaac gccatgcttg gtaggtgtgg   7320
gaatggggac gcgttctgtt cagaggtcat cccggaagcc tgccgtgtgc agactggagg   7380
cagggaggat tgtttgaagg ttacgcaaga gtccaggcac acagtcacgg gaacacgtgc   7440
tcagggagca gctcggcaaa tccatgggtg gggtggggct gaggggtgtg tctaagagac   7500
actgaggagg ctctgtcaag atgttaacct cgtgagggac agagagccag gcgggaggtg   7560
aaagacaaga ctgtggagaa agaggttcag tggcgcatag tgatttttct taccacaaca   7620
acctccttga ggtctttccc ttcgggttca gggagaggtg atagatgggg ggattgctca   7680
```

-continued

```
gccctggcac tgactggtca caggggcaga ggccagcccg agggttgccc ggttgagggt   7740
ggcagcacac tgtgcagggc agagcaggga cacatggact tagcctgctg tccctaggag   7800
aagtgctggg aggagcgctc actgagaagg agggtcctgc agaaggcaaa ggcaagaaag   7860
ccagtggcat ctgaaatggg tctcccttcg aaagagagca catccacctg acccagaccg   7920
cagagccagg ccaggaggaa gaggaggaag aataaaaaag ccaaccacat cgggactcaa   7980
aggaagccca ggatcctcgc cggcctccac cgcatgctgc cctgaccctg ccccacttcc   8040
taactttgct ggcctcagtt tccgtcaaag gaggcagcca cttcctgccc acatggtctg   8100
tccagtgagg agatcggggg ctgtctcggg acctctaggt ttccctttag caatgatgtt   8160
ctatttacat gacctcagca ggcagctaga tgtgtcccac tagagaggac ctgaggatct   8220
ggggcctgat gggctccagg gtaccgtctg cccagtgctt gctgtgctcc tgagcatggg   8280
gcgctggccc tggtggtttc catgacacca ggtcctgact tgacctcgac agatttacct   8340
agcctccgga tgagaatggt gagctgtgca tgtcagacga gcagagggaa gacggcagcc   8400
actctcatgt caaatcccag cgtcttttgg gaggcagctt ccctttttta gtttagtttg   8460
ttggaagaaa agaattgtcc ctttcccccc tctaaactaa aagccttgcc agcccaggtg   8520
ggcagcaccg aggtccctgc agggaacgtg caaggggaac cctgcagttt cccgctcaca   8580
tgcccttccg agactgagtg ctccgaggac tgaggacgag aaatatgcca ggtctgccac   8640
tgccttctta cgagacccgg acccagggga ggcacagcca tgcccagctc ctgcctgcca   8700
gttctgtcct cccagctgcc ctactttcat gctgggacct ccaattcagt acaaagggag   8760
acctcactgt ttctgaacca tctctactca gactcccaag tgccacgtgc ccaggggact   8820
gttctgtgac aaacttatac acaacttcac cctattctcc taagaacaac cgcagaatag   8880
gcctttcagg atgagtggga ggacagccga gggcagggat gtgctagtgt aaggtcgagg   8940
cagagggtgg gctgctgtca tggaaagacc ccaggtaact gcgtcacaca caaatttgtg   9000
tccttctccc acaacgggct ctcccgagtt ctctgtcatc tgcacggccc tgtgagcagg   9060
aggggaaaca gagggctcac ccctgccccc aaggcccagt gtgcaaatcc attcatcaca   9120
acgaggttgt gtgagtctcc ccagtagcaa gggctgctga ggaatggagc cctcgtttcc   9180
ggggcctgcg tggcccactc tgtattctat gactgtgatg ggggagggtg ggggccacag   9240
gacagctggt gggctctgcc atggctgggg ctagacatgg attaaaaagt gagtatgagc   9300
aggggcctct aggagtggtg ggatagtgcg gtggtggcca catgtcattc tacgtgcgtc   9360
caaacctaca gaatgtaaaa caccaggagg gagactcaaa gaaaactatc aactttgagt   9420
gctgaggacg tgtcagtgta ggttcgtcag ttgcaacaaa tgggccacgc tggtgtgaga   9480
tgttgatcac gggggaggct gtgtagtggg ggacaagagt tatatgggaa ctttctgtac   9540
tttctgctcg attttgctgt gaacctaaag tcactctaaa aaataacatc tcttaaattt   9600
tttaaaaagt gagtgtgtca aaccacagcc tttgggtcag gacagttcta ggtttgagtt   9660
gacctggcag gtaccagtgg cttatgtccc ttaaggtgac agatgcaaaa cccccggttt   9720
ggtgcctggc atgttgtgtg tcttgcaggt ggcggttagg gctgcctcag tgaactcaaa   9780
tggctgcatt ttacaggaga aatatttgag ccacacttgc ggtcctgtgg ccaggagaat   9840
gcagagtggc ctgggggggg ccaaggaagg aggctgaggc agggcgaggg gcaggatctg   9900
ggcctttggt gtctgccagc cctcattcct gcccctgtct tgggtgactc ttccctccct   9960
```

-continued

