



US 20110046005A1

(19) **United States**

(12) **Patent Application Publication**  
**ADORJAN et al.**

(10) **Pub. No.: US 2011/0046005 A1**  
(43) **Pub. Date: Feb. 24, 2011**

(54) **METHOD AND NUCLEIC ACIDS FOR THE ANALYSIS OF COLON CELL PROLIFERATIVE DISORDERS**

(30) **Foreign Application Priority Data**

Feb. 27, 2002 (EP) ..... 02004551.4

(75) Inventors: **PETER ADORJAN**, BERLIN (DE); **MATTHIAS BURGER**, BERLIN (DE); **SABINE MAIER**, PRINCETON, NJ (US); **RALF LESCHE**, BERLIN (DE); **SUSAN COTTRELL**, SEATTLE, WA (US); **SUZANNE MOONEY**, SEATTLE, WA (US)

**Publication Classification**

(51) **Int. Cl.**  
*C40B 30/04* (2006.01)  
*C07K 2/00* (2006.01)  
*C07H 21/04* (2006.01)  
*C40B 50/14* (2006.01)  
*C40B 40/06* (2006.01)  
*G01N 33/53* (2006.01)  
*C12Q 1/68* (2006.01)

Correspondence Address:  
**DAVIS WRIGHT TREMAINE, LLP/Seattle**  
**1201 Third Avenue, Suite 2200**  
**SEATTLE, WA 98101-3045 (US)**

(52) **U.S. Cl.** ..... **506/9**; 536/23.1; 530/322; 536/24.33; 506/30; 506/16; 436/94; 435/6

(73) Assignee: **EPIGENOMICS AG**, BERLIN (DE)

(21) Appl. No.: **12/794,449**

(57) **ABSTRACT**

(22) Filed: **Jun. 4, 2010**

Provided are methods and nucleic acids for detecting, differentiating or distinguishing between colon cell proliferative disorders by analysis of one or more of the genes Versican, TPEF, H-Cadherin, Calcitonin, and EYA4. Further provided are novel nucleic acid sequences useful for the cell proliferative disorder specific analysis of said genes as well as methods, assays and kits thereof.

**Related U.S. Application Data**

(63) Continuation of application No. 10/506,089, filed on Sep. 6, 2005, now abandoned, filed as application No. PCT/EP03/02034 on Feb. 27, 2003.