```
gtctcctgtc tggatttcag ggaagatgaa gggcttctcc ctgctggccg agccccagga  10020
gttctgggtg gacaacagca cctcagtgtc tgttcccatg ctctctggca tgggcacctt  10080
ccagcactgg agtgacatcc aggacaactt ctcggtgact caagtgccct tcactgagag  10140
cgcctgcctg ctgctgatcc agcctcacta tgcctctgac ctggacaagg tggagggtct  10200
cactttccag caaaactccc tcaactggat gaagaaactg tctccccggt aggagcctcc  10260
cggtctcccc tggaatgtgg gagccacact gtcctgccca ggctgggggc ggggtgggga  10320
gtagacacac ctgagctgag ccttgggtgc agagcagggc agggccgcgg tggcacgggg  10380
ctgggcaggc ggcctgtgtg tctgtctacc agtcctccat ccagccagca cccagctctc  10440
cagttagtgt ctgtctttca agtgcaggca aggtaaagga ggagaggaag aatgcttttt  10500
ctacacttac acttgcctgg tagttttgga gggggagaaa acattgcaat ccgccctctg  10560
agagaggacc attttggtcc cacacctgac acacagcaca cctgtgacat ccaagagctt  10620
cttggaactg acttgccagg agggttcgga cttcgcgtga gcgggggtgg ggccttctca  10680
gggagcgtcc cttgactcca gaacgccctt gctggcggct ggcggctggg tggggatagg  10740
tgttgttagc tcctctttcc tgctgcaatt cctttccaca gagccctgga ctcaaactac  10800
acatcacccc agatcatcga ggcctggaaa tctgctccca gaggcaggca ttgagtgaca  10860
cgatggcttg acatcaactc tgggtgtttt ttatgtttta aaaattgtga tggtaaaata  10920
tacgtaacaa aatttgccat cgtaaccatt ttcgagtgca cagttcagtg gtactaggcc  10980
cattcacact gttgtgcagc catcaccccc gtccatctcc atttatcttc tcaacttccc  11040
aaactgaagc tctgtcctgc tgaaacacta actctccatt tccccttccc cttggccccg  11100
gcaaccacca cgatgtcctc gaggttcacc catgttgtag cacatgtcag aatgtccttc  11160
cttttgaagg ctgaataata ttccattgca tgtggttacc accttttgtg tatccactca  11220
tccatcgatg gacacgtggg ttgcttccac ctttgagctg ctgtgaatag tgcagtgtac  11280
cctgtaaaca tgggtgtact gtcagctctt ataagtgctt gatacatcac tggaaatgtc  11340
catgggctct gaaggatgcc aaaagatgga agaggctcta tacgaagatc aatcgagttg  11400
acatagcaac gtgtccagca cgaggttgac actgtaccct cctgcctctc tccttttcat  11460
gggtgtcatg tcatcaagaa cactgctgtg gcagtagtaa gacacagtgc attatttcag  11520
agaatagcat ttaaaaatta cccaagtaac acaccttcaa tgcagccaac ctaaaaacag  11580
aatgcaccaa aggacaacca ttcctaggtc ctcatcggta aatcttctat gtccctcaca  11640
tagtattgca aatgacatga aggattttta ttgtaggttt tgctgaaatt ttccccaagg  11700
gggaggatga cttagttggg tgatgggggg agcaaacatc cctgtcgtca gggttgggtg  11760
caaggagcat aagcctgcct ggcctctggg agagccctca ctgtgtggcc tggagccttc  11820
ctaactgtgc atcatctccc caggaccatc cacctgacca tgccccaact ggtgctgcaa  11880
ggatcttatg acctgcagga cctgctcgcc caggctgagc tgcccgccat tctgcacacc  11940
gagctgaacc tgcaaaaatt gagcaatgac cgcatcaggg tgggggaggt atgtgtgagc  12000
ctgtgtctgt gcctgacctg ggttccaagt gtgcacaggg tgggaggcat ggatgtaagg  12060
gacacagagg aggctatggg tggggccagc agggcaagag ggagcggaga gtagggccaa  12120
aggtgggaga gaagtagcca gagcattctg gggccttcca ggtgcagagc agcaaatccc  12180
tccccatccc tgctgtgcct cctcctgcta ggtgtgtgtt ccatggtcct gcttggcctt  12240
```

-continued

```
gccttgcctc agggtcctcc agggttccta tagtggagtt gaaaccggga tgaagacagc  12300
aagcacccct ggacctggtg ccctgggccc agccccttct tcagggaaat gctgagcagc  12360
agacagaatg tccccctgcc atgtggcacc atgcacatct gcagctacca aggatgtgcc  12420
ttgatgttct gggccctgtg ctcagtgctg gggagaaagt gggagttctt acgggggcca  12480
gcgggaagag ccctctgtgc taagttagct aagccctggc actggtgggc catggccaag  12540
ggagccagga attctgcctg ggacatcagg gcagaatgtg aagatgggag gatgtaaggg  12600
gtgtgttagg gaggagccgg catgtgagtt tggccattgt ggccaattaa cggtcatcta  12660
cacacagaca cacccttgcc tacactgagg ggcaggcata cactgtgcat cctcctggca  12720
ggctggaaaa tgtccccctc caggacagtg cacagcacag aggtcctgag cccaccccgg  12780
ccctctagcc ctcagcaccc tgggtcaccc agtgcgccct cagaatgatc ctgatgtctg  12840
ctgctttgca ggtgctgaac agcatttttt ttgagcttga agcggatgag agagagccca  12900
cagagtctac ccaacagctt aacaagcctg aggtcttgga ggtgaccctg aaccgcccat  12960
tcctgtttgc tgtgtatgat caaagcgcca ctgccctgca cttcctgggc cgcgtggcca  13020
acccgctgag cacagcatga ggccagggcc ccagaacaca gtgcctggca aggcctctgc  13080
ccctggcctt tgaggcaaag gccagcagca gataacaacc ccggacaaat cagcgatgtg  13140
tcacccccag tctcccacct tttcttctaa tgagtcgact ttgagctgga aagcagccgt  13200
ttctccttgg tctaagtgtg ctgcatggag tgagcagtag aagcctgcag cggcacaaat  13260
gcacctccca gtttgctggg tttattttag agaatggggg tggggaggca agaaccagtg  13320
tttagcgcgg gactactgtt ccaaaaagaa ttccaaccga ccagcttgtt tgtgaaacaa  13380
aaaagtgttc ccttttcaag ttgagaacaa aaattgggtt ttaaaattaa agtatacatt  13440
tttgcattgc cttcggtttg tatttagtgt cttgaatgta agaacatgac ctccgtgtag  13500
tgtctgtaat accttagttt tttccacaga tgcttgtgat tttgaacaa tacgtgaaag  13560
atgcaagcac ctgaatttct gtttgaatgc ggaaccatag ctggttattt ctcccttgtg  13620
ttagtaataa acgtcttgcc acaataagcc tccaaaaatt ttatctttca tttagcagcc  13680
aaacagatgt atacaattca gcagatagac tgtgcaaacg aaagtgcttt cctggacttt  13740
ggatggaatt tccatgggag gtctgagcca gtacttagca gtcctttgaa gttttaggtg  13800
atgcttttct ctggacactt ccattggtaa gcagtggtgg ccatctgtgt gatggacagg  13860
gggcgggaag agggtgacag ggaaggcccc ataccccatg tggcacctgg gaaaggaacc  13920
aggcagatgg gacttcttcc gtcctggtga cacagggcca gactgctgct ggtattgtgc  13980
cccgggagtg gaaggtagag aaataaatct tcacaaataa atatttgcaa ttttccccca  14040
tctgttgagt gcctctgcct gctcctcctc gatgggatta ggcccacagt tcggaatctt  14100
ggggagagcc aaggaagcgg taggcaccca gtaggcccac ggccgtcggc tgatagcaat  14160
ggtgatgctg tcctacctac ttgtgtaagg cattcgatct tcctcccttc catacatatt  14220
gaaataaata agccgcgcaa tgtgttagct attgatcaga actaaagtga agtcagccac  14280
ggggattaca aatctcggct tctcccctca tgttcctgag agtcttcccc tggttttgaa  14340
cacatctccc tagctcgatg tcaaggtgag ggattctgtc ggcaacagca gtgcccttag  14400
ttgcttcgtc gtaactcccc gtcaccggtt ttattcagtt accttccagt cccactctca  14460
gagcttcctg gcttgttctg ctctcaaagc gggtagagct ggcacacatg gactctccga  14520
```

-continued

```
aacggctgca agatgccaag tttctcggaa gaactggaag cacagagacc agaagtgcct  14580
taaggtctcg ctattcagtg tggcgcttag accggcagtg gcggcagctg ccctgggagc  14640
ttgttagaat gtggcttctc acgcccctcc tggacctaca gagtcagaat ctgcagtttt  14700
acaggaggtc caggcttgga agttgctcgt agagacctga gacagcgcag ccacgtgctg  14760
gaaacaaagc atttaagttt gtgactttat tttaaaaggc agcaggcagt cgacaaacca  14820
atttcttcta cttagaggcg gcttcggctt ctggaagtcg ctaggagtat aaagttgcca  14880
accagcgctg ttctcccgct gttttctgtg cacttataaa tgggaagtta ggtcaggata  14940
gatctctcag ctattacaag gatacaaaat acgaacattc tacaagttac ttaacacaca  15000
cacacacaca cacacacaca cacacacaca caaaattaat tccacaggtc agtttctctg  15060
aaacattttt tcactaaatt ctaagtcttc ctggagttgc aagtgcctat ctcctagaca  15120
aggcaattac tcaccaacta aaatcactgt caatctgaga tttcggctgg gcatgagacc  15180
atggtcaggg gatgctttga acagcctctg aggaaattag tgagtttgaa aaatggaaag  15240
atttttatta ctcacttggc agtaaaacct gatggggaca gacgtcaggc tgtttaagat  15300
cctcagaaga aaaagttgat agtgtgaata ttcctaaatt tgccacacga agatgtacat  15360
gtgattataa ggtgctgttg cagaagcccc tgggggtgtt atgggatata cactatatgg  15420
gccactttac cttcctaaaa tctgaaaaac ttcaactact gaaacatgga ctgaaggttt  15480
tgaatagtgg atggtgaatt tgaataccat cccgtgtgat tttttttct agcagacttt  15540
agttttttag agcagtttta agcccacacc aaaactgaga ggaagataca gcaatttctc  15600
atatacccccc tactaccttc cagtctcc                                    15628
```

<210> SEQ ID NO 5
<211> LENGTH: 2137
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

```
cttgtccact ccagtgtggc atcatgtggc agctgctcct cccaactgct ctgctacttc    60
tagtttcagc tggcatgcgg actgaagatc tcccaaaggc tgtggtgttc ctggagcctc   120
aatggtacag ggtgctcgag aaggacagtg tgactctgaa gtgccaggga gcctactccc   180
ctgaggacaa ttccacacag tggtttcaca atgagagcct catctcaagc caggcctcga   240
gctacttcat tgacgctgcc acagttgacg acagtggaga gtacaggtgc cagacaaacc   300
tctccaccct cagtgacccg gtgcagctag aagtccatat cggctggctg ttgctccagg   360
cccctcggtg ggtgttcaag gaggaagacc ctattcacct gaggtgtcac agctggaaga   420
acactgctct gcataaggtc acatatttac agaatggcaa aggcaggaag tattttcatc   480
ataattctga cttctacatt ccaaaagcca cactcaaaga cagcggctcc tacttctgca   540
gggggcttgt tgggagtaaa aatgtgtctt cagagactgt gaacatcacc atcactcaag   600
gtttgtcagt gtcaaccatc tcatcattct ttccacctgg gtaccaagtc tctttctgct   660
tggtgatggt actccttttt gcagtggaca caggactata tttctctgtg aagacaaaca   720
ttcgaagctc aacaagagac tggaaggacc ataaatttaa atggagaaag gaccctcaag   780
acaaatgacc cccatcccat gggggtaata agagcagtag cagcagcatc tctgaacatt   840
tctctggatt tgcaacccca tcatcctcag gcctctctac aagcagcagg aaacatagaa   900
```

-continued

```
ctcagagcca gatcccttat ccaactctcg acttttcctt ggtctccagt ggaagggaaa    960

agcccatgat cttcaagcag ggaagcccca gtgagtagct gcattcctag aaattgaagt   1020

ttcagagcta cacaaacact ttttctgtcc caaccgttcc ctcacagcaa agcaacaata   1080

caggctaggg atggtaatcc tttaaacata caaaaattgc tcgtgttata aattacccag   1140

tttagagggg aaaaaaaaac aattattcct aaataaatgg ataagtagaa ttaatggttg   1200

aggcaggacc atacagagtg tgggaactgc tggggatcta gggaattcag tgggaccaat   1260

gaaagcatgg ctgagaaata gcaggtagtc caggatagtc taagggaggt gttcccatct   1320

gagcccagag ataagggtgt cttcctagaa cattagccgt agtggaatta acaggaaatc   1380

atgagggtga cgtagaattg agtcttccag ggactctat cagaactgga ccatctccaa    1440

gtatataacg atgagtcctc ttaatgctag gagtagaaaa tggtcctagg aaggggactg   1500

aggattgcgg tggggggtgg ggtggaaaag aaagtacaga acaaaccctg tgtcactgtc   1560

ccaagttgct aagtgaacag aactatctca gcatcagaat gagaaagcct gagaagaaag   1620

aaccaaccac aagcacacag gaaggaaagc gcaggaggtg aaaatgcttt cttggccagg   1680

gtagtaagaa ttagaggtta atgcagggac tgtaaaacca ccttttctgc ttcaatatct   1740

aattcctgtg tagctttgtt cattgcattt attaaacaaa tgttgtataa ccaatactaa   1800

atgtactact gagcttcgct gagttaagtt atgaaacttt caaatccttc atcatgtcag   1860

ttccaatgag gtggggatgg agaagacaat tgttgcttat gaaagaaagc tttagctgtc   1920

tctgttttgt aagctttaag cgcaacattt cttggttcca ataaagcatt ttacaagatc   1980

ttgcatgcta ctcttagata gaagatggga aaaccatggt aataaaatat gaatgataaa   2040

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   2100

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa                            2137
```

<210> SEQ ID NO 6
<211> LENGTH: 6930
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