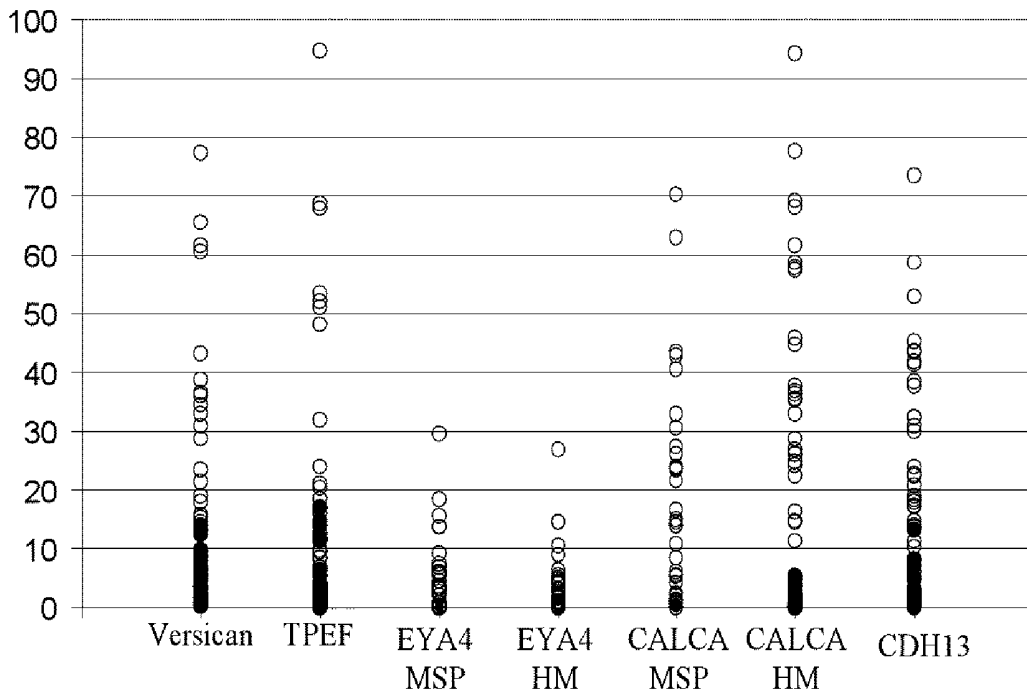


FIGURE 1

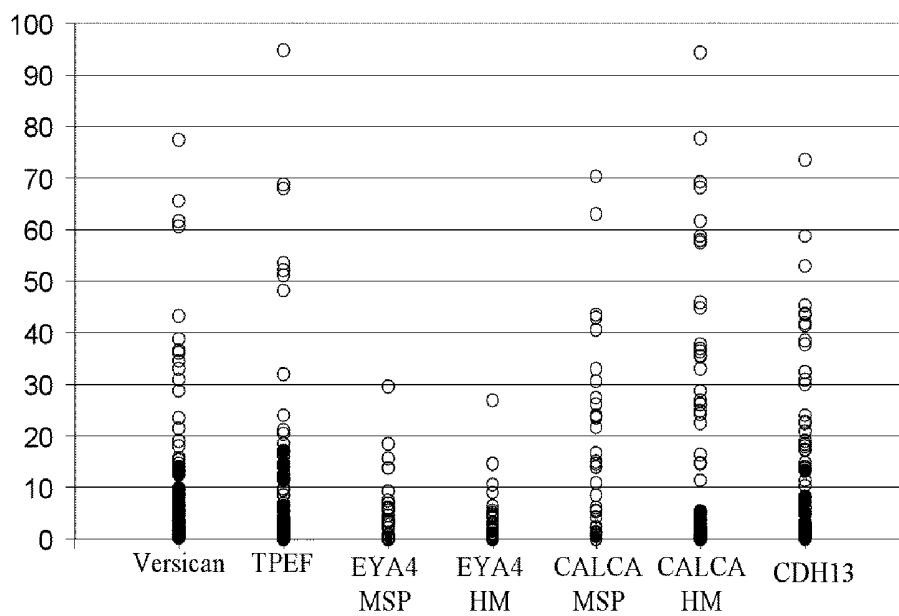


FIGURE 2

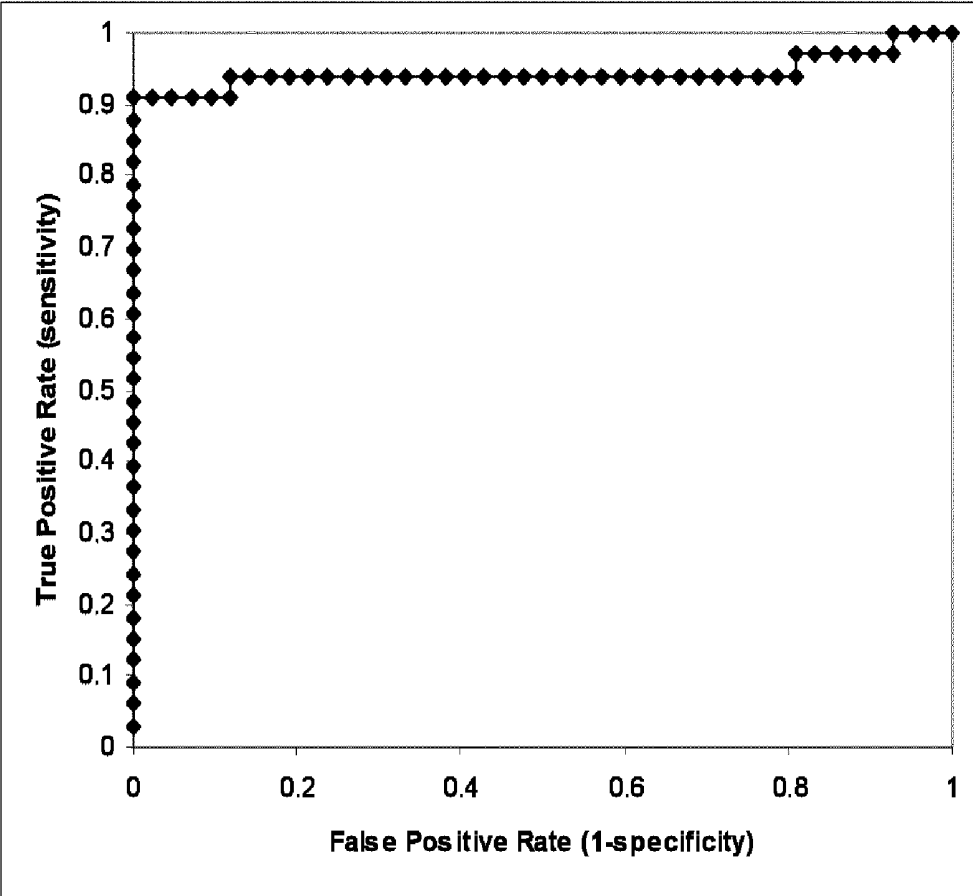


FIGURE 3

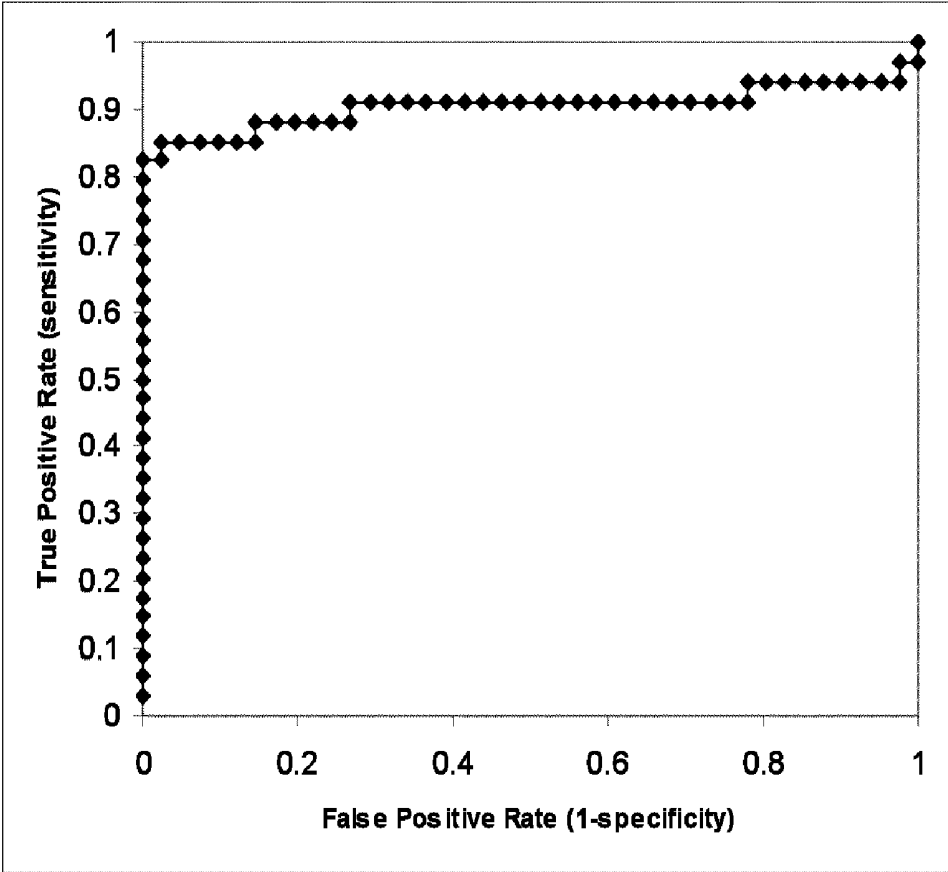


FIGURE 4

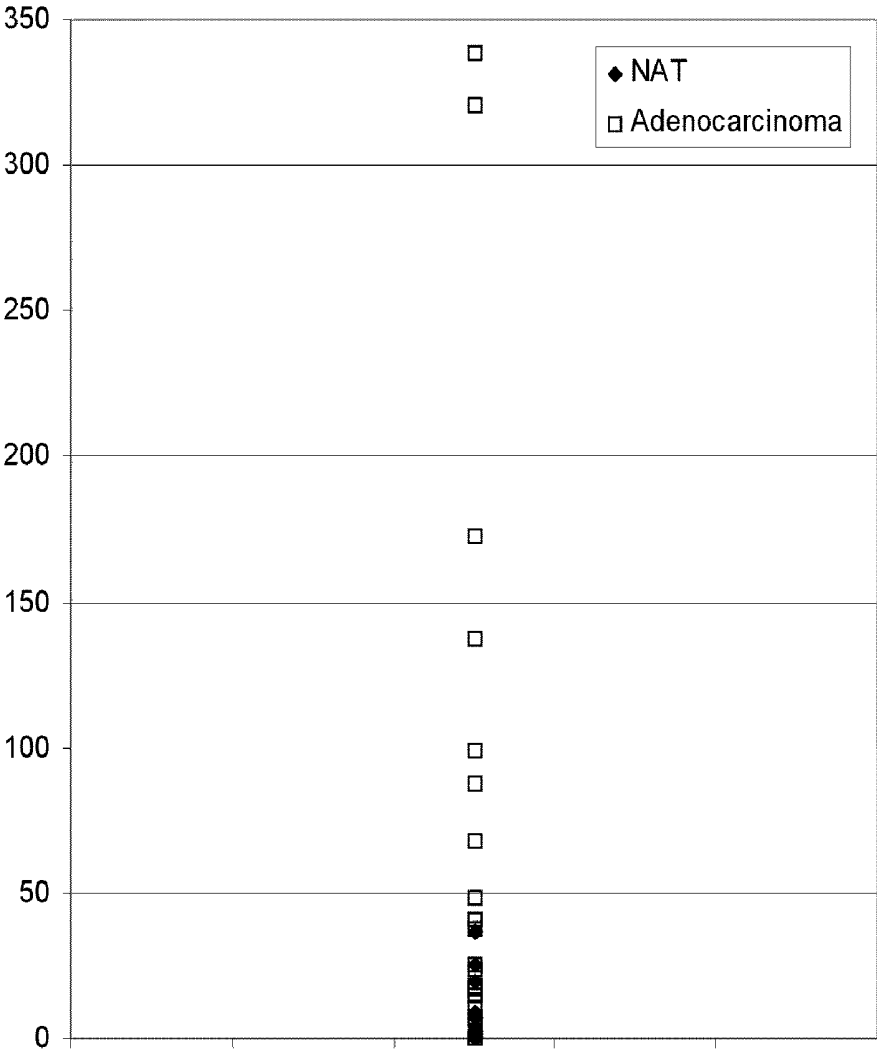


FIGURE 5

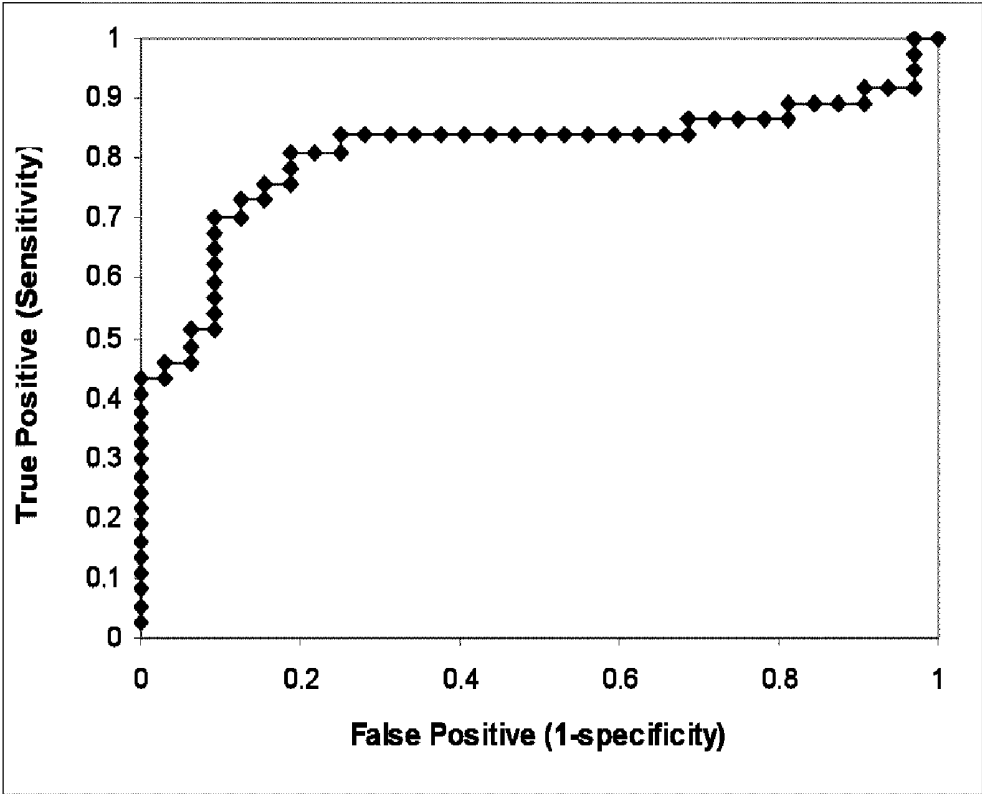


FIGURE 6

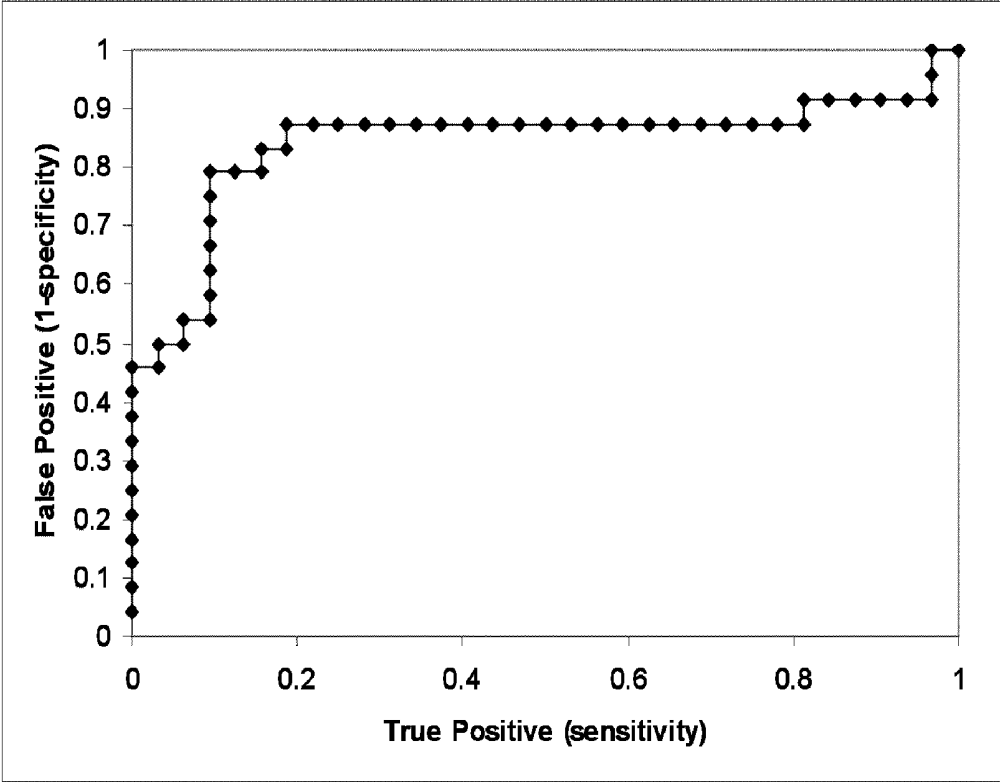


FIGURE 7

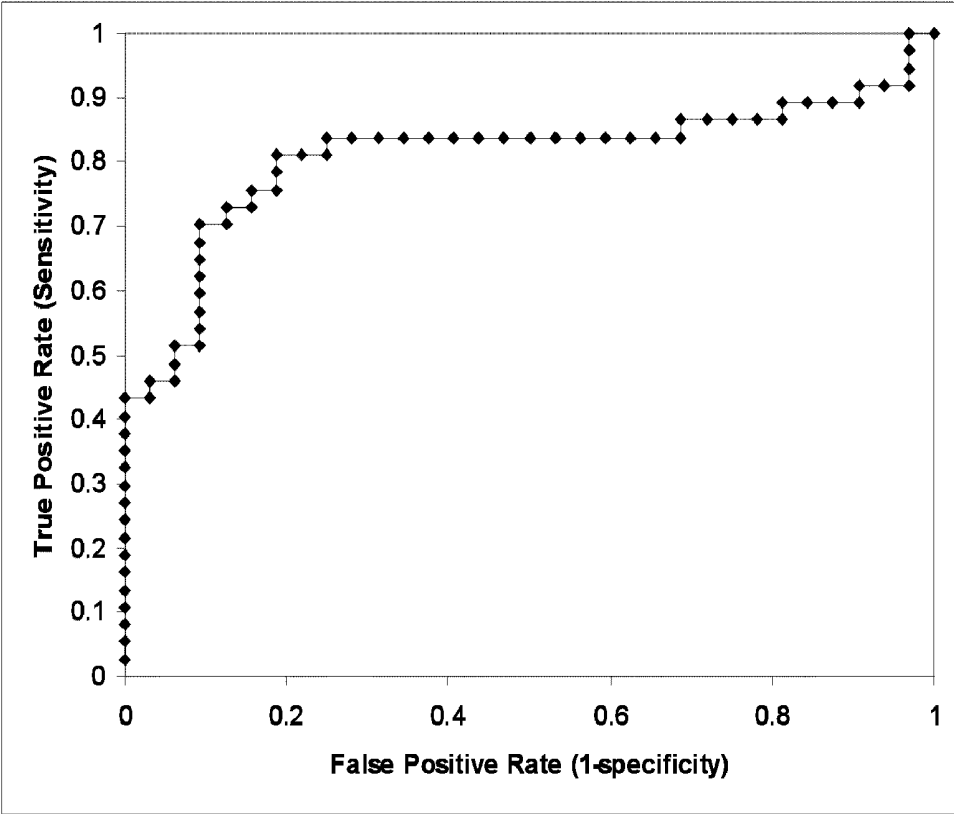


FIGURE 8

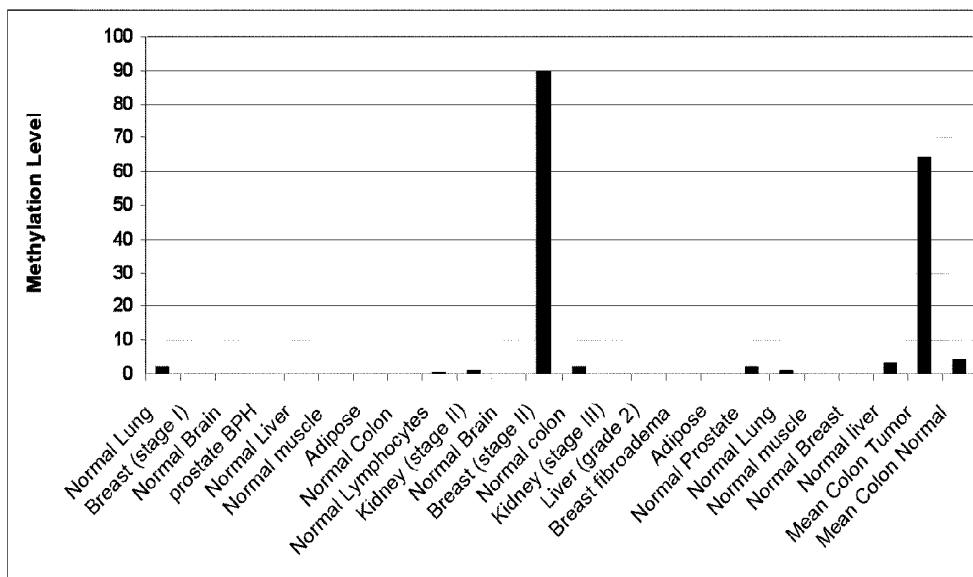


FIGURE 9

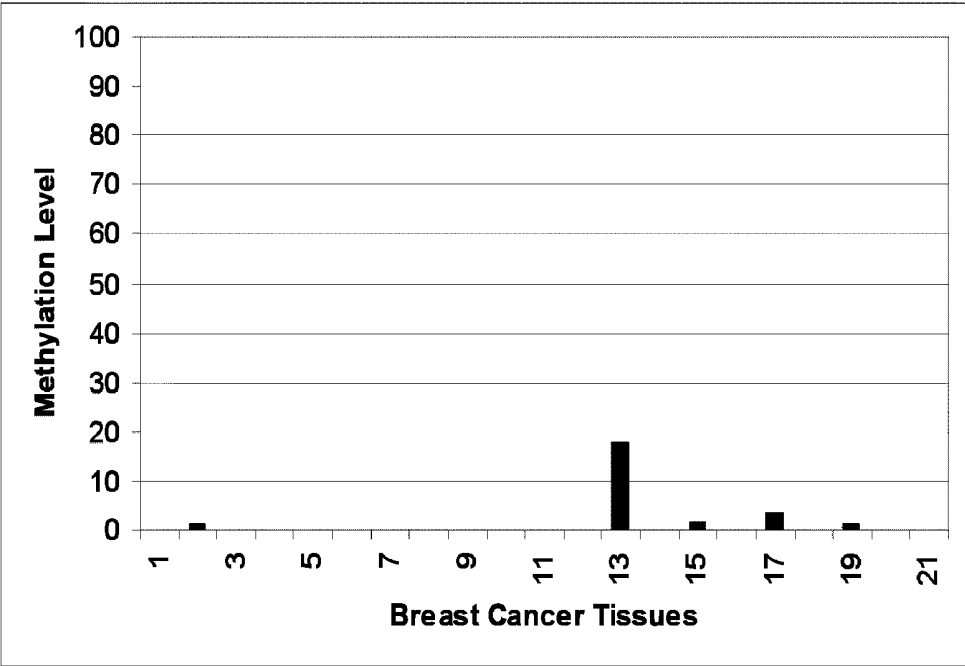


FIGURE 10

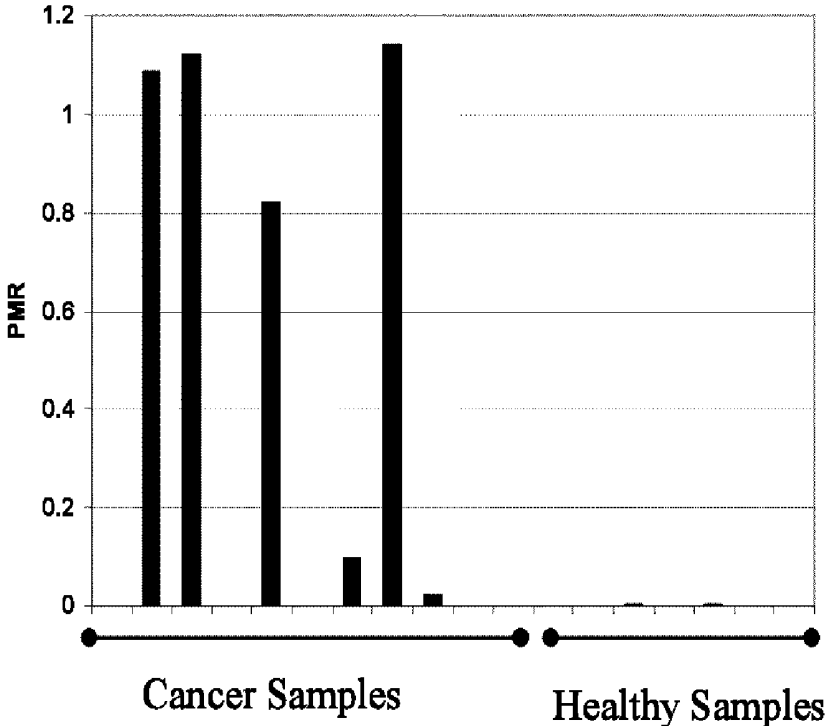


FIGURE 11

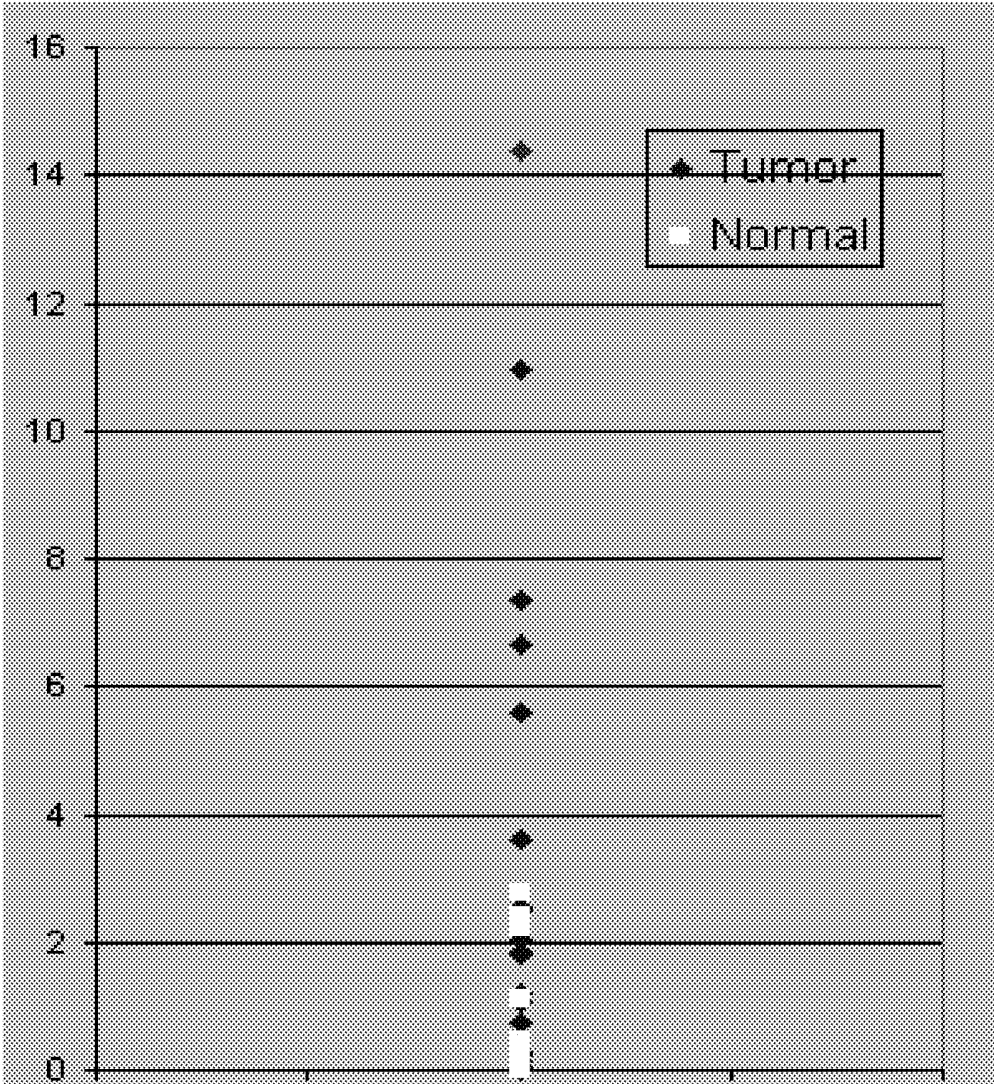


FIGURE 12

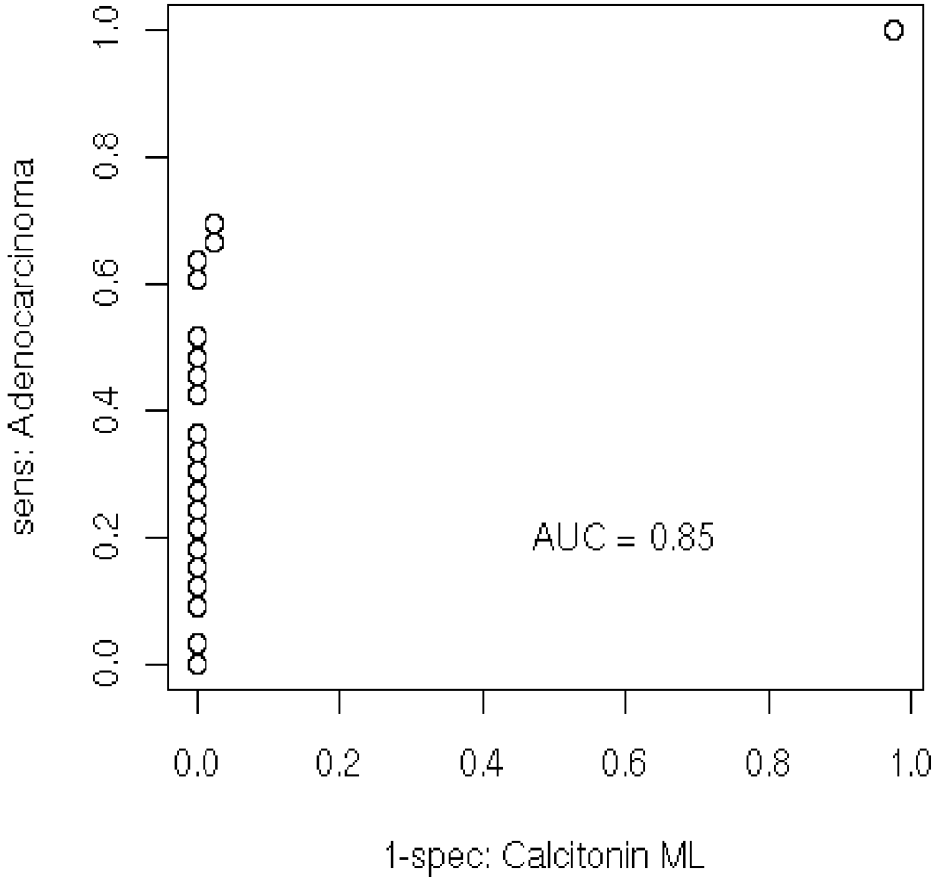


FIGURE 13

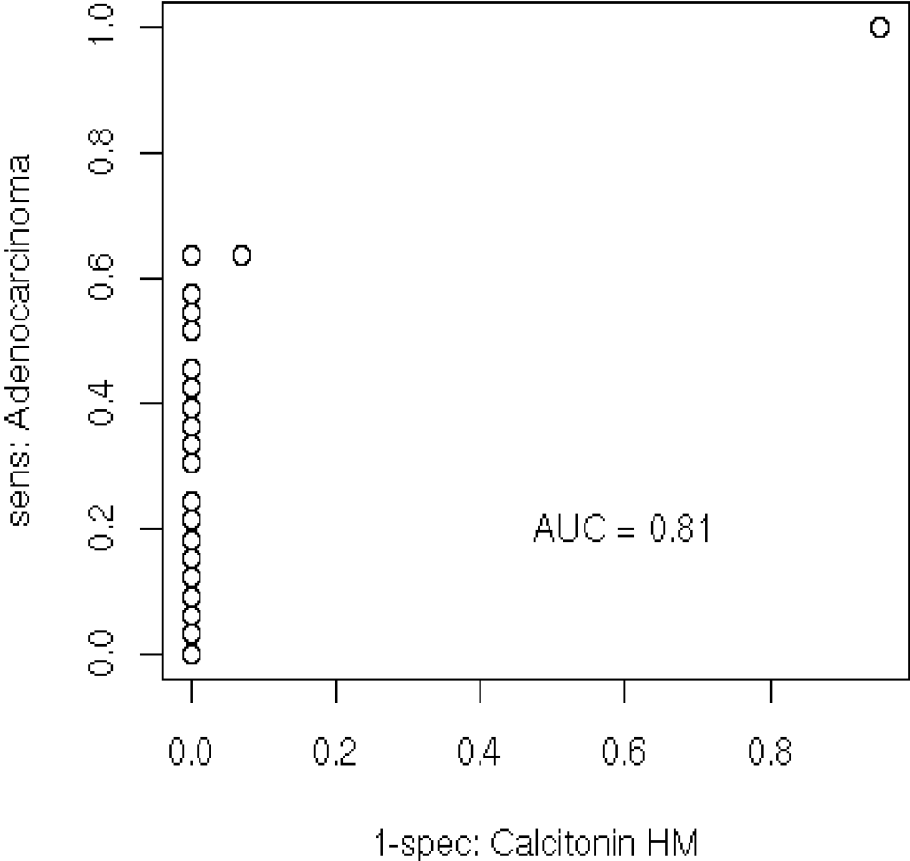


Figure 14

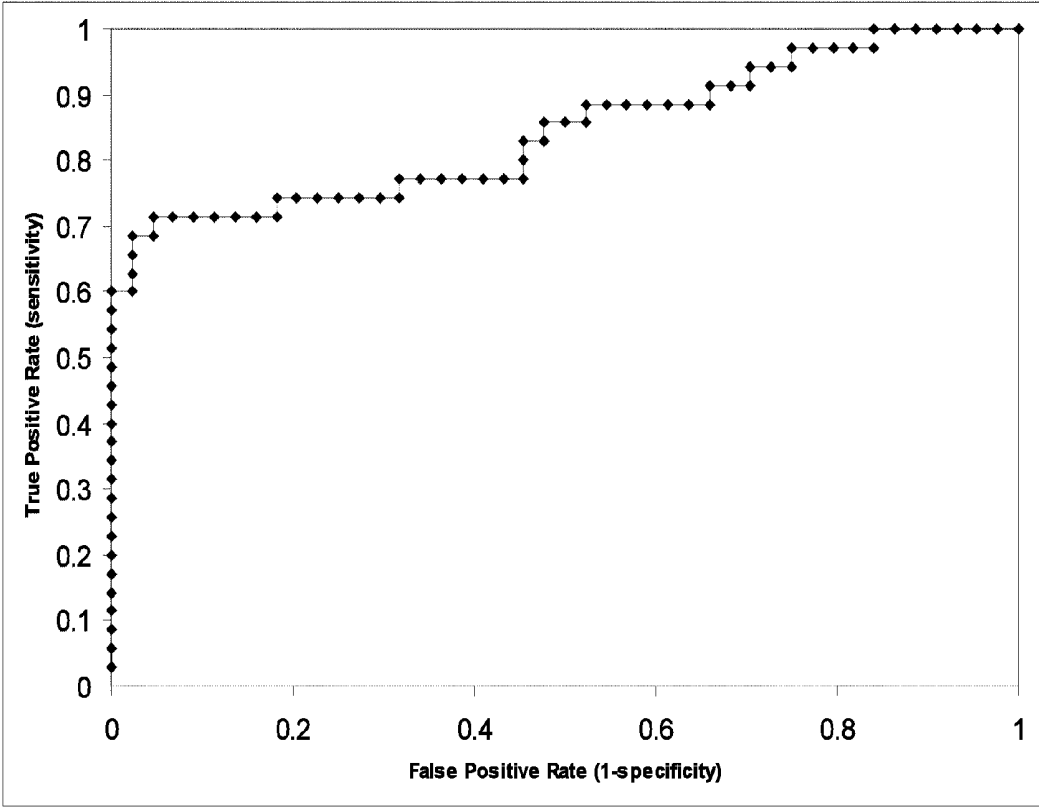


Figure 15

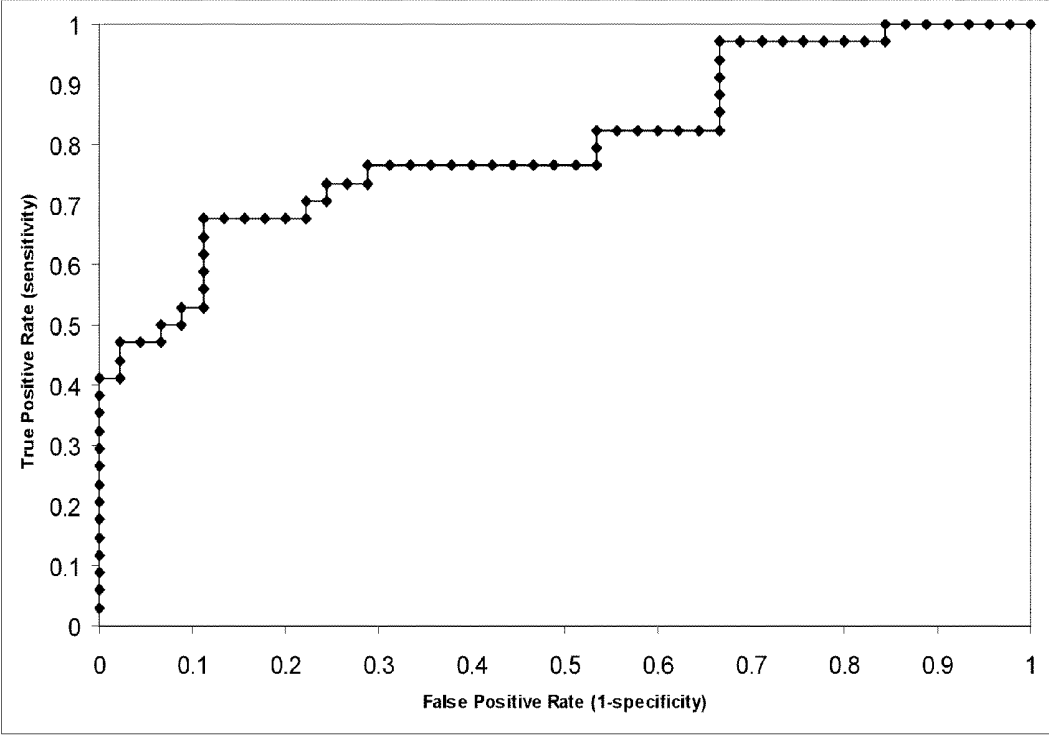


Figure 16

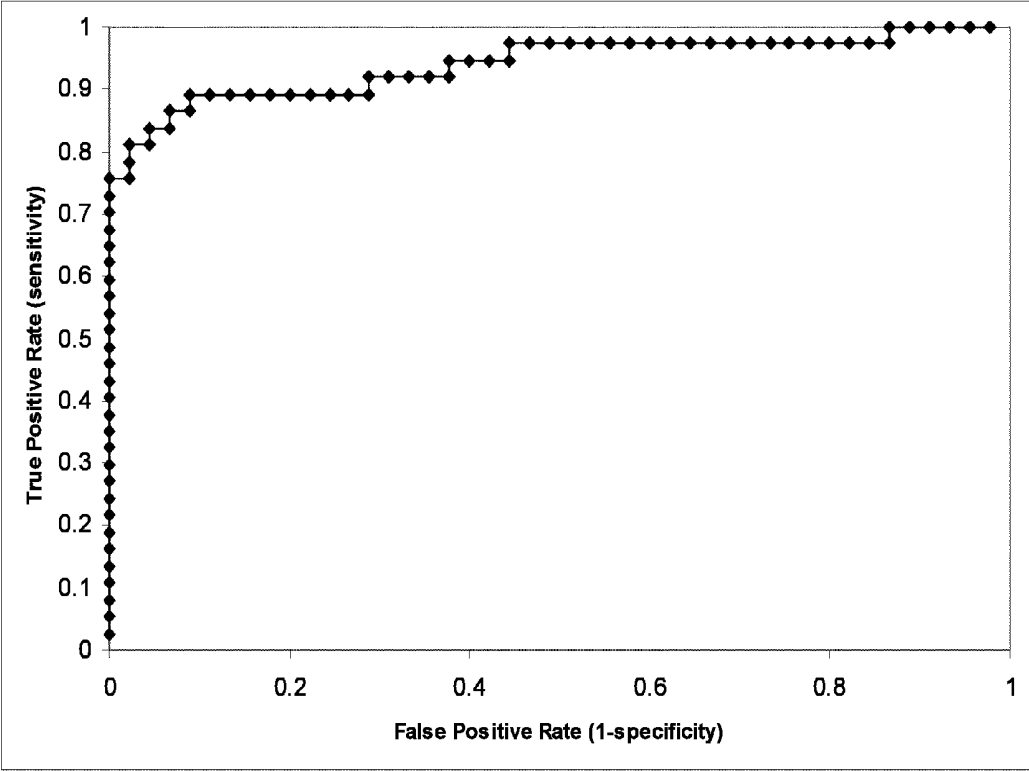




FIGURE 18

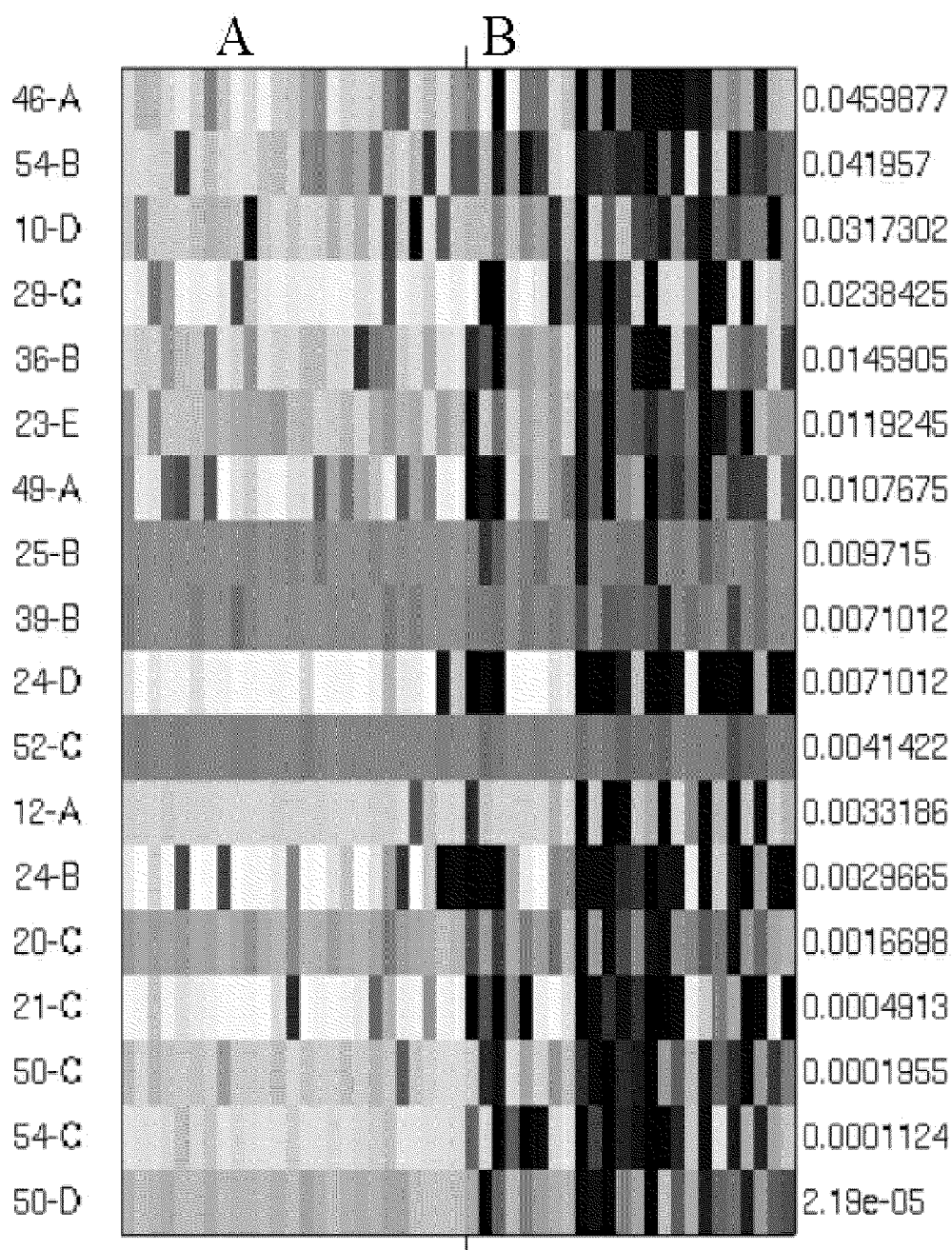


FIGURE 19

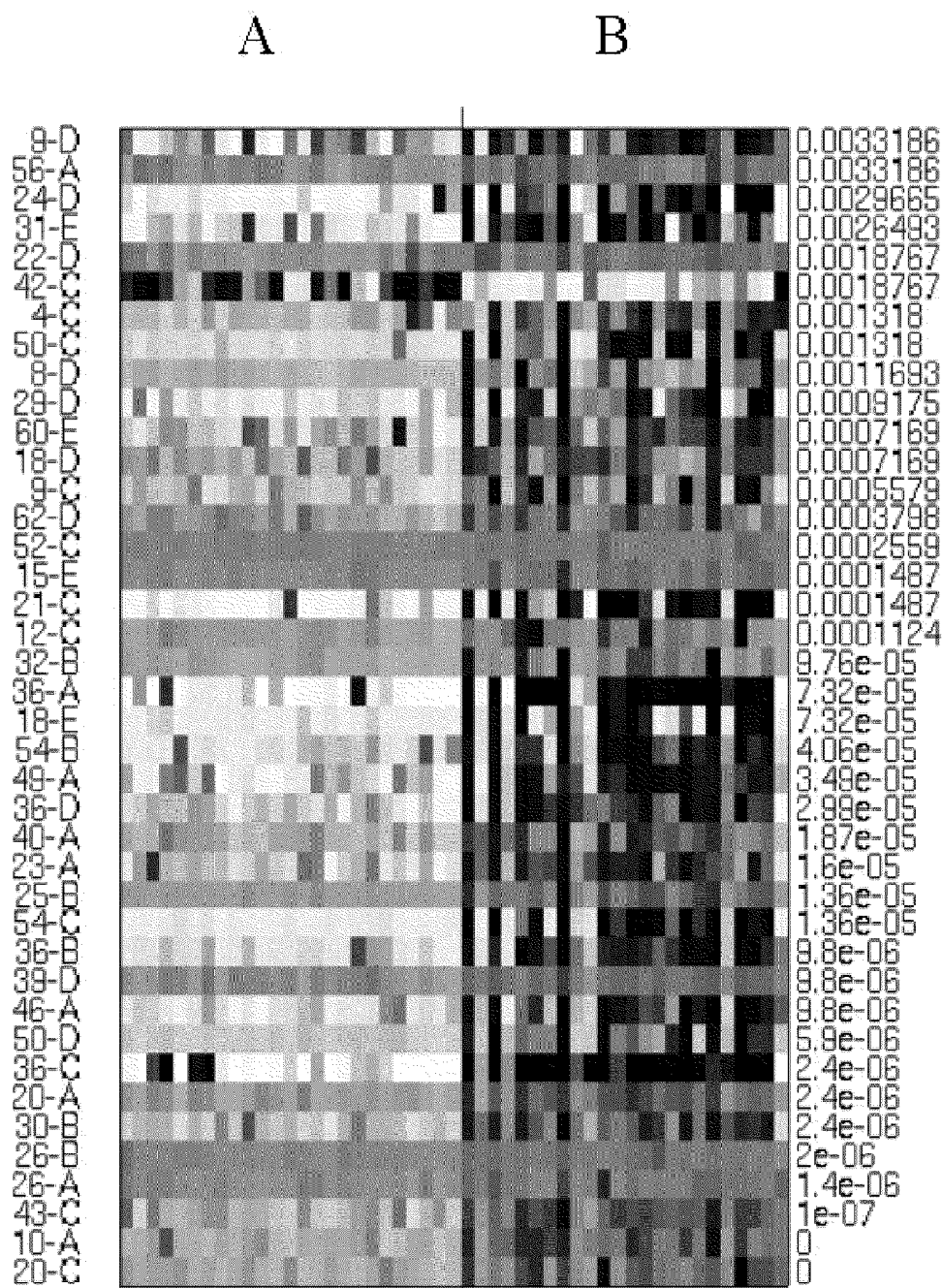


FIGURE 20

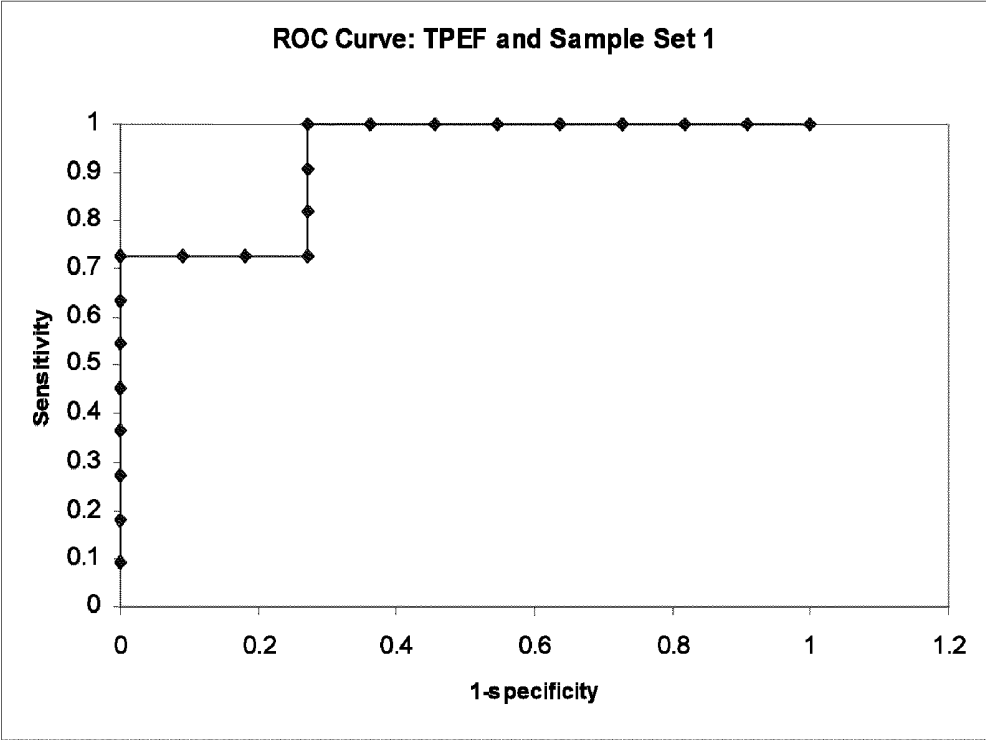


FIGURE 21

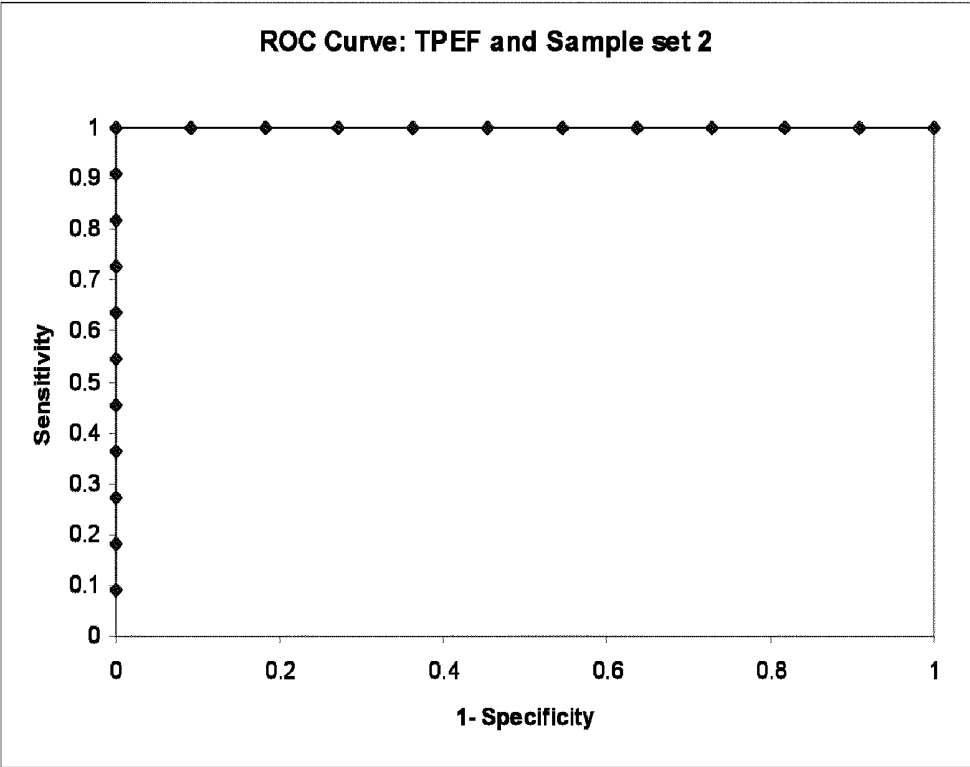


Figure 22

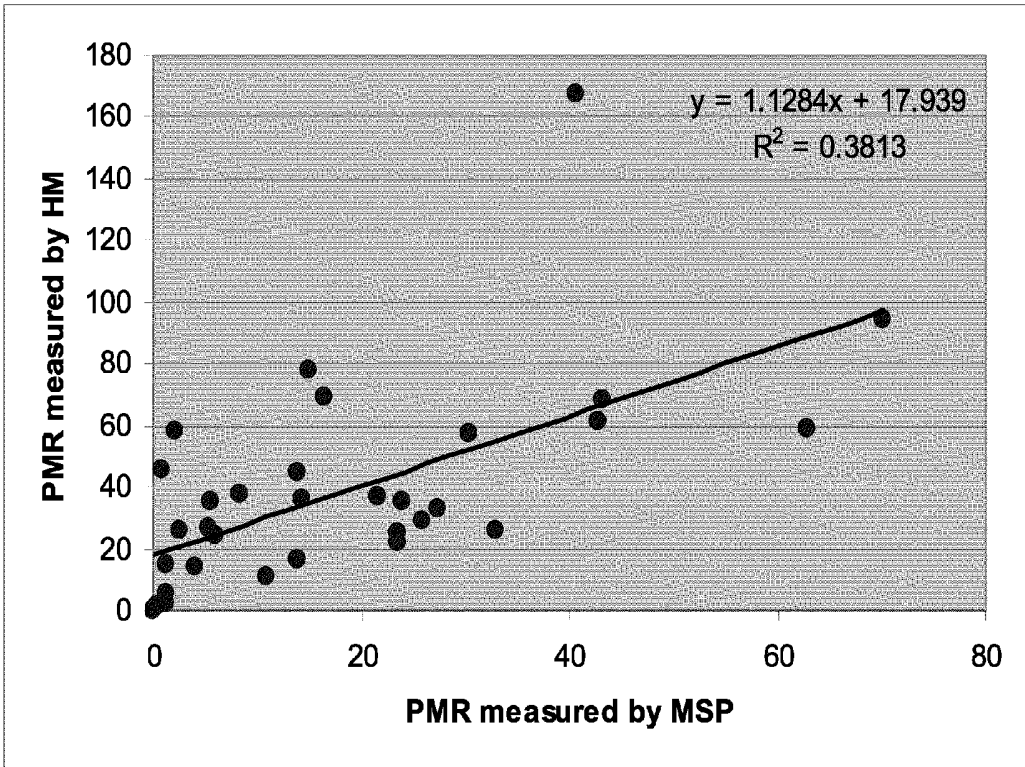
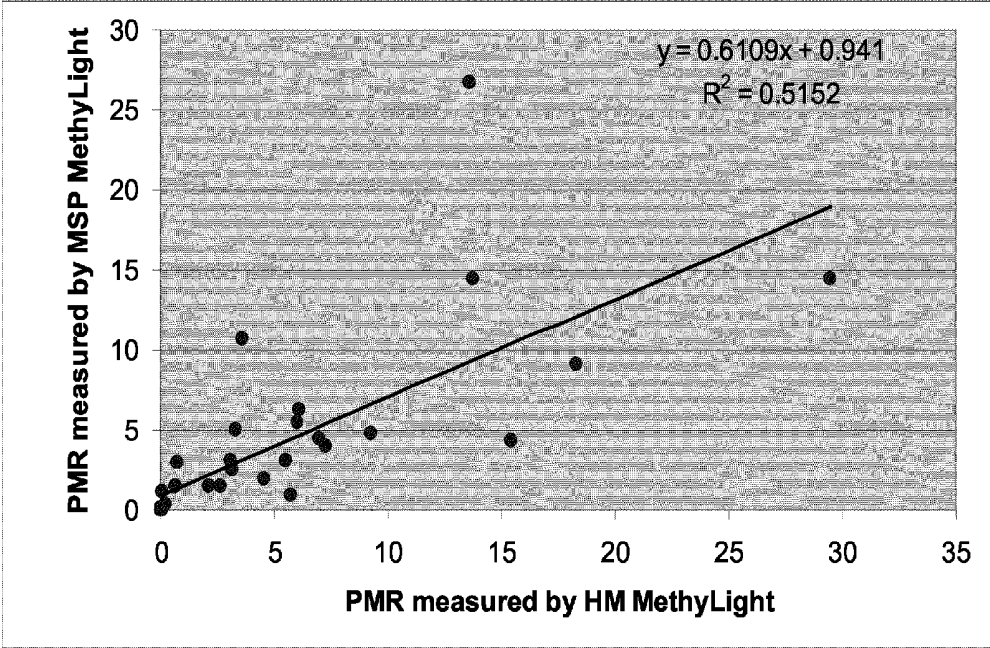


Figure 23



## METHOD AND NUCLEIC ACIDS FOR THE ANALYSIS OF COLON CELL PROLIFERATIVE DISORDERS

### CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of U.S. patent application Ser. No. 10/506,089, which is the United States nationalization pursuant to 35 U.S.C. §371 of International application number PCT/EP2003/002034, filed 27 Feb. 2003 and published 4 Sep. 2003 as WO 2003/072820, which claims priority to European Application EP 02004551.4, filed 27 Feb. 2002, both of which are incorporated by reference herein in their entireties.

### FIELD OF THE INVENTION

[0002] Aspects of the invention relate generally to cellular proliferative disorders (e.g., cancer) and genomic DNA methylation, and more particularly to modified and genomic sequences, to oligonucleotides and/or PNA-oligomers for detecting the cytosine methylation state of genomic DNA, as well as to methods for ascertaining genetic and/or epigenetic parameters of genes for use in the differentiation, diagnosis, prognosis, treatment and/or monitoring of cellular proliferative disorders (e.g., cancer) including colon cell proliferative disorders (e.g., adenocarcinomas, polyps, squamous cell cancers, carcinoid tumours, sarcomas and lymphomas), or the predisposition to colon cell proliferative disorders.