```
gaccgttgct tggcagacac tggatggtta tgagcctgaa caagctgaaa aggggcagga     60

aaagaagtgg aggcagcatt cttcctattt aaagctgcat cgcttgaaaa aagttttcgc    120

agactgtgct ggagctggtg ctgaaaaagg gggtttgcag aggctgccct ggggctggtg    180

ctgaaagaag agcccacagc tgacttcatg gtgctacaat aacctcagaa tctacttttc    240

actctcagga gaacccacat gtctaatatt tagacatgat ggcaaactgg gcggaagcaa    300

gacctctcct cattcttatt gttttattag ggcaatttgt ctcaataaaa gcccaggaag    360

aagacgagga tgaaggatat ggtgaagaaa tagcctgcac tcagaatggc cagatgtact    420

taaacaggga catttggaaa cctgcccctt gtcagatctg tgtctgtgac aatggagcca    480

ttctctgtga caagatagaa tgccaggatg tgctggactg tgccgaccct gtaacgcccc    540

ctggggaatg ctgtcctgtc tgttcacaaa cacctggagg tggcaataca aattttggta    600

gaggaagaaa gggacaaaag ggagaaccag gattagtgcc tgttgtaaca ggcatacgtg    660

gtcgtccagg accggcagga cctccaggat cacagggacc aagaggagag cgagggccaa    720

aaggaagacc tggccctcgt ggacctcagg gaattgatgg agaaccaggt gttcctggtc    780
```

-continued

```
aacctggtgc tccaggacct cctggacatc cgtcccaccc aggacccgat ggcttgagca    840
ggccgttttc agctcaaatg gctgggttgg atgaaaaatc tggacttggg agtcaagtag    900
gactaatgcc tggctctgtg ggtcctgttg gcccaagggg accacagggt ttacaaggac    960
agcaaggtgg tgcaggacct acaggacctc ctggtgaacc tggtgatcct ggaccaatgg   1020
gtccgattgg ttcacgtgga ccagagggcc ctcctggtaa acctggggaa gatggtgaac   1080
ctggcagaaa tggaaatcct ggtgaagtgg gatttgcagg atctccggga gctcgtggat   1140
ttcctggggc tcctggtctt ccaggtctga agggtcaccg aggacacaaa ggtcttgaag   1200
gccctaaagg tgaagttgga gcacctggtt ccaagggtga agctggcccc actggtccaa   1260
tgggtgccat gggtcctctg ggtccgaggg gaatgccagg agagagaggg agacttgggc   1320
cacagggtgc tcctggacaa cgaggtgcac atggtatgcc tggaaaacct ggaccaatgg   1380
gtcctcttgg gataccaggc tcttctggtt ttccaggaaa tcctggaatg aagggagaag   1440
caggtcctac aggggcgcga ggccctgaag gtcctcaggg gcagagaggt gaaactgggc   1500
cccaggtcc agttggctct ccaggtcttc ctggtgcaat aggaactgat ggtactcctg    1560
gtgccaaagg cccaacgggc tctccgggta cctctggtcc tcctggctca gcagggcctc   1620
ctggatctcc aggacctcag ggtagcactg gtcctcaggg aattcgaggc caaccgggtg   1680
atccaggagt tccaggtttc aaaggagaag ctggcccaaa aggggaacca gggccacatg   1740
gtattcaggg tccgataggc ccacccggtg aagaaggcaa aagaggtccc agaggtgacc   1800
caggaacagt tggtcctcca gggccagtgg gagaaagggg tgctcctggc aatcgtggtt   1860
ttccaggctc tgatggttta cctgggccaa agggtgctca aggagaacgg ggtcctgtag   1920
gttcttcagg acccaaagga agccaggggg atccaggacg tccaggggaa cctgggcttc   1980
caggtgctcg ggtttgaca ggaaatcctg gtgttcaagg tcctgaagga aaacttggac    2040
ctttgggtgc gccaggggaa gatggccgtc caggtcctcc aggctccata ggaatcagag   2100
ggcagcccgg gagcatgggc cttccaggcc ccaaaggtag cagtggtgac cctgggaaac   2160
ctggagaagc aggaaatgct ggagttcctg ggcagagggg agctcctgga aaagatggtg   2220
aagttggtcc ttctggtcct gtgggcccgc cgggtctagc tggtgaaaga ggagaacaag   2280
gacctccagg ccccacaggt tttcaggggc ttcctggtcc tccagggcct cctggagaag   2340
gtggaaaacc aggtgatcaa ggtgttcctg gagatcccgg agcagttggc ccgttaggac   2400
ctagaggaga acgaggaaat cctggggaaa gaggagaacc tgggataact ggactccctg   2460
gtgagaaggg aatggctgga ggacatggtc ctgatggccc aaaaggcagt ccaggtccat   2520
ctgggacccc tggagataca ggcccaccag gtcttcaagg tatgccggga gaaagaggaa   2580
ttgcaggaac tcctggcccc aagggtgaca gaggtggcat aggagaaaaa ggtgctgaag   2640
gcacagctgg aaatgatggt gcaagaggtc ttccaggtcc tttgggccct ccaggtccgg   2700
caggtcctac tggagaaaag ggtgaacctg gtcctcgagg tttagttggc cctcctggct   2760
cccggggcaa tcctggttct cgaggtgaaa atgggccaac tggagctgtt ggttttgccg   2820
gaccccaggg tcctgacgga cagcctggag taaaaggtga acctggagag ccaggacaga   2880
agggagatgc tggttctcct ggaccacaag gtttagcagg atcccctggc cctcatggtc   2940
ctaatggtgt tcctggacta aaaggtggtc gaggaaccca aggtccgcct ggtgctacag   3000
gatttcctgg ttctgcgggc agagttggac ctccaggccc tgctggagct ccaggacctg   3060
```

-continued