### BACKGROUND

[0003] Colon cancer is the fourth leading cause of cancer mortality in men and women. The 5-year survival rate is 61% over all stages with early detection being a prerequisite for curative therapy of the disease. Up to 95% of all colorectal cancers are adenocarcinomas of varying differentiation grades.

[0004] Sporadic colon cancer develops in a multistep process starting with the pathological transformation of normal colonic epithelium to an adenoma which consecutively progresses to invasive cancer. The progression rate of colonic adenomas is currently predicted based on their histological appearance, location, degree of spread and extent of bowel involvement. For example, tubular-type benign adenomas rarely progress to malignant tumours, whereas villous benign adenomas, particularly if larger than 2 cm in diameter, have a significant malignant potential.

[0005] During progression from benign proliferative lesions to malignant neoplasms several genetic and epigenetic alterations are known to occur. Somatic mutation of the APC gene seems to be one of the earliest events in 75 to 80% of colorectal adenomas and carcinomas. Activation of K-RAS is thought to be a critical step in the progression towards a malignant phenotype. Consecutively, mutations in other oncogenes as well as alterations leading to inactivation of tumor suppressor genes accumulate.

[0006] Aberrant DNA methylation within CpG islands is among the earliest and most common alterations in human cancers leading to abrogation or overexpression of a broad spectrum of genes. In addition, abnormal methylation has been shown to occur in CpG rich regulatory elements in intronic and coding parts of genes for certain tumours. In contrast to the specific hypermethylation of tumour suppressor genes, an overall hypomethylation of DNA can be

observed in tumour cells. This decrease in global methylation can be detected early, far before the development of frank tumour formation. Also, correlation between hypomethylation and increased gene expression was reported for many oncogenes. In colon cancer, aberrant DNA methylation constitutes one of the most prominent alterations and inactivates many tumour suppressor genes such as p14ARF, p16INK4a, THBS1, MINT2, and MINT31 and DNA mismatch repair genes such as hMLH1.