```
cgggacccct aggggaaccc gggaaggagg gacctccagg tcttcgtggg gaccctggct   3120
ctcatgggcg tgtgggagat cgaggaccag ctggcccccc tggtggccca ggagacaaag   3180
gggacccagg agaagatggg caacctggtc cagatggccc ccctggtcca gctggaacga   3240
ccgggcagag aggaattgtt ggcatgcctg ggcaacgtgg agagagaggc atgcccggcc   3300
taccaggccc agcgggaaca ccaggaaaag taggaccaac tggtgcaaca ggagataaag   3360
gtccacctgg acctgtgggg cccccaggct ccaatggtcc tgtaggggaa cctggaccag   3420
aaggtccagc tggcaatgat ggtaccccag gacgggatgg tgctgttgga gaacgtggtg   3480
atcgtggaga ccctgggcct gcaggtctgc caggctctca gggtgcccct ggaactcctg   3540
gccctgtggg tgctccagga gatgcaggac aaagaggaga tccgggttct cggggtccta   3600
taggaccacc tggtcgagct gggaaacgtg gattacctgg accccaagga cctcgtggtg   3660
acaaaggtga tcatggagac cgaggcgaca gaggtcagaa gggccacaga ggctttactg   3720
gtcttcaggg tcttcctggc cctcctggtc caaatggtga acaaggaagt gctggaatcc   3780
ctggaccatt tggcccaaga ggtcctccag gcccagttgg tccttcaggt aaagaaggaa   3840
accctgggcc acttgggcca attggacctc caggtgtacg aggcagtgta ggagaagcag   3900
gacctgaggg ccctcctggt gagcctggcc cacctggccc tccgggtccc cctggccacc   3960
ttacagctgc tcttggggat atcatggggc actatgatga aagcatgcca gatccacttc   4020
ctgagtttac tgaagatcag gcggctcctg atgacaaaaa caaaacggac ccaggggttc   4080
atgctaccct gaagtcactc agtagtcaga ttgaaaccat gcgcagcccc gatggctcga   4140
aaaagcaccc agcccgcacg tgtgatgacc taaagctttg ccattccgca aagcagagtg   4200
gtgaatactg gattgatcct aaccaaggat ctgttgaaga tgcaatcaaa gtttactgca   4260
acatggaaac aggagaaaca tgtatttcag caaacccatc cagtgtacca cgtaaaacct   4320
ggtgggccag taaatctcct gacaataaac ctgtttggta tggtcttgat atgaacagag   4380
ggtctcagtt cgcttatgga gaccaccaat cacctaatac agccattact cagatgactt   4440
ttttgcgcct tttatcaaaa gaagcctccc agaacatcac ttacatctgt aaaaacagtg   4500
taggatacat ggacgatcaa gctaagaacc tcaaaaaagc tgtggttctc aaaggggcaa   4560
atgacttaga tatcaaagca gagggaaata ttagattccg gtatatcgtt cttcaagaca   4620
cttgctctaa gcggaatgga aatgtgggca agactgtctt tgaatataga acacagaatg   4680
tggcacgctt gcccatcata gatcttgctc ctgtggatgt tggcggcaca gaccaggaat   4740
tcggcgttga aattgggcca gtttgttttg tgtaaagtaa gccaagacac atcgacaatg   4800
agcaccacca tcaatgacca ccgccattca caagaacttt gactgtttga agttgatcct   4860
gagactcttg aagtaatggc tgatcctgca tcagcattgt atatatggtc ttaagtgcct   4920
ggcctcctta tccttcagaa tatttatttt acttacaatc ctcaagtttt aattgatttt   4980
aaatattttt caatacaaca gtttaggttt aagatgacca atgacaatga ccacctttgc   5040
agaaagtaaa ctgattgaat aaataaatct ccgttttctt caatttattt cagtgtaatg   5100
aaaaagttgc ttagtattta tgaggaaatt cttcttcctg gcaggtagct taaagagtgg   5160
ggtatataga gccacaacac atgtttattt tgcttggctg cagttgaaaa atagaaatta   5220
gtgccctttt gtgacctctc attccaagat tgtcaattaa aaatgagttt aaaatgttta   5280
acttgtgatc gagacctaca tgcatgtctt gatattgtgt aactataata gagactcttt   5340
```

-continued

```
aaggagaatc ttaaaaaaaa aaaaacgttt ctcactgtct taaatagaat ttttaaatag   5400
tatatattca gtggcatttt ggagaacaaa gtgaatttac ttcgacttct taaatttttg   5460
taaaagacta taagtttaga catctttctc attcaaattt aaagatatct ttctcctctt   5520
gatcaatcta tcaatattga tagaagtcac actagtatat accatttaat acatttacac   5580
tttcttattt aagaagatat tgaatgcaaa ataattgaca tatagaactt tacaaacata   5640
tgtccaagga ctctaaattg agactcttcc acatgtacaa tctcatcatc ctgaagccta   5700
taatgaagaa aaagatctag aaactgagtt gtggagctga ctctaatcaa atgtgatgat   5760
tggaattaga ccatttggcc tttgaacttt cataggaaaa atgacccaac atttcttagc   5820
atgagctacc tcatctctag aagctgggat ggacttacta ttcttgttta tattttagat   5880
actgaaaggt gctatgcttc tgttattatt ccaagactgg agataggcag ggctaaaaag   5940
gtattattat ttttccttta atgatggtgc taaaattctt cctataaaat tccttaaaaa   6000
taaagatggt ttaatcacta ccattgtgaa aacataactg ttagacttcc cgtttctgaa   6060
agaaagagca tcgttccaat gcttgttcac tgttcctctg tcatactgta tctggaatgc   6120
tttgtaatac ttgcatgctt cttagaccag aacatgtagg tccccttgtg tctcaatact   6180
tttttttct taattgcatt tgttggctct attttaattt ttttctttta aaataaacag   6240
ctgggaccat cccaaaagac aagccatgca tacaactttg gtcatgtatc tctgcaaagc   6300
atcaaattaa atgcacgctt ttgtcatgtc agtggttttt gttttgtgaa attcctttga   6360
ccatattaga tctatttcat ttccaatagt gaaaaggaga tgtggtggta tactttgttt   6420
gccatttgtt taaaagatac aacggatacc ttctatcatg tatgtactgg cttataaatg   6480
aaaatctatc tacaacatta cccacaaagg caacatgaca ccaattatca ctgcctctgc   6540
ccttaaaaat gtcagagtag tattattgat aaaaagggca agcaatagat ttttcatgac   6600
tgaataaact gtaataataa aacatatgtc tcaaagtgta tcacatatga atttagccta   6660
attgttttca gtttcattct caatatttag tttacaacat cattttcccc taaactggtt   6720
atattttgac ctgtatatct taaatttgag tatttatatg cctaaataca tgtgtgagtt   6780
ttgtttgact tccaagtcca aactataaga ttatataagt tcatatagat gaatcagaaa   6840
tatgtggtaa tactattaag tcacaaacac taacaatttc caactataga aataacagtt   6900
cttatttgga ttttgggaat gctaccaata                                    6930
```

<210> SEQ ID NO 7
<211> LENGTH: 510
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

```
tgaggctgcc ttataaagca ccaagaggct gccagtggga cattttctcg gccctgccag     60
cccccaggag gaaggtgggt ctgaatctag caccatgacg gaactagaga cagccatggg    120
catgatcata gacgtctttt cccgatattc gggcagcgag ggcagcacgc agaccctgac    180
caagggggag ctcaaggtgc tgatggagaa ggagctacca ggcttcctgc agagtggaaa    240
agacaaggat gccgtggata aattgctcaa ggacctggac gccaatggag atgcccaggt    300
ggacttcagt gagttcatcg tgttcgtggc tgcaatcacg tctgcctgtc acaagtactt    360
tgagaaggca ggactcaaat gatgccctgg agatgtcaca gattcctggc agagccatgg    420
```

-continued

```
tcccaggctt cccaaaagtg tttgttggca attattcccc taggctgagc ctgctcatgt    480

acctctgatt aataaatgct tatgaaatga                                      510


<210> SEQ ID NO 8
<211> LENGTH: 2013
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

accccgtcca gcttcatccg cagaggagcc tcggccaggc ttgccagggc gcccccagcc     60

cctccccagg ccgcgagcgc ccctgccgcg gtgcctggcc tccccgccca gactgcaggg    120

acagcacccg gtaactgcga gtggagcgga ggacccgagc ggctgaggag agaggaggcg    180

gcggcttagc tgctacgggg tccggccggc gccctcccga gggggctca ggaggaggaa    240

ggaggacccg tgcgagaatg cctctgccct ggagccttgc gctcccgctg ctgctctcct    300

gggtggcagg tggtttcggg aacgcggcca gtgcaaggca tcacgggttg ttagcatcgg    360

cacgtcagcc tggggtctgt cactatggaa ctaaactggc ctgctgctac ggctggagaa    420

gaaacagcaa gggagtctgt gaagctacat gcgaacctgg atgtaagttt ggtgagtgcg    480

tgggaccaaa caaatgcaga tgctttccag gatacaccgg gaaaacctgc agtcaagatg    540

tgaatgagtg tggaatgaaa ccccggccat gccaacacag atgtgtgaat acacacggaa    600

gctacaagtg cttttgcctc agtggccaca tgctcatgcc agatgctacg tgtgtgaact    660

ctaggacatg tgccatgata aactgtcagt acagctgtga agacacagaa gaagggccac    720

agtgcctgtg tccatcctca ggactccgcc tggccccaaa tggaagagac tgtctagata    780

ttgatgaatg tgcctctggt aaagtcatct gtccctacaa tcgaagatgt gtgaacacat    840

ttggaagcta ctactgcaaa tgtcacattg gtttcgaact gcaatatatc agtggacgat    900

atgactgtat agatataaat gaagagaaaa tgaaagaggg gcttgaggat gagaaaagag    960

aagagaaagc cctgaagaat gacatagagg agcgaagcct gcgaggagat gtgtttttcc   1020

ctaaggtgaa tgaagcaggt gaattcggcc tgattctggt ccaaaggaaa gcgctaactt   1080

ccaaactgga acataaagca gatttaaata tctcggttga ctgcagcttc aatcatggga   1140

tctgtgactg gaaacaggat agagaagatg attttgactg gaatcctgct gatcgagata   1200

atgctattgg cttctatatg gcagttccgg ccttggcagg tcacaagaaa gacattggcc   1260

gattgaaact tctcctacct gacctgcaac cccaaagcaa cttctgtttg ctctttgatt   1320

accggctggc cggagacaaa gtcgggaaac ttcgagtgtt tgtgaaaaac agtaacaatg   1380

ccctggcatg ggagaagacc acgagtgagg atgaaaagtg gaagacaggg aaaattcagt   1440

tgtatcaagg aactgatgct accaaaagca tcatttttga agcagaacgt ggcaagggca   1500

aaaccggcga aatcgcagtg gatggcgtct tgcttgtttc aggcttatgt ccagatagcc   1560

ttttatctgt ggatgactga atgttactat ctttatattt gactttgtat gtcagttccc   1620

tggttttttt gatattgcat cataggacct ctggcatttt agaattacta gctgaaaaat   1680

tgtaatgtac caacagaaat attattgtaa gatgcctttc ttgtataaga tatgccaata   1740

tttgctttaa atatcatatc actgtatctt ctcagtcatt tctgaatctt tccacattat   1800

attataaaat atggaaatgt cagtttatct cccctcctca gtatatctga tttgtataag   1860

taagttgatg agcttctctc tacaacattt ctagaaaata gaaaaaaaag cacagagaaa   1920
```

-continued