[0007] In the molecular evolution of colorectal cancer, DNA methylation errors have been suggested to play two distinct roles. In normal colonic mucosa cells, methylation errors accumulate as a function of age or as time-dependent events predisposing these cells to neoplastic transformation. For example, hypermethylation of several loci could be shown to be already present in adenomas, particularly in the tubulovillous and villous subtype. At later stages, increased DNA methylation of CpG islands plays an important role in a subset of tumours affected by the so called CpG island methylator phenotype (CIMP). Most CIMP+ tumours, which constitute about 15% of all sporadic colorectal cancers, are characterized by micro satellite instability (MIN) due to hypermethylation of the hMLH1 promoter and other DNA mismatch repair genes. By contrast, CIMP- colon cancers evolve along a more classic genetic instability pathway (CIN), with a high rate of p53 mutations and chromosomal changes.

[0008] However, the molecular subtypes do not only show varying frequencies regarding molecular alterations. According to the presence of either micro satellite instability or chromosomal aberrations, colon cancer can be subclassified into two classes, which also exhibit significant clinical differences. Almost all MIN tumours originate in the proximal colon (ascending and transversum), whereas 70% of CIN tumours are located in the distal colon and rectum. This has been attributed to the varying prevalence of different carcinogens in different sections of the colon. Methylating carcinogens, which constitute the prevailing carcinogen in the proximal colon have been suggested to play a role in the pathogenesis of MIN cancers, whereas CIN tumours are thought to be more frequently caused by adduct-forming carcinogens, which occur more frequently in distal parts of the colon and rectum. Moreover, MIN tumours have a better prognosis than do tumours with a CIN phenotype and respond better to adjuvant chemotherapy.

[0009] The identification of markers for the differentiation of colon carcinoma as well as for early detection are main goals of current research.

[0010] The alpha-calcitonin gene encodes a small family of peptides: calcitonin, katacalcin, and calcitonin gene-related peptide (CGRP). Calcitonin is concerned with skeletal integrity, the secretion of calcitonin is, in part, oestrogen dependent, and it appears likely that a postmenopausal decline in calcitonin secretion is a factor in the development of postmenopausal osteoporosis.

[0011] Investigation of the Calcitonin gene has revealed that hypermethylation of the promoter region of the gene is present in neoplastic cells of several cancer types, particularly acute leukemias. The major part of said research was carried out using methylation sensitive enzyme based methods, this identified the general phenomenon of hypermethylation within the promoter and first exon regions of the gene in multiple types of cancers. However said methods do not allow for a targeted analysis of selected CpG positions. The obser-

vations of hypermethylation were enabled only by the assumption of comethylation within the region. Comethylation is a phenomenon whereby methylation of one CpG position is taken as indicative of methylation of all CpG positions within the region. Examples of research carried out using restriction enzyme based methods include the following:

**[0012]** Hiltunen M O, Koistinaho J, Alhonen L, Myohanen S, Marin S, Kosma V M, Paakkonen M, Janne J. Hypermethylation of the WT1 and calcitonin gene promoter regions at chromosome 11p in human colorectal cancer. *Br J Cancer*. 1997; 76(9):1124-30.

**[0013]** Silverman A L, Park J G, Hamilton S R, Gazdar A F, Luk G D, Baylin S B. Abnormal methylation of the calcitonin gene in human colonic neoplasms. *Cancer Res*. 1989 Jul. 1; 49(13):3468-73.

**[0014]** The gene "Versican," (NM\_004385) also known as CSPG2 encodes a large chondroitin sulfate proteoglycan. This gene is known to exhibit aberrant methylation patterns and conflicting opinions on this matter have been published. For instance Adany R and Iozzo R V. ("Altered methylation of versican proteoglycan gene in human colon carcinoma." *Biochem Biophys Res Commun* 1990 Sep. 28; 171(3):1402-13) observed a correlation between hypomethylation and colonic neoplasms. However, more recently Issa et. al. ("Accelerated Age-related CpG Island Methylation in Ulcerative Colitis," *Cancer Research* 61, 3573-3577, May 1, 2001) described an observed hypermethylation of dysplastic mucosa as compared to non-UC control mucosa (58% versus 31%,  $P=0.01$ ) or compared with adjacent uninvolved mucosa (58% versus 35%,  $P=0.06$ ). Therefore it would seem that although aberrant methylation of this gene has been observed in colorectal cell proliferative disorders, the characterisation of this aberrant methylation is as yet not obvious and it would appear that the commonly held assumption of co-methylation does not hold in the case of this gene.

**[0015]** The gene TPEF (also known as TMEFF2) NM\_016192 encodes a transmembrane protein containing EGF and follistatin domains. It was initially identified on the basis of its methylation properties by Jones et. al. ("The Gene for a Novel Transmembrane Protein Containing Epidermal Growth Factor and Follistatin Domains Is Frequently Hypermethylated in Human Tumor Cells," *Cancer Research* 60, 4907-4912, Sep. 1, 2000). It was therein shown that the 5' region of the gene contained a CpG island, a 3' region of which was shown to exhibit significant hypermethylation in tumor cell lines. Although significant said observation was carried out by means of arbitrarily primed PCR, a methodology that is not suitable for application in a clinical or diagnostic setting.

**[0016]** EYA4 is the most recently identified member of the vertebrate Eya (eyes-absent) gene family, a group of four transcriptional activators that interact with other proteins in a conserved regulatory hierarchy to ensure normal embryologic development. The EYA4 gene is mapped to 6q22.3 and encodes a 640 amino acid protein. The structure of EYA4 conforms to the basic pattern established by EYA1-3, and includes a highly conserved 271 amino acid C-terminus called the eya-homologous region (eyaHR; alternatively referred to as the eya domain or eya homology domain 1) and a more divergent proline-serine-threonine (PST)-rich (34-41%) transactivation domain at the N-terminus. EYA proteins interact with members of the SIX and DACH protein families during early embryonic development. Mutations in the EYA4 gene are responsible for postlingual, progressive, autosomal

dominant hearing loss at the DFNA10 locus (Wayne S, Robertson N G, DeClau F, Chen N, Verhoeven K, Prasad S, Tranebjarg L, Morton C C, Ryan A F, Van Camp G, Smith R J: Mutations in the transcriptional activator EYA4 cause late-onset deafness at the DFNA10 locus. *Hum Mol Genet* 2001 Feb. 1; 10(3):195-200 with further references). A link between the Methylation of Cytosine positions in the EYA 4 gene and cancer has not yet been established.

**[0017]** The cadherins are a family of cell surface glycoproteins responsible for selective cell recognition and adhesion. Several family members, including CDH1 (E-cadherin) and CDH13 (H-cadherin, NM\_001257) are located on the long arm of chromosome 16, while another gene cluster resides on the short arm of chromosome 5. The chromosomal locations of several of the cadherins are sites of frequent loss of heterozygosity in many tumor types. Deletions of 16q are frequent in breast, lung, and other carcinomas. Loss of expression of cadherins has been described in many epithelial cancers, and it may play a role in tumour cell invasion and metastasis. CDH13 expression is diminished in breast and lung cancers. In ovarian tumours, the combination of deletion and aberrant methylation has been reported to inactivate CDH13. Aberrant methylation of CDH13 has also been reported in lung cancers.

**[0018]** 5-methylcytosine is the most frequent covalent base modification in the DNA of eukaryotic cells. It plays a role, for example, in the regulation of the transcription, in genetic imprinting, and in tumorigenesis. Therefore, the identification of 5-methylcytosine as a component of genetic information is of considerable interest. However, 5-methylcytosine positions cannot be identified by sequencing since 5-methylcytosine has the same base pairing behaviour as cytosine. Moreover, the epigenetic information carried by 5-methylcytosine is completely lost during PCR amplification.

**[0019]** A relatively new and currently the most frequently used method for analysing DNA for 5-methylcytosine is based upon the specific reaction of bisulfite with cytosine which, upon subsequent alkaline hydrolysis, is converted to uracil which corresponds to thymidine in its base pairing behaviour. However, 5-methylcytosine remains unmodified under these conditions. Consequently, the original DNA is converted in such a manner that methylcytosine, which originally could not be distinguished from cytosine by its hybridisation behaviour, can now be detected as the only remaining cytosine using "normal" molecular biological techniques, for example, by amplification and hybridisation or sequencing. All of these techniques are based on base pairing which can now be fully exploited. In terms of sensitivity, the prior art is defined by a method which encloses the DNA to be analysed in an agarose matrix, thus preventing the diffusion and renaturation of the DNA (bisulfite only reacts with single-stranded DNA), and which replaces all precipitation and purification steps with fast dialysis (Olek A, Oswald J, Walter J. A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res*. 1996 Dec. 15; 24(24):5064-6). Using this method, it is possible to analyse individual cells, which illustrates the potential of the method. However, currently only individual regions of a length of up to approximately 3000 base pairs are analysed, a global analysis of cells for thousands of possible methylation events is not possible. However, this method cannot reliably analyse very small fragments from small sample quantities either. These are lost through the matrix in spite of the diffusion protection.

**[0020]** An overview of the further known methods of detecting 5-methylcytosine may be gathered from the following review article: Rein, T., DePamphilis, M. L., Zorbas, H., *Nucleic Acids Res.* 1998, 26, 2255.

**[0021]** To date, barring few exceptions (e.g., Zeschnigk M, Lich C, Buiting K, Doerfler W, Horsthemke B. A single-tube PCR test for the diagnosis of Angelman and Prader-Willi syndrome based on allelic methylation differences at the SNRPN locus. *Eur J Hum Genet.* 1997 March-April; 5(2): 94-8) the bisulfite technique is only used in research. Always, however, short, specific fragments of a known gene are amplified subsequent to a bisulfite treatment and either completely sequenced (Olek A, Walter J. The pre-implantation ontogeny of the H19 methylation imprint. *Nat Genet.* 1997 November; 17(3):275-6) or individual cytosine positions are detected by a primer extension reaction (Gonzalzo M L, Jones P A. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* 1997 Jun. 15; 25(12):2529-31, WO 95/00669) or by enzymatic digestion (Xiong Z, Laird P W. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* 1997 Jun. 15; 25(12):2532-4). In addition, detection by hybridization has also been described (Olek et al., WO 99/28498).

**[0022]** Further publications dealing with the use of the bisulfite technique for methylation detection in individual genes are: Grigg G, Clark S. Sequencing 5-methylcytosine residues in genomic DNA. *Bioessays.* 1994 June; 16(6):431-6, 431; Zeschnigk M, Schmitz B, Dittrich B, Buiting K, Horsthemke B, Doerfler W. Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum Mol Genet.* 1997 March; 6(3):387-95; Feil R, Charlton J, Bird A P, Walter J, Reik W. Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. *Nucleic Acids Res.* 1994 Feb. 25; 22(4):695-6; Martin V, Ribieras S, Song-Wang X, Rio M C, Dante R. Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. *Gene.* 1995 May 19; 157(1-2):261-4; WO 97/46705 and WO 95/15373.

**[0023]** An overview of the Prior Art in oligomer array manufacturing can be gathered from a special edition of *Nature Genetics* (*Nature Genetics Supplement*, Volume 21, January 1999), published in January 1999, and from the literature cited therein.

**[0024]** Fluorescently labeled probes are often used for the scanning of immobilized DNA arrays. The simple attachment of Cy3 and Cy5 dyes to the 5'-OH of the specific probe are particularly suitable for fluorescence labels. The detection of the fluorescence of the hybridized probes may be carried out, for example via a confocal microscope. Cy3 and Cy5 dyes, besides many others, are commercially available.

**[0025]** Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-TOF) is a very efficient development for the analysis of biomolecules (Karas M, Hillenkamp F. Laser desorption ionisation of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem.* 1988 Oct. 15; 60(20):2299-301). An analyte is embedded in a light-absorbing matrix. The matrix is evaporated by a short laser pulse thus transporting the analyte molecule into the vapor phase in an unfragmented manner. The analyte is ionised by collisions with matrix molecules. An applied voltage accelerates the

ions into a field-free flight tube. Due to their different masses, the ions are accelerated at different rates. Smaller ions reach the detector sooner than bigger ones.

**[0026]** MALDI-TOF spectrometry is excellently suited to the analysis of peptides and proteins. The analysis of nucleic acids is somewhat more difficult (Gut I G, Beck S. DNA and Matrix Assisted Laser Desorption Ionisation Mass Spectrometry. *Current Innovations and Future Trends.* 1995, 1; 147-57). The sensitivity to nucleic acids is approximately 100 times worse than to peptides and decreases disproportionately with increasing fragment size. For nucleic acids having a multiply negatively charged backbone, the ionisation process via the matrix is considerably less efficient. In MALDI-TOF spectrometry, the selection of the matrix plays an eminently important role. For the desorption of peptides, several very efficient matrixes have been found which produce a very fine crystallisation. There are now several responsive matrixes for DNA, however, the difference in sensitivity has not been reduced. The difference in sensitivity can be reduced by chemically modifying the DNA in such a manner that it becomes more similar to a peptide. Phosphorothioate nucleic acids in which the usual phosphates of the backbone are substituted with thiophosphates can be converted into a charge-neutral DNA using simple alkylation chemistry (Gut I G, Beck S. A procedure for selective DNA alkylation and detection by mass spectrometry. *Nucleic Acids Res.* 1995 Apr. 25; 23(8):1367-73). The coupling of a charge tag to this modified DNA results in an increase in sensitivity to the same level as that found for peptides. A further advantage of charge tagging is the increased stability of the analysis against impurities which make the detection of unmodified substrates considerably more difficult.

**[0027]** Genomic DNA is obtained from DNA of cell, tissue or other test samples using standard methods. This standard methodology is found in references such as Sambrook, Fritsch and Maniatis eds., *Molecular Cloning: A Laboratory Manual*, 1989.

#### DETAILED DESCRIPTION

**[0028]** The invention provides a method for the analysis of biological samples for features associated with the development of colon cell proliferative disorders, characterised in that the nucleic acid of at least one member of the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 is/are contacted with a reagent or series of reagents capable of distinguishing between methylated and non methylated CpG dinucleotides within the genomic sequence of interest.

**[0029]** The genes that form the basis of the present invention may also be used to form a "gene panel," i.e. a collection comprising the particular genetic sequences of the present invention and/or their respective informative methylation sites. The formation of gene panels allows for a quick and specific analysis of specific aspects of breast cancer treatment. The gene panel(s) as described and employed in this invention can be used with surprisingly high efficiency for the improved diagnosis, treatment and monitoring of colon cell proliferative disorders.

**[0030]** The invention provides significant improvements over the state of the art in that there are currently no markers used to detect colorectal cancer from body fluid samples. Current methods used to detect and diagnose colon cell proliferative disorders include colonoscopy, sigmoidoscopy, and fecal occult blood colon cancer. In comparison to these meth-

ods, the disclosed invention is much less invasive than colonoscopy, and as, if not more sensitive than sigmoidoscopy and FOBT. Compared to the previous descriptions of these markers in the literature, the described invention provides significant advantages in terms of sensitivity and specificity due to the advantageous combination of using a gene panel and highly sensitive assay techniques.

**[0031]** The present invention makes available a method for ascertaining genetic and/or epigenetic parameters of genomic DNA. The method is for use in the improved diagnosis, treatment and monitoring of colon cell proliferative disorders, more specifically by enabling the improved identification of and differentiation between subclasses of said disorder and the genetic predisposition to said disorders. The invention presents improvements over the state of the art in that it enables a highly specific classification of colon carcinomas, thereby allowing for improved and informed treatment of patients.

**[0032]** In one aspect of the invention, the disclosed matters provides novel nucleic acid sequences useful for the analysis of methylation within said gene, other aspects provide novel uses of the gene and the gene product as well as methods, assays and kits directed to detecting, differentiating and distinguishing colon cell proliferative disorders. The method and nucleic acids according to the invention may be used for the analysis of colon cell proliferative disorders taken from the group comprising adenocarcinomas, polyps, squamous cell cancers, carcinoid tumours, sarcomas and lymphomas.

**[0033]** In one embodiment the method discloses the use of one or more genes selected from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 as markers for the differentiation, detection and distinguishing of colon cell proliferative disorders. Said use of the gene may be

cific methylation pattern. Due to the degeneracy of the genetic code, the sequence as identified in SEQ ID NOS:1 to SEQ ID NO:5 should be interpreted so as to include all substantially similar and equivalent sequences upstream of the promoter region of a gene which encodes a polypeptide with the biological activity of that encoded by the genes Versican, TPEF, H-Cadherin, Calcitonin and EYA4.

**[0035]** In a preferred embodiment of the method, the objective of the invention is achieved by analysis of a nucleic acid comprising a sequence of at least 18 bases in length according to one of SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto.

**[0036]** The sequences of SEQ ID NOS:6 to SEQ ID NO:25 provide modified versions of the nucleic acid according to SEQ ID NOS:1 to SEQ ID NO:5, wherein the conversion of said sequence results in the synthesis of a nucleic acid having a sequence that is unique and distinct from SEQ ID NOS:1 to SEQ ID NO:5 as follows. (see also the following TABLE 1): SEQ ID NOS:1 to SEQ ID NO:5, sense DNA strand of Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter and regulatory elements; SEQ ID NOS:6 to SEQ ID NO:15, converted SEQ ID NOS:1 to SEQ ID NO:5 and complementary sequences, wherein "C" or "T," but "CpG" remains "CpG" (i.e., corresponds to case where all "C" residues of CpG dinucleotide sequences are methylated and are thus not converted); SEQ ID NOS:16 to SEQ ID NO:25, converted SEQ ID NOS:1 to SEQ ID NO:5 and complementary sequences, wherein "C" converted to "T" for all "C" residues, including those of "CpG" dinucleotide sequences (i.e., corresponds to case where, for SEQ ID NOS:1 to SEQ ID NO:5, all "C" residues of CpG dinucleotide sequences are unmethylated);

TABLE 1

Description of SEQ ID NO: 1 to SEQ ID NO: 25		
SEQ ID NO	Relationship to SEQ ID NO 1 to SEQ ID NO 5	Nature of cytosine base conversion
SEQ ID NOS: 1 to SEQ ID NO: 5	Sense strand (Versican, TPEF, H-Cadherin, Calcitonin and EYA4 including promoter and regulatory elements)	None; untreated sequence
SEQ ID NOS: 6 to 15	Converted methylated strand	"C" to "T," but "CpG" remains "CpG" (all "C" residues of CpGs are methylated)
SEQ ID NOS: 16 to 25	Converted sense strand	"C" to "T" for all "C" residues (all "C" residues of CpGs are unmethylated)

enabled by means of analysis of the methylation status of one or more genes selected from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter or regulatory elements.