```
tgtttaactg tttgactctt atgatacttc ttggaaacta tgacatcaaa gatagacttt   1980

tgcctaagtg gcttagctgg gtctttcata gcc                                2013


<210> SEQ ID NO 9
<211> LENGTH: 1236
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

ctgcggcggc ctcggagcgc ggcggagcca gacgctgacc acgttcctct cctcggtctc     60

ctccgcctcc agctccgcgc tgcccggcag ccgggagcca tgcgacccca gggccccgcc    120

gcctccccgc agcggctccg cggcctcctg ctgctcctgc tgctgcagct gcccgcgccg    180

tcgagcgcct ctgagatccc caaggggaag caaaaggcgc agctccggca gagggaggtg    240

gtggacctgt ataatggaat gtgcttacaa gggccagcag gagtgcctgg tcgagacggg    300

agccctgggg ccaatggcat tccgggtaca cctgggatcc caggtcggga tggattcaaa    360

ggagaaaagg gggaatgtct gagggaaagc tttgaggagt cctggacacc caactacaag    420

cagtgttcat ggagttcatt gaattatggc atagatcttg ggaaaattgc ggagtgtaca    480

tttacaaaga tgcgttcaaa tagtgctcta agagttttgt tcagtggctc acttcggcta    540

aaatgcagaa atgcatgctg tcagcgttgg tatttcacat tcaatggagc tgaatgttca    600

ggacctcttc ccattgaagc tataatttat ttggaccaag gaagccctga aatgaattca    660

acaattaata ttcatcgcac ttcttctgtg gaaggacttt gtgaaggaat tggtgctgga    720

ttagtggatg ttgctatctg ggttggcact tgttcagatt acccaaaagg agatgcttct    780

actggatgga attcagtttc tcgcatcatt attgaagaac taccaaaata aatgctttaa    840

ttttcatttg ctacctcttt ttttattatg ccttggaatg gttcacttaa atgacatttt    900

aaataagttt atgtatacat ctgaatgaaa agcaaagcta aatatgttta cagaccaaag    960

tgtgatttca cactgttttt aaatctagca ttattcattt tgcttcaatc aaaagtggtt   1020

tcaatatttt ttttagttgg ttagaatact ttcttcatag tcacattctc tcaacctata   1080

atttggaata ttgttgtggt cttttgtttt ttctcttagt atagcatttt taaaaaaata   1140

taaaagctac caatctttgt acaatttgta aatgttaaga attttttta tatctgttaa   1200

ataaaaatta tttccaacaa aaaaaaaaaa aaaaaa                             1236


<210> SEQ ID NO 10
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AZGP1 forward primer

<400> SEQUENCE: 10

ctctgcggaa atacctgaaa                                                 20


<210> SEQ ID NO 11
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AZGP1 reverse primer

<400> SEQUENCE: 11
```

-continued

```
tgaagaacat ctccccgtaa                                        20


<210> SEQ ID NO 12
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CXCL3 forward primer

<400> SEQUENCE: 12

ggtgctcccc ttgttcag                                          18


<210> SEQ ID NO 13
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CXCL3 reverse primer

<400> SEQUENCE: 13

agggaattca cctcaaga                                          18


<210> SEQ ID NO 14
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CXCL6 forward primer

<400> SEQUENCE: 14

agatccctgg acccagta                                          18


<210> SEQ ID NO 15
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CXCL6 reverse primer

<400> SEQUENCE: 15

ttgccaaagg gttcaata                                          18


<210> SEQ ID NO 16
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AGT forward primer

<400> SEQUENCE: 16

gctgcaaaac ttgacacc                                          18


<210> SEQ ID NO 17
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AGT reverse primer