**[0034]** The objective of the invention may be achieved by analysis of the methylation state of the CpG dinucleotides within one or more of the genomic sequence according to SEQ ID NOS:1 to SEQ ID NO:5 and sequences complementary thereto. SEQ ID NOS:1 to SEQ ID NO:5 disclose the nucleic acid sequences of the genes from the group consisting of Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter and regulatory elements, wherein said fragment comprises CpG dinucleotides exhibiting a disease spe-

**[0037]** Significantly, heretofore, the nucleic acid sequences and molecules according to SEQ ID NOS:6 to SEQ ID NO:25 were not implicated in or connected with the ascertainment of colon cell proliferative disorders.

**[0038]** The described invention further disclose an oligonucleotide or oligomer for detecting the cytosine methylation state within pretreated DNA, according to SEQ ID NOS:6 to SEQ ID NO:25. Said oligonucleotide or oligomer comprising a nucleic acid sequence having a length of at least nine (9) nucleotides which hybridizes, under moderately stringent or stringent conditions (as defined herein above), to a pretreated nucleic acid sequence according to SEQ ID NOS:6 to SEQ ID NO:25 and/or sequences complementary thereto.

**[0039]** Thus, the present invention includes nucleic acid molecules (e.g., oligonucleotides and peptide nucleic acid (PNA) molecules (PNA-oligomers)) that hybridise under moderately stringent and/or stringent hybridisation conditions to all or a portion of the sequences of SEQ ID NOS:6 to SEQ ID NO:25, or to the complements thereof. The hybridising portion of the hybridising nucleic acids is typically at least 9, 15, 20, 25, 30 or 35 nucleotides in length. However, longer molecules have inventive utility, and are thus within the scope of the present invention.

**[0040]** Preferably, the hybridising portion of the inventive hybridising nucleic acids is at least 95%, or at least 98%, or 100% identical to the sequence, or to a portion thereof of SEQ ID NOS:6 to SEQ ID NO:25, or to the complements thereof.

**[0041]** Hybridising nucleic acids of the type described herein can be used, for example, as a primer (e.g., a PCR primer), or a diagnostic and/or prognostic probe or primer. Preferably, hybridisation of the oligonucleotide probe to a nucleic acid sample is performed under stringent conditions and the probe is 100% identical to the target sequence. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or  $T_m$ , which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions.

**[0042]** For target sequences that are related and substantially identical to the corresponding sequence of SEQ ID NOS:1 to SEQ ID NO:5 (such as allelic variants and SNPs), rather than identical, it is useful to first establish the lowest temperature at which only homologous hybridisation occurs with a particular concentration of salt (e.g., SSC or SSPE). Then, assuming that 1% mismatching results in a 1° C. decrease in the  $T_m$ , the temperature of the final wash in the hybridisation reaction is reduced accordingly (for example, if sequences having >95% identity with the probe are sought, the final wash temperature is decreased by 5° C.). In practice, the change in  $T_m$  can be between 0.5° C. and 1.5° C. per 1% mismatch.

**[0043]** Examples of inventive oligonucleotides of length X (in nucleotides), as indicated by polynucleotide positions with reference to, e.g., SEQ ID NOS:1 to SEQ ID NO:5, include those corresponding to sets of consecutively overlapping oligonucleotides of length X, where the oligonucleotides within each consecutively overlapping set (corresponding to a given X value) are defined as the finite set of Z oligonucleotides from nucleotide positions:

**[0044]**  $n$  to  $(n+(X-1))$ ;

**[0045]** where  $n=1, 2, 3, \dots (Y-(X-1))$ ;

**[0046]** where Y equals the length (nucleotides or base pairs) of SEQ ID NOS:1 to SEQ ID NO:5;

**[0047]** where X equals the common length (in nucleotides) of each oligonucleotide in the set (e.g.,  $X=20$  for a set of consecutively overlapping 20-mers); and

**[0048]** where the number (Z) of consecutively overlapping oligomers of length X for a given SEQ ID NO of length Y is equal to  $Y-(X-1)$ . For example  $Z=2,785-19=2,766$  for either sense or antisense sets of SEQ ID NOS:1 to SEQ ID NO:5, where  $X=20$ .

**[0049]** Preferably, the set is limited to those oligomers that comprise at least one CpG, TpG or CpA dinucleotide.

**[0050]** The present invention encompasses, for each of SEQ ID NOS:6 to SEQ ID NO:25 (sense and antisense), multiple consecutively overlapping sets of oligonucleotides or modified oligonucleotides of length X, where, e.g.,  $X=9, 10, 17, 20, 22, 23, 25, 27, 30$  or 35 nucleotides.

**[0051]** The oligonucleotides or oligomers according to the present invention constitute effective tools useful to ascertain genetic and epigenetic parameters of the genomic sequence corresponding to SEQ ID NOS:1 to SEQ ID NO:5. Preferred sets of such oligonucleotides or modified oligonucleotides of length X are those consecutively overlapping sets of oligomers corresponding to SEQ ID NOS:1 to SEQ ID NO:25 (and to the complements thereof). Preferably, said oligomers comprise at least one CpG, TpG or CpA dinucleotide. Included in these preferred sets are the preferred oligomers corresponding to SEQ ID NOS:11 to SEQ ID NO:15.

**[0052]** Particularly preferred oligonucleotides or oligomers according to the present invention are those in which the cytosine of the CpG dinucleotide (or the thymine of the TpG or the adenosine of the CpA dinucleotide) sequences is within the middle third of the oligonucleotide; that is, where the oligonucleotide is, for example, 13 bases in length, the CpG, TpG or CpA dinucleotide is positioned within the fifth to ninth nucleotide from the 5'-end.

**[0053]** The oligonucleotides of the invention can also be modified by chemically linking the oligonucleotide to one or more moieties or conjugates to enhance the activity, stability or detection of the oligonucleotide. Such moieties or conjugates include chromophores, fluorophors, lipids such as cholesterol, cholic acid, thioether, aliphatic chains, phospholipids, polyamines, polyethylene glycol (PEG), palmitoyl moieties, and others as disclosed in, for example, U.S. Pat. Nos. 5,514,758, 5,565,552, 5,567,810, 5,574,142, 5,585,481, 5,587,371, 5,597,696 and 5,958,773. The probes may also exist in the form of a PNA (peptide nucleic acid) which has particularly preferred pairing properties. Thus, the oligonucleotide may include other appended groups such as peptides, and may include hybridization-triggered cleavage agents (Krol et al., *BioTechniques* 6:958-976, 1988) or intercalating agents (Zon, *Pharm. Res.* 5:539-549, 1988). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a chromophore, fluorophor, peptide, hybridisation-triggered cross-linking agent, transport agent, hybridisation-triggered cleavage agent, etc.

**[0054]** The oligonucleotide may also comprise at least one art-recognised modified sugar and/or base moiety, or may comprise a modified backbone or non-natural internucleoside linkage.

**[0055]** The oligomers according to the present invention are normally used in so called "sets" which in one embodiment contain at least one oligomer for analysis of each of the CpG dinucleotides of a genomic sequence comprising SEQ ID NOS:1 to SEQ ID NO:5 and sequences complementary thereto or to their corresponding CG, TG or CA dinucleotide within the pretreated nucleic acids according to SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto. Preferred is a set which contains at least one oligomer for each of the CpG dinucleotides within one or more genes selected from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter and regulatory elements in both the pretreated and genomic versions of each gene, SEQ ID NOS:1 to SEQ ID NO:25 respectively. However, it is anticipated that for economic or other factors it may be preferable to analyse a limited selection of the CpG dinucleotides within said sequences and the contents of the set of oligonucleotides should be altered accordingly. Therefore, the present invention moreover relates to a set of at least 4 oligonucleotides and/or PNA-oligomers used for detecting the cytosine methylation state in pretreated genomic DNA

(SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto) and genomic DNA (SEQ ID NOS:1 to SEQ ID NO:5 and sequences complementary thereto). These probes enable diagnosis and/or therapy of genetic and epigenetic parameters of cell proliferative disorders. The set of oligomers may also be used for detecting single nucleotide polymorphisms (SNPs) in pretreated genomic DNA (SEQ ID NOS:6 to SEQ ID NO:25, and sequences complementary thereto) and genomic DNA (SEQ ID NOS:1 to SEQ ID NO:5, and sequences complementary thereto).

**[0056]** Moreover, the present invention makes available a set of at least two oligonucleotides which can be used as so-called "primer oligonucleotides" for amplifying DNA sequences of one of SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto, or segments thereof.

**[0057]** In the case of the sets of oligonucleotides according to the present invention, it is preferred that at least one and more preferably all members of the set of oligonucleotides is bound to a solid phase.

**[0058]** According to the present invention, it is preferred that an arrangement of different oligonucleotides and/or PNA-oligomers (a so-called "array") made available by the present invention is present in a manner that it is likewise bound to a solid phase. This array of different oligonucleotide- and/or PNA-oligomer sequences can be characterized in that it is arranged on the solid phase in the form of a rectangular or hexagonal lattice. The solid phase surface is preferably composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold. However, nitrocellulose as well as plastics such as nylon which can exist in the form of pellets or also as resin matrices may also be used.

**[0059]** Therefore, a further subject matter of the present invention is a method for manufacturing an array fixed to a carrier material for analysis in connection with cell proliferative disorders, in which method at least one oligomer according to the present invention is coupled to a solid phase. Methods for manufacturing such arrays are known, for example, from U.S. Pat. No. 5,744,305 by means of solid-phase chemistry and photolabile protecting groups.

**[0060]** A further subject matter of the present invention relates to a DNA chip for the analysis of cell proliferative disorders. DNA chips are known, for example, in U.S. Pat. No. 5,837,832.

**[0061]** The present invention further provides a method for conducting an assay in order to ascertain genetic and/or epigenetic parameters of one or more genes selected from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter and regulatory elements. Most preferably the assays according to the following method are used in order to detect methylation within one or more genes selected from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 wherein said methylated nucleic acids are present in a solution further comprising an excess of background DNA, wherein the background DNA is present in between 100 to 1000 times the concentration of the DNA to be detected. Said method comprises contacting a nucleic acid sample obtained from a subject with at least one reagent or a series of reagents, wherein said reagent or series of reagents, distinguishes between methylated and non-methylated CpG dinucleotides within the target nucleic acid.

**[0062]** Preferably, said method comprises the following steps: In the first step, a sample of the tissue to be analysed is obtained. The source may be any suitable source, preferably,

the source of the sample is selected from the group consisting of histological slides, biopsies, paraffin-embedded tissue, bodily fluids, stool, blood, serum, plasma, urine, sputum and combinations thereof. Preferably, the source is biopsies, bodily fluids, urine, or blood.

**[0063]** The DNA is then isolated from the sample. Extraction may be by means that are standard to one skilled in the art, including the use of detergent lysates, sonification and vortexing with glass beads. Once the nucleic acids have been extracted, the genomic double stranded DNA is used in the analysis.

**[0064]** In the second step of the method, the genomic DNA sample is treated in such a manner that cytosine bases which are unmethylated at the 5'-position are converted to uracil, thymine, or another base which is dissimilar to cytosine in terms of hybridisation behavior. This will be understood as "pretreatment" herein.

**[0065]** The above described treatment of genomic DNA is preferably carried out with bisulfite (hydrogen sulfite, disulfite) and subsequent alkaline hydrolysis which results in a conversion of non-methylated cytosine nucleobases to uracil or to another base which is dissimilar to cytosine in terms of base pairing behaviour. Enclosing the DNA to be analysed in an agarose matrix, thereby preventing the diffusion and renaturation of the DNA (bisulfite only reacts with single-stranded DNA), and replacing all precipitation and purification steps with fast dialysis (Olek A, et al., A modified and improved method for bisulfite based cytosine methylation analysis, *Nucleic Acids Res.* 24:5064-6, 1996). It is further preferred that the bisulfite treatment is carried out in the presence of a radical trap or DNA denaturing agent.

**[0066]** In the third step of the method, fragments of the pretreated DNA are amplified. Wherein the source of the DNA is free DNA from serum, or DNA extracted from paraffin it is particularly preferred that the size of the amplificate fragment is between 100 and 200 base pairs in length, and wherein said DNA source is extracted from cellular sources (e.g. tissues, biopsies, cell lines) it is preferred that the amplificate is between 100 and 350 base pairs in length. It is particularly preferred that said amplicates comprise at least one 20 base pair sequence comprising at least three CpG dinucleotides. Said amplification is carried out using sets of primer oligonucleotides according to the present invention, and a preferably heat-stable polymerase. The amplification of several DNA segments can be carried out simultaneously in one and the same reaction vessel, in one embodiment of the method preferably two or more fragments are amplified simultaneously. Typically, the amplification is carried out using a polymerase chain reaction (PCR). The set of primer oligonucleotides includes at least two oligonucleotides whose sequences are each reverse complementary, identical, or hybridise under stringent or highly stringent conditions to an at least 18-base-pair long segment of the base sequences of SEQ ID NO 6 to SEQ ID NO 25 and sequences complementary thereto.

**[0067]** In an alternate embodiment of the method, the methylation status of preselected CpG positions within the nucleic acid sequences comprising SEQ ID NOS:6 to SEQ ID NO:25 may be detected by use of methylation-specific primer oligonucleotides. This technique (MSP) has been described in U.S. Pat. No. 6,265,171 to Herman. The use of methylation status specific primers for the amplification of bisulfite treated DNA allows the differentiation between methylated and unmethylated nucleic acids. MSP primers pairs contain at

least one primer which hybridizes to a bisulfite treated CpG dinucleotide. Therefore, the sequence of said primers comprises at least one CpG, TpG or CpA dinucleotide. MSP primers specific for non-methylated DNA contain a "T" at the 3' position of the C position in the CpG. Preferably, therefore, the base sequence of said primers is required to comprise a sequence having a length of at least 18 nucleotides which hybridizes to a pretreated nucleic acid sequence according to SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto, wherein the base sequence of said oligomers comprises at least one CpG, TpG or CpA dinucleotide. In this embodiment of the method according to the invention it is particularly preferred that the MSP primers comprise between 2 and 5 CpG, TpG or CpA dinucleotides. It is further preferred that said dinucleotides are located within the 3' half of the primer e.g. wherein a primer is 18 bases in length the specified dinucleotides are located within the first 9 bases from the 3' end of the molecule. In addition to the CpG, TpG or CpA dinucleotides it is further preferred that said primers should further comprise several bisulfite converted bases (i.e. cytosine converted to thymine, or on the hybridizing strand, guanine converted to adenosine). In a further preferred embodiment said primers are designed so as to comprise no more than 2 cytosine or guanine bases.

**[0068]** In one embodiment of the method the primers may be selected from the group consisting of SEQ ID NOS:34 to SEQ ID NO:49, SEQ ID NOS:96, 97, 101, 102, 106 and SEQ ID NO:107.

**[0069]** The fragments obtained by means of the amplification can carry a directly or indirectly detectable label. Preferred are labels in the form of fluorescence labels, radionuclides, or detachable molecule fragments having a typical mass which can be detected in a mass spectrometer. Where said labels are mass labels, it is preferred that the labelled amplicates have a single positive or negative net charge, allowing for better detectability in the mass spectrometer. The detection may be carried out and visualised by means of, e.g., matrix assisted laser desorption/ionisation mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI).

**[0070]** Matrix Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-TOF) is a very efficient development for the analysis of biomolecules (Karas & Hillenkamp, *Anal Chem.*, 60:2299-301, 1988). An analyte is embedded in a light-absorbing matrix. The matrix is evaporated by a short laser pulse thus transporting the analyte molecule into the vapour phase in an unfragmented manner. The analyte is ionised by collisions with matrix molecules. An applied voltage accelerates the ions into a field-free flight tube. Due to their different masses, the ions are accelerated at different rates. Smaller ions reach the detector sooner than bigger ones. MALDI-TOF spectrometry is well suited to the analysis of peptides and proteins. The analysis of nucleic acids is somewhat more difficult (Gut & Beck, *Current Innovations and Future Trends*, 1:147-57, 1995). The sensitivity with respect to nucleic acid analysis is approximately 100-times less than for peptides, and decreases disproportionately with increasing fragment size. Moreover, for nucleic acids having a multiply negatively charged backbone, the ionisation process via the matrix is considerably less efficient. In MALDI-TOF spectrometry, the selection of the matrix plays an eminently important role. For the desorption of peptides, several very efficient matrixes have been found which produce a very fine crystallisation. There are now several responsive matrixes for

DNA, however, the difference in sensitivity between peptides and nucleic acids has not been reduced. This difference in sensitivity can be reduced, however, by chemically modifying the DNA in such a manner that it becomes more similar to a peptide. For example, phosphorothioate nucleic acids, in which the usual phosphates of the backbone are substituted with thiophosphates, can be converted into a charge-neutral DNA using simple alkylation chemistry (Gut & Beck, *Nucleic Acids Res.* 23: 1367-73, 1995). The coupling of a charge tag to this modified DNA results in an increase in MALDI-TOF sensitivity to the same level as that found for peptides. A further advantage of charge tagging is the increased stability of the analysis against impurities, which makes the detection of unmodified substrates considerably more difficult.

**[0071]** In a particularly preferred embodiment of the method the amplification of step three is carried out in the presence of at least one species of blocker oligonucleotides. The use of such blocker oligonucleotides has been described by Yu et al., *BioTechniques* 23:714-720, 1997. The use of blocking oligonucleotides enables the improved specificity of the amplification of a subpopulation of nucleic acids. Blocking probes hybridised to a nucleic acid suppress, or hinder the polymerase mediated amplification of said nucleic acid. In one embodiment of the method blocking oligonucleotides are designed so as to hybridise to background DNA. In a further embodiment of the method said oligonucleotides are designed so as to hinder or suppress the amplification of unmethylated nucleic acids as opposed to methylated nucleic acids or vice versa.

**[0072]** Blocking probe oligonucleotides are hybridised to the bisulfite treated nucleic acid concurrently with the PCR primers. PCR amplification of the nucleic acid is terminated at the 5' position of the blocking probe, such that amplification of a nucleic acid is suppressed where the complementary sequence to the blocking probe is present. The probes may be designed to hybridize to the bisulfite treated nucleic acid in a methylation status specific manner. For example, for detection of methylated nucleic acids within a population of unmethylated nucleic acids, suppression of the amplification of nucleic acids which are unmethylated at the position in question would be carried out by the use of blocking probes comprising a "TpG" at the position in question, as opposed to a "CpG." In one embodiment of the method the sequence of said blocking oligonucleotides should be identical or complementary to molecule is complementary or identical to a sequence at least 18 base pairs in length selected from the group consisting of SEQ ID NOS:6 to SEQ ID NO:25, preferably comprising one or more CpG, TpG or CpA dinucleotides. In one embodiment of the method the sequence of said oligonucleotides is selected from the group consisting SEQ ID NOS:85 to SEQ ID NO:87, SEQ ID NOS:98, 103 and SEQ ID NO:108, and sequences complementary thereto.

**[0073]** For PCR methods using blocker oligonucleotides, efficient disruption of polymerase-mediated amplification requires that blocker oligonucleotides not be elongated by the polymerase. Preferably, this is achieved through the use of blockers that are 3'-deoxyoligonucleotides, or oligonucleotides derivatized at the 3' position with other than a "free" hydroxyl group. For example, 3'-O-acetyl oligonucleotides are representative of a preferred class of blocker molecule.

**[0074]** Additionally, polymerase-mediated decomposition of the blocker oligonucleotides should be precluded. Preferably, such preclusion comprises either use of a polymerase

lacking 5'-3' exonuclease activity, or use of modified blocker oligonucleotides having, for example, thioate bridges at the 5'-terminii thereof that render the blocker molecule nuclease-resistant. Particular applications may not require such 5' modifications of the blocker. For example, if the blocker- and primer-binding sites overlap, thereby precluding binding of the primer (e.g., with excess blocker), degradation of the blocker oligonucleotide will be substantially precluded. This is because the polymerase will not extend the primer toward, and through (in the 5'-3' direction) the blocker—a process that normally results in degradation of the hybridized blocker oligonucleotide.

**[0075]** A particularly preferred blocker/PCR embodiment, for purposes of the present invention and as implemented herein, comprises the use of peptide nucleic acid (PNA) oligomers as blocking oligonucleotides. Such PNA blocker oligomers are ideally suited, because they are neither decomposed nor extended by the polymerase.

**[0076]** In one embodiment of the method, the binding site of the blocking oligonucleotide is identical to, or overlaps with that of the primer and thereby hinders the hybridisation of the primer to its binding site. In a further preferred embodiment of the method, two or more such blocking oligonucleotides are used. In a particularly preferred embodiment, the hybridisation of one of the blocking oligonucleotides hinders the hybridisation of a forward primer, and the hybridisation of another of the probe (blocker) oligonucleotides hinders the hybridisation of a reverse primer that binds to the amplificate product of said forward primer.

**[0077]** In an alternative embodiment of the method, the blocking oligonucleotide hybridises to a location between the reverse and forward primer positions of the treated background DNA, thereby hindering the elongation of the primer oligonucleotides.

**[0078]** It is particularly preferred that the blocking oligonucleotides are present in at least 5 times the concentration of the primers.

**[0079]** In the fourth step of the method, the amplicates obtained during the third step of the method are analysed in order to ascertain the methylation status of the CpG dinucleotides prior to the treatment.

**[0080]** In embodiments where the amplicates were obtained by means of MSP amplification and/or blocking oligonucleotides, the presence or absence of an amplicate is in itself indicative of the methylation state of the CpG positions covered by the primers and/or blocking oligonucleotide, according to the base sequences thereof. All possible known molecular biological methods may be used for this detection, including, but not limited to gel electrophoresis, sequencing, liquid chromatography, hybridisations, real time PCR analysis or combinations thereof. This step of the method further acts as a qualitative control of the preceding steps.

**[0081]** In the fourth step of the method amplicates obtained by means of both standard and methylation specific PCR are further analysed in order to determine the CpG methylation status of the genomic DNA isolated in the first step of the method. This may be carried out by means of base-based methods such as, but not limited to, array technology and probe based technologies as well as by means of techniques such as sequencing and template directed extension.

**[0082]** In one embodiment of the method, the amplicates synthesised in step three are subsequently hybridised to an array or a set of oligonucleotides and/or PNA probes. In this

context, the hybridisation takes place in the following manner: the set of probes used during the hybridisation is preferably composed of at least 2 oligonucleotides or PNA-oligomers; in the process, the amplicates serve as probes which hybridise to oligonucleotides previously bonded to a solid phase; the non-hybridised fragments are subsequently removed; said oligonucleotides contain at least one base sequence having a length of at least 9 nucleotides which is reverse complementary or identical to a segment of the base sequences specified in the SEQ ID NOS:2 to SEQ ID NO:5; and the segment comprises at least one CpG, TpG or CpA dinucleotide.

**[0083]** In a preferred embodiment, said dinucleotide is present in the central third of the oligomer. For example, wherein the oligomer comprises one CpG dinucleotide, said dinucleotide is preferably the fifth to ninth nucleotide from the 5'-end of a 13-mer. One oligonucleotide exists for the analysis of each CpG dinucleotide within the sequence according to SEQ ID NOS:1 to SEQ ID NO:5, and the equivalent positions within SEQ ID NOS:6 to SEQ ID NO:25. Said oligonucleotides may also be in the form of peptide nucleic acids. The non-hybridised amplicates are then removed. The hybridised amplicates are detected. In this context, it is preferred that labels attached to the amplicates are identifiable at each position of the solid phase at which an oligonucleotide sequence is located. In one embodiment of the method said oligonucleotides may be selected from the group comprising SEQ ID NOS:50-77, and SEQ ID NOS:88 and 89.

**[0084]** In yet a further embodiment of the method, the genomic methylation status of the CpG positions may be ascertained by means of oligonucleotide probes that are hybridised to the bisulfite treated DNA concurrently with the PCR amplification primers (wherein said primers may either be methylation specific or standard).

**[0085]** A particularly preferred embodiment of this method is the use of fluorescence-based Real Time Quantitative PCR (Heid et al., *Genome Res.* 6:986-994, 1996; also see U.S. Pat. No. 6,331,393). There are two preferred embodiments of utilising this method. One embodiment, known as the TaqMan™ assay employs a dual-labelled fluorescent oligonucleotide probe. The TaqMan™ PCR reaction employs the use of a nonextendible interrogating oligonucleotide, called a TaqMan™ probe, which is designed to hybridise to a GpC-rich sequence located between the forward and reverse amplification primers. The TaqMan™ probe further comprises a fluorescent “reporter moiety” and a “quencher moiety” covalently bound to linker moieties (e.g., phosphoramidites) attached to the nucleotides of the TaqMan™ oligonucleotide. Hybridised probes are displaced and broken down by the polymerase of the amplification reaction thereby leading to an increase in fluorescence. For analysis of methylation within nucleic acids subsequent to bisulfite treatment, it is required that the probe be methylation specific, as described in U.S. Pat. No. 6,331,393, (hereby incorporated by reference in its entirety) also known as the MethyLight™ assay. The second preferred embodiment of this technology is the use of dual-probe technology (Lightcycler™), each carrying donor or recipient fluorescent moieties, hybridisation of two probes in proximity to each other is indicated by an increase or fluorescent amplification primers. Both these techniques may be adapted in a manner suitable for use with bisulfite treated DNA, and moreover for methylation analysis within CpG dinucleotides. In one embodiment of the method the sequence

of said probe oligonucleotides may be selected from the group comprising SEQ ID NOS:78-84, 90, 99, 100, 104, 105, 109 and SEQ ID NO:110.

**[0086]** In a further preferred embodiment of the method, the fourth step of the method comprises the use of template-directed oligonucleotide extension, such as MS-SNuPE as described by Gonzalgo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997. In said embodiment it is preferred that the MS-SNuPE primer is identical or complementary to a sequence at least nine but preferably no more than twenty five nucleotides in length of one or more of the sequences taken from the group of SEQ ID NOS:2 to SEQ ID NO:5.

**[0087]** In yet a further embodiment of the method, the fourth step of the method comprises sequencing and subsequent sequence analysis of the amplificate generated in the third step of the method (Sanger F., et al., *Proc Natl Acad Sci USA* 74:5463-5467, 1977).

**[0088]** Additional embodiments of the invention provide a method for the analysis of the methylation status of genomic DNA according to the invention (SEQ ID NOS:1 to SEQ ID NO:5) without the need for pretreatment.

**[0089]** In the first step of such additional embodiments, the genomic DNA sample is isolated from tissue or cellular sources. Preferably, such sources include cell lines, histological slides, body fluids, or tissue embedded in paraffin. Extraction may be by means that are standard to one skilled in the art, including but not limited to the use of detergent lysates, sonification and vortexing with glass beads. Once the nucleic acids have been extracted, the genomic double-stranded DNA is used in the analysis.

**[0090]** In a preferred embodiment, the DNA may be cleaved prior to the treatment, and this may be by any means standard in the state of the art, in particular with methylation-sensitive restriction endonucleases.

**[0091]** In the second step, the DNA is then digested with one or more methylation sensitive restriction enzymes. The digestion is carried out such that hydrolysis of the DNA at the restriction site is informative of the methylation status of a specific CpG dinucleotide.

**[0092]** In the third step, which is optional but a preferred embodiment, the restriction fragments are amplified. This is preferably carried out using a polymerase chain reaction, and said amplicates may carry suitable detectable labels as discussed above, namely fluorophore labels, radionuclides and mass labels.

**[0093]** In the final step the amplicates are detected. The detection may be by any means standard in the art, for example, but not limited to, gel electrophoresis analysis, hybridisation analysis, incorporation of detectable tags within the PCR products, DNA array analysis, MALDI or ESI analysis.

**[0094]** The present invention enables diagnosis and/or prognosis of events which are disadvantageous to patients or individuals in which important genetic and/or epigenetic parameters within the Versican, TPEF, H-Cadherin, Calcitonin and EYA4 and their promoter or regulatory elements may be used as markers. Said parameters obtained by means of the present invention may be compared to another set of genetic and/or epigenetic parameters, the differences serving as the basis for a diagnosis and/or prognosis of events which are disadvantageous to patients or individuals.

**[0095]** Specifically, the present invention provides for diagnostic and/or prognostic cancer assays based on measurement of differential methylation of Versican, TPEF, H-Cadherin,

Calcitonin and/or EYA4 CpG dinucleotide sequences. Preferred gene sequences useful to measure such differential methylation are represented herein by SEQ ID NOS:1 to SEQ ID NO:25. Typically, such assays involve obtaining a tissue sample from a test tissue, performing an assay to measure the methylation status of at least one of the inventive Versican, TPEF, H-Cadherin, Calcitonin and/or EYA4 specific CpG dinucleotide sequences derived from the tissue sample, relative to a control sample, and making a diagnosis or prognosis based thereon.

**[0096]** In particular preferred embodiments, inventive oligomers are used to assess Versican, TPEF, H-Cadherin, Calcitonin and/or EYA4 specific CpG dinucleotide methylation status, such as those based on SEQ ID NOS:1 to SEQ ID NO:25, including the representative preferred oligomers corresponding to SEQ ID NOS:ALL OLIGOS, or arrays thereof, as well as a kit based thereon are useful for the diagnosis and/or prognosis of cancer and/or other prostate cell proliferative disorders.

**[0097]** The present invention moreover relates to a diagnostic agent and/or therapeutic agent for the diagnosis and/or therapy colon cell proliferative disorders, the diagnostic agent and/or therapeutic agent being characterised in that at least one primer or probe based on SEQ ID NOS:1 to SEQ ID NO:25 is used for manufacturing it, possibly together with suitable additives and ancillary agents.

**[0098]** Moreover, an additional aspect of the present invention is a kit comprising, for example: a bisulfite-containing reagent as well as at least one oligonucleotide whose sequences in each case correspond, are complementary, or hybridise under stringent or highly stringent conditions to a 18-base long segment of the sequences SEQ ID NOS:1 to SEQ ID NO:5. Said kit may further comprise instructions for carrying out and evaluating the described method. In a further preferred embodiment, said kit may further comprise standard reagents for performing a CpG position-specific methylation analysis, wherein said analysis comprises one or more of the following techniques: MS-SNuPE, MSP, MethyLight, HeavyMethyl™, COBRA, and nucleic acid sequencing. However, a kit along the lines of the present invention can also contain only part of the aforementioned components.

**[0099]** Typical reagents (e.g., as might be found in a typical COBRA-based kit) for COBRA analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); restriction enzyme and appropriate buffer; gene-hybridisation oligo; control hybridisation oligo; kinase labelling kit for oligo probe; and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kits (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

**[0100]** Typical reagents (e.g., as might be found in a typical MethyLight-based kit) for MethyLight analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); Taq-Man® probes; optimized PCR buffers and deoxynucleotides; and Taq polymerase.

**[0101]** Typical reagents (e.g., as might be found in a typical Ms-SNuPE-based kit) for Ms-SNuPE analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimised PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for specific

gene; reaction buffer (for the Ms-SNuPE reaction); and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

[0102] Typical reagents (e.g., as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

#### DEFINITIONS

[0103] The term “CpG island” refers to a contiguous region of genomic DNA that satisfies the criteria of (1) having a frequency of CpG dinucleotides corresponding to an “Observed/Expected Ratio” > 0.6, and (2) having a “GC Content” > 0.5. CpG islands are typically, but not always, between about 0.2 to about 1 kb in length.

[0104] The term “methylation state” or “methylation status” refers to the presence or absence of 5-methylcytosine (“5-mCyt”) at one or a plurality of CpG dinucleotides within a DNA sequence. Methylation states at one or more particular palindromic CpG methylation sites (each having two CpG CpG dinucleotide sequences) within a DNA sequence include “unmethylated,” “fully-methylated” and “hemi-methylated.”

[0105] The term “hemi-methylation” or “hemimethylation” refers to the methylation state of a palindromic CpG methylation site, where only a single cytosine in one of the two CpG dinucleotide sequences of the palindromic CpG methylation site is methylated (e.g., 5'-CC<sup>M</sup>GG-3' (top strand); 3'-GGCC-5' (bottom strand)).

[0106] The term “hypermethylation” refers to the average methylation state corresponding to an increased presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

[0107] The term “hypomethylation” refers to the average methylation state corresponding to a decreased presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

[0108] The term “microarray” refers broadly to both “DNA microarrays,” and “DNA chip(s),” as recognized in the art, encompasses all art-recognized solid supports, and encompasses all methods for affixing nucleic acid molecules thereto or synthesis of nucleic acids thereon.

[0109] “Genetic parameters” are mutations and polymorphisms of genes and sequences further required for their regulation. To be designated as mutations are, in particular, insertions, deletions, point mutations, inversions and polymorphisms and, particularly preferred, SNPs (single nucleotide polymorphisms).

[0110] “Epigenetic parameters” are, in particular, cytosine methylations. Further epigenetic parameters include, for example, the acetylation of histones which, however, cannot be directly analysed using the described method but which, in turn, correlate with the DNA methylation.