<400> SEQUENCE: 17

attgcctgta gcctgtca                                          18
```

-continued

<210> SEQ ID NO 18
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: FCGR3A forward primer

<400> SEQUENCE: 18

gcttgttggg agtaaaaatg 20

<210> SEQ ID NO 19
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: FCGR3A reverse primer

<400> SEQUENCE: 19

tccagtcttg ttgagctt 18

<210> SEQ ID NO 20
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Col5A2 forward primer

<400> SEQUENCE: 20

gacctcgtgg tgacaaaggt 20

<210> SEQ ID NO 21
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Col5A2 reverse primer

<400> SEQUENCE: 21

agccgcctga tcttcagtaa 20

<210> SEQ ID NO 22
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: S100P forward primer

<400> SEQUENCE: 22

agacagccat gggcatgat 19

<210> SEQ ID NO 23
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: S100P reverse primer

<400> SEQUENCE: 23

tcatttgagt cctgccttct c 21

<210> SEQ ID NO 24
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: EGFL6 forward primer

<400> SEQUENCE: 24

gcatgaaaaa gaaggcaaaa 20

<210> SEQ ID NO 25
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: EGFL6 reverse primer

<400> SEQUENCE: 25

tgtcattctt cagggctttc 20

<210> SEQ ID NO 26
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CTHRC1 forward primer

<400> SEQUENCE: 26

tcatcgcact tcttctgtgg a 21

<210> SEQ ID NO 27
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: CTHRC1 reverse primer

<400> SEQUENCE: 27

gccaacccag atagcaacat c 21

<210> SEQ ID NO 28
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: beta-actin forward primer

<400> SEQUENCE: 28

gatcattgct cctcctgagc 20

<210> SEQ ID NO 29
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: beta-actin reverse primer

<400> SEQUENCE: 29

actcctgctt gctgatccac 20

1-17. (canceled)

18. A method of diagnosing colon cancer in a subject, comprising:
- a) measuring a level of CXCL3 (C-X-C chemokine ligand 3) mRNA or protein in a biological sample from the subject; and
- b) determining the presence of colon cancer in the subject, wherein an increase in the level of CXCL3 mRNA or protein as compared to a normal control subject indicates the presence of colon cancer.

19. The method according to claim **18**, wherein the biological sample is selected from the group consisting of tissue, cell, whole blood, serum, plasma, saliva, sputum, cerebrospinal fluid and urine.

20. The method according to claim **18**, wherein the level of CXCL3 mRNA is measured by a RT-PCR (reverse transcription-polymerase chain reaction), Competitive RT-PCR, Real-Time RT-PCR, RPA (RNase protection assay), or Northern blotting.

21. The method according to claim **18**, wherein the level of CXCL3 mRNA is measured by using a primer set comprising a forward primer of SEQ. ID. NO: 12 and a reverse primer of SEQ. ID. NO: 13.

22. The method according to claim **18**, wherein the level of CXCL3 protein is measured by an immunodot assay, a luminex assay, an ELISA assay, a protein microarray assay, an immunochromatographic strip assay, or western blot assay.

23. The method according to claim **18**, wherein the level of CXCL3 protein is measured by using an antibody specific to the protein.

24. A method of diagnosing colon cancer in a subject, comprising:
- a) measuring levels of CXCL3 mRNA and one or more additional mRNAs in a biological sample from the subject, wherein the additional mRNAs are selected from the group consisting of AZGP1 (alpha-2-glycoprotein 1, zinc-binding) mRNA, CXCL6 [chemokine (C-X-C motif) ligand 6, granulocyte chemotactic protein 2] mRNA, AGT [angiotensinogen(serpin peptidase inhibitor, clade A, member 8)] mRNA, FCGR3A (Fc fragment of IgG, low affinity Ma, receptor) mRNA, Col5A2 (collagen, type V, alpha 2) mRNA, S100P (S 100 calcium binding protein P) mRNA, EGFL6 (EGF-like-domain, multiple 6) mRNA, and CTHRC 1 (collagen triple helix repeat containing 1) mRNA; and
- b) determining the presence of colon cancer in the subject, wherein an increase in the levels of CXCL3 mRNA and one or more additional mRNAs as compared to a normal control subject indicates the presence of colon cancer.

25. The method according to claim **24**, wherein the biological sample is selected from the group consisting of tissue, cell, whole blood, serum, plasma, saliva, sputum, cerebrospinal fluid and urine.

26. The method according to claim **24**, wherein the levels of mRNAs are measured by a RT-PCR (reverse transcription-polymerase chain reaction), Competitive RT-PCR, Real-Time RT-PCR, RPA (RNase protection assay), or Northern blotting.

27. The method according to claim **24**, wherein the levels of mRNAs are measured by using primer sets selected from the group consisting of the primer sets of following 1)-9):
1) SEQ. ID. NO: 12 (forward) and SEQ. ID. NO: 13 (reverse) for CXCL3;
2) SEQ. ID. NO: 10 (forward) and SEQ. ID. NO: 11 (reverse) for AZGP1;
3) SEQ. ID. NO: 14 (forward) and SEQ. ID. NO: 15 (reverse) for CXCL6;
4) SEQ. ID. NO: 16 (forward) and SEQ. ID. NO: 17 (reverse) for AGT;
5) SEQ. ID. NO: 18 (forward) and SEQ. ID. NO: 19 (reverse) for FCGR3A;
6) SEQ. ID. NO: 20 (forward) and SEQ. ID. NO: 21 (reverse) for Col5A2;
7) SEQ. ID. NO: 22 (forward) and SEQ. ID. NO: 23 (reverse) for S100P;
8) SEQ. ID. NO: 24 (forward) and SEQ. ID. NO: 25 (reverse) for EGFL6; and
9) SEQ. ID. NO: 26 (forward) and SEQ. ID. NO: 27 (reverse) for CTHRC1.

28. A method of diagnosing colon cancer in a subject, comprising:
- a) measuring levels of CXCL3 protein and one or more additional proteins in a biological sample from the subject, wherein the additional proteins are selected from the group consisting of AZGP1 (alpha-2-glycoprotein 1, zinc-binding) protein, CXCL6 [chemokine (C-X-C motif) ligand 6, granulocyte chemotactic protein 2] protein, AGT [angiotensinogen(serpin peptidase inhibitor, clade A, member 8)] protein, FCGR3A (Fc fragment of IgG, low affinity Ma, receptor) protein, Col5A2 (collagen, type V, alpha 2) protein, S100P (S 100 calcium binding protein P) protein, EGFL6 (EGF-like-domain, multiple 6) protein, and CTHRC 1 (collagen triple helix repeat containing 1) protein; and
- b) determining the presence of colon cancer in the subject, wherein an increase in the levels of CXCL3 protein and one or more additional proteins as compared to a normal control subject indicates the presence of colon cancer,

29. The method according to claim **28**, wherein the biological sample is selected from the group consisting of tissue, cell, whole blood, serum, plasma, saliva, sputum, cerebrospinal fluid and urine.

30. The method according to claim **28**, wherein the levels of proteins are measured by an immunodot assay, a luminex assay, an ELISA assay, a protein microarray assay, an immunochromatographic strip assay, or western blot assay.

31. The method according to claim **28**, wherein the levels of proteins are measured by using antibodies specific to the proteins.

* * * * *

patsnap

专利名称(译)	结肠癌诊断试剂盒使用结肠癌相关标记物及其诊断方法		
公开(公告)号	US20110251097A1	公开(公告)日	2011-10-13
申请号	US13/126500	申请日	2009-04-20
[标]申请(专利权)人(译)	韩国生命工学研究院		
申请(专利权)人(译)	韩国研究学院生物科学与生物		
当前申请(专利权)人(译)	韩国研究学院生物科学与生物		
[标]发明人	SONG EUN YOUNG LEE HEE GU YEOM YOUNG IL KIM JAE WHA JI NA YOUNG CHUNG KYUNG SOOK WON MISUN KIM SEON YOUNG KIM JOO HEON KIM YOUNG HO CHUN HO KYUNG		
发明人	SONG, EUN YOUNG LEE, HEE GU YEOM, YOUNG IL KIM, JAE WHA JI, NA YOUNG CHUNG, KYUNG-SOOK WON, MISUN KIM, SEON-YOUNG KIM, JOO HEON KIM, YOUNG HO CHUN, HO KYUNG		
IPC分类号	C40B30/04 G01N33/53 G01N33/566 C12Q1/68		
CPC分类号	C12Q2600/158 C12Q1/6886		
优先权	1020080103387 2008-10-22 KR		
外部链接	Espacenet USPTO		

摘要(译)

本发明涉及用于诊断结肠癌的组合物。该组合物包含至少一种标记物，用于测量至少一种对结肠癌特异的基因的mRNA或蛋白质表达水平。它可以筛选仅在结肠癌组织或血液中特异性过表达的基因。本发明可以同时定量分析基因的mRNA表达水平和由该基因编码的蛋白质的表达水平，从而以高水平的可靠性诊断早期结肠癌。

[Fig. 1]