[0111] The term “bisulfite reagent” refers to a reagent comprising bisulfite, disulfite, hydrogen sulfite or combinations

thereof, useful as disclosed herein to distinguish between methylated and unmethylated CpG dinucleotide sequences.

[0112] The term “Methylation assay” refers to any assay for determining the methylation state of one or more CpG dinucleotide sequences within a sequence of DNA.

[0113] The term “MS.AP-PCR” (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction) refers to the art-recognized technology that allows for a global scan of the genome using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides, and described by Gonzalogo et al., *Cancer Research* 57:594-599, 1997.

[0114] The term “MethylLight” refers to the art-recognized fluorescence-based real-time PCR technique described by Eads et al., *Cancer Res.* 59:2302-2306, 1999.

[0115] The term “HeavyMethyl” assay, in the embodiment thereof implemented herein, refers to a HeavyMethyl™ MethylLight assay, which is a variation of the MethylLight assay, wherein the MethylLight assay is combined with methylation specific blocking probes covering CpG positions between the amplification primers.

[0116] The term “Ms-SNuPE” (Methylation-sensitive Single Nucleotide Primer Extension) refers to the art-recognized assay described by Gonzalogo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997.

[0117] The term “MSP” (Methylation-specific PCR) refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996, and by U.S. Pat. No. 5,786,146.

[0118] The term “COBRA” (Combined Bisulfite Restriction Analysis) refers to the art-recognized methylation assay described by Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997.

[0119] The term “hybridisation” is to be understood as a bond of an oligonucleotide to a complementary sequence along the lines of the Watson-Crick base pairings in the sample DNA, forming a duplex structure.

[0120] “Stringent hybridisation conditions,” as defined herein, involve hybridising at 68° C. in 5×SSC/5×Denhardt’s solution/1.0% SDS, and washing in 0.2×SSC/0.1% SDS at room temperature, or involve the art-recognized equivalent thereof (e.g., conditions in which a hybridisation is carried out at 60° C. in 2.5×SSC buffer, followed by several washing steps at 37° C. in a low buffer concentration, and remains stable). Moderately stringent conditions, as defined herein, involve including washing in 3×SSC at 42° C., or the art-recognized equivalent thereof. The parameters of salt concentration and temperature can be varied to achieve the optimal level of identity between the probe and the target nucleic acid. Guidance regarding such conditions is available in the art, for example, by Sambrook et al., 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, *Current Protocols in Molecular Biology*, (John Wiley & Sons, N.Y.) at Unit 2.10.

[0121] “Background DNA” as used herein refers to any nucleic acids which originate from sources other than colon cells.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0122] FIG. 1 shows the level of methylation determined by different MSP MethylLight assays and HeavyMethyl MethylLight assays. The Y-axis shows the degree of methylation. Tumor samples are represented by white points, and normal

colon tissue samples by black points. A significantly higher degree of methylation was observed in tumor samples than in healthy tissue samples.

**[0123]** FIG. 2 shows the Receiver Operating Characteristic curve (ROC curve) of the EYA4-MSP-Methyl-Light-Assay for adenocarcinomas according to Example 1. The AUC for the MSP-Methyl-Light-Assay is: 0.94.

**[0124]** FIG. 3 shows the Receiver Operating Characteristic curve (ROC curve) of the EYA4-HM-Methyl-Light-Assay for Adenocarcinoma according to Example 2. The AUC for the HM-Methyl-Light-Assay is: 0.91.

**[0125]** FIG. 4 shows the level of methylation determined by a EYA4-HeavyMethyl MethylLight™ assay according to example 2, testing an additional set of colon samples (25 adenocarcinoma, 33 normals, and 13 adenomas). The Y-axis shows the degree of methylation within the region of the EYA4 gene investigated. Adenocarcinoma samples are represented by white squares, and normal colon tissue samples by black diamonds. A significantly higher degree of methylation was observed in tumor samples than in healthy tissue samples. The level of significance as measured using a t-test was 0.00424.

**[0126]** FIG. 5 shows the Receiver Operating Characteristic curve (ROC curve) of the EYA4-HM-Methyl-Light-Assay for Adenocarcinoma and Adenoma according to Example 2 (additional sets of samples). The area under an ROC curve (AUC) is a measure for the accuracy of a diagnostic test. The AUC for the HM-Methyl-Light-Assay is 0.81.

**[0127]** FIG. 6 shows the Receiver Operating Characteristic curve (ROC curve) of the EYA4-HM-Methyl-Light-Assay for Adenocarcinoma only according to Example 2 (additional sets of samples). The area under an ROC curve (AUC) is a measure for the accuracy of a diagnostic test. The AUC for the HM-Methyl-Light-Assay is: 0.844.

**[0128]** FIG. 7 shows the Receiver Operating Characteristic curve (ROC curve) of the EYA4-HM-Methyl-Light-Assay for Adenomas according to Example 2 (additional sets of samples). The area under an ROC curve (AUC) is a measure for the accuracy of a diagnostic test. The AUC for the HM-Methyl-Light-Assay is: 0.748.

**[0129]** FIG. 8 shows the level of methylation in different tumor and healthy tissues determined by a EYA4-HeavyMethyl MethylLight™ assay according to example 3. The Y-axis shows the degree of methylation within the region of the EYA4 gene investigated. Besides the colon cancer samples only one of the two breast cancer tissues were methylated.

**[0130]** FIG. 9 shows the level of methylation in different breast cancer tissues determined by a EYA4-HeavyMethyl MethylLight™ assay according to example 3. Only one was methylated.

**[0131]** FIG. 10 shows the level of methylation in serum samples determined by a EYA4-HeavyMethyl MethylLight™ assay according to example 4. The Y-axis shows the degree of methylation within the region of the EYA4 gene investigated.

**[0132]** FIG. 11 shows the Receiver Operating Characteristic curve (ROC curve) of the Calcitonin-MSP-Methyl-Light-Assay according to Example 5. The area under an ROC curve (AUC) is a measure for the accuracy of a diagnostic test. The AUC for the HM-Methyl-Light-Assay is: 0.85.

**[0133]** FIG. 12 shows the ROC curve of the Calcitonin-HM-Methyl-Light-Assay according to Example 6. The AUC is: 0.81.

**[0134]** FIG. 13 shows the ROC curve of the Versican-MSP-Methyl-Light-Assay according to Example 9. The AUC is: 0.84.

**[0135]** FIG. 14 shows the ROC curve of the TBEF-MSP-Methyl-Light-Assay according to Example 10. The AUC is: 0.80.

**[0136]** FIG. 15 shows the ROC curve of the Cadherin-MSP-Methyl-Light-Assay according to Example 12. The AUC is: 0.94.

**[0137]** FIG. 16 shows the differentiation of healthy tissue from non healthy tissue wherein the non healthy specimens are obtained from either colon adenoma or colon carcinoma tissue (Example 13). The evaluation is carried out using informative CpG positions from 27 genes. Informative CpG positions from the genes Versican, TPEF, EYA4 and H-Cadherin are further described in Table 3.

**[0138]** FIG. 17 shows the differentiation of healthy tissue from carcinoma tissue using informative CpG positions from 15 genes (Example 13). Informative CpG positions from the genes Versican, TPEF, EYA4 and H-Cadherin are further described in Table 4.

**[0139]** FIG. 18 shows the differentiation of healthy tissue from adenoma tissue using informative CpG positions from 40 genes (Example 13). Informative CpG positions from the genes Versican, TPEF, EYA4 and H-Cadherin are further described in Table 5.

**[0140]** FIG. 19 shows the ROC curve of the TPEF-MSP-Methyl-Light-Assay according to Example 11 (first sample set). The AUC is: 0.93.

**[0141]** FIG. 20 shows the ROC curve of the TPEF-MSP-Methyl-Light-Assay according to Example 11 (second sample set). The AUC is: 1.

**[0142]** FIG. 21 shows the ROC curve of a combined EYA4-Calcitonin-Heavymethyl-MethylLight-Assay according to Example 6. The AUC is: 0.97.

**[0143]** FIG. 22 shows the regression plot of the percentage methylation within the EYA 4 gene calculated in each sample using the MSP and HeavyMethyl variants of the MethylLight assay.

**[0144]** FIG. 23 shows the regression plot of the percentage methylation within the Calcitonin gene calculated in each sample using the MSP and HeavyMethyl variants of the MethylLight assay.

## EXAMPLES

**[0145]** The following examples describe the analysis of the methylation status of the genes EYA 4, Calcitonin, TPEF, H-Cadherin and Versican in healthy and sick colon cell proliferative disorder samples. The initial link between said genes and colon cell proliferative disorders was initially carried by means of hybridisation analysis as described in examples 13 onwards. The genes EYA 4, Calcitonin, TPEF, H-Cadherin and Versican were then selected from the larger set of genes analysed in said examples, and the correlation between methylation status and colon cell proliferative disorder states was validated by analysis of samples using other methylation analysis techniques, namely the MSP-MethylLight and HeavyMethyl MethylLight assays. Please note that the term 'MethylLight' is used to describe real time PCR

analysis of bisulfite treated DNA using probes of both the Taqman single probe) and Lightcycler (dual probe) technologies.

#### Example 1

**[0146]** Analysis of methylation within colon cancer using an MSP-MethylLight assay (EYA4) DNA was extracted from 33 colon adenocarcinoma samples and 43 colon normal adjacent tissues using a Qiagen extraction kit. The DNA from each sample was treated using a bisulfite solution (hydrogen sulfite, disulfite) according to the agarose-bead method (Olek et al 1996). The treatment is such that all non methylated cytosines within the sample are converted to thymidine. Conversely, 5-methylated cytosines within the sample remain unmodified.

**[0147]** The methylation status was determined with a MSP-MethylLight assay designed for the CpG island of interest and a control fragment from the beta actin gene (Eads et al., 2001). The CpG island assay covers CpG sites in both the primers and the Taqman style probe, while the control gene does not. The control gene is used as a measure of total DNA concentration, and the CpG island assay (methylation assay) determines the methylation levels at that site.

**[0148]** Methods: The EYA4 gene CpG island assay was performed using the following primers and probes: Forward Primer: CGGAGGGTACGGAGATTACG (SEQ ID NO:40); Reverse Primer: CGACGACGCGCGAAA (SEQ ID NO:41); and Probe: CGAAACCCTAAATATC-CCGAATAACGCCG (SEQ ID NO:81). The corresponding control assay was performed using the following primers and probes: Primer: TGGTGATGGAGGAGTTAGTAAGT (SEQ ID NO:91); Primer: AACCAATAAACCTACTC-CTCCCTTAA (SEQ ID NO:92); and Probe: ACCACCAC-CCAACACACAATAACAAACACA (SEQ ID NO:93). The reactions were run in triplicate on each DNA sample with the following assay conditions: Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM Magnesium Chloride; 1 unit of taq polymerase; 200 μM dNTPs; of DNA, in a final reaction volume of 20 μl); Cycling conditions: (95° C. for 10 minutes; then 50 cycles of: 95° C. for 15 seconds; 60° C. for 1 minute).

**[0149]** The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001). Results.

**[0150]** Results. The mean PMR for normal samples was 0.15, with a standard deviation of 0.18. The mean PMR for tumour samples was 17.98, with a standard deviation of 18.18. The overall difference in methylation levels between tumour and normal samples is significant in a t-test ( $p=0.00000312$ ). The results are shown in FIG. 1. A Receiver Operating Characteristic curve (ROC curve) of the assay was also determined. A ROC is a plot of the true positive rate against the false positive rate for the different possible cut-points of a diagnostic test. It shows the tradeoff between sensitivity and specificity depending on the selected cutpoint (any increase in sensitivity will be accompanied by a decrease in specificity). The area under an ROC curve (AUC) is a measure for the accuracy of a diagnostic test (the larger the area the better, optimum is 1, a random test would have a ROC curve lying on the diagonal with an area of 0.5; for reference: J. P. Egan. Signal Detection Theory and ROC Analysis, Aca-

demic Press, New York, 1975). The AUC for the MSP-Methyl-Light-Assay is: 0.94 (FIG. 2).

#### Example 2

**[0151]** Methylation within colon cancer was analysed using a EYA4-HeavyMethyl MethylLight assay. The same DNA samples were also used to analyse methylation of the CpG island with a HeavyMethyl MethylLight (or HM MethylLight) assay, also referred to as the HeavyMethyl assay. The methylation status was determined with a HM MethylLight assay designed for the CpG island of interest and the same control gene assay described above. The CpG island assay covers CpG sites in both the blockers and the Taqman style probe, while the control gene does not.

**[0152]** Methods. The CpG island assay (methylation assay) was performed using the following primers and probes:

(SEQ ID NO: 44)

Forward Primer: GGTGATTGTTTATTGTTATGGTTTG;

(SEQ ID NO: 45)

Reverse Primer: CCCCTCAACCTAAAACTACAAC;

(SEQ ID NO: 87)

Forward Blocker: GTTATGGTTTGTGATTTGTGTGGG;

(SEQ ID NO: 86)

Reverse Blocker: AAACCTACAACCACTCAAATCAACCCA;

and

(SEQ ID NO: 84)

Probe: AAAATTACGACGACGCCACCCGAAA.

**[0153]** The reactions were each run in triplicate on each DNA sample with the following assay conditions:

**[0154]** Reaction solution: (400 nM primers; 400 nM probe; 10 μM both blockers; 3.5 mM magnesium chloride; 1×ABI Taqman buffer; 1 unit of ABI TaqGold polymerase; 200 μM dNTPs; and 7 μl of DNA, in a final reaction volume of 20 μl);

**[0155]** Cycling conditions: (95° C. for 10 minutes); (95° C. for 15 seconds, 64° C. for 1 minute (2 cycles)); (95° C. for 15 seconds, 62° C. for 1 minute (2 cycles)); (95° C. for 15 seconds, 60° C. for 1 minute (2 cycles)); and (95° C. for 15 seconds, 58° C. for 1 minute, 60° C. for 40 seconds (41 cycles)).

**[0156]** Results. The mean PMR for normal samples was 1.12 with a standard deviation of 1.45. The mean PMR for tumour samples was 38.23 with a standard deviation of 33.22. The overall difference in methylation levels between tumour and normal samples is significant in a t-test ( $p=0.000000326$ ). The results are shown in FIG. 1.

**[0157]** A ROC curve of the assay was also determined. The AUC for the MSP-Methyl-Light-Assay is 0.91 (FIG. 3).

**[0158]** The assay was tested on an additional set of colon samples (25 adenocarcinoma, 33 normals, and 13 adenomas). The results showed a significant difference again (FIG. 4). The ROC are shown in FIG. 5-7.

**[0159]** The MSP and HeavyMethyl variants of the MethylLight assay were determined to be equivalent for the analysis of methylation in the gene EYA4, FIG. 22 shows the regression plot of the percentage methylation detected in each sample using the two methods.

#### Example 3

**[0160]** The EYA4-HeavyMethyl-MethylLight-assay was also tested against a panel of other tissues (FIG. 8). Besides

the colon cancer samples only one of the two breast cancer tissues were methylated. However, on a panel of 21 additional breast tumours (different stages), only one was methylated (FIG. 9). So the marker is specific for colon tumour samples. All primers, probes, blockers and reaction conditions were identical to those used in the analysis of the colon cancer samples (Example 2).

#### Example 4

**[0161]** Twelve of the colon tissues analysed by real-time PCR also had paired serum taken before surgery. We extracted DNA from 1 ml of that serum using a Qiagen UltraSens DNA extraction kit, bisulfite treated the DNA sample, and ran the EYA4-HeavyMethyl-MethylLight-assay on those samples. The control gene did not amplify for three of the cancer serum samples and three of the normal serum samples, so we can conclude that the sample preparation did not work in these cases. In the other cases, there was evidence of higher methylation in the cancer samples than the normal samples (FIG. 10).

#### Example 5

**[0162]** Analysis of methylation within colon cancer using a Calcitonin-MSP-MethylLight Assay The colon cancer samples described in Example 1 were also analysed using a Calcitonin-MSP-MethylLight Assay, with a Taqman® style probe. The sample preparation was carried out as described above (Example 1) The assay was performed using the following primers and probes:

Primer: AGGTTATCGTCGTCGAGTGT; (SEQ ID NO: 34)

Primer: TCACTCAAACGTATCCAAACCTA; (SEQ ID NO: 35)  
and

Probe: CGAATCTCTCGAACGATCGCATCCA. (SEQ ID NO: 78)

The corresponding control assay was performed as described above (Example 1).

**[0163]** The reactions were run in triplicate on each DNA sample with the following assay conditions:

**[0164]** Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM Magnesium Chloride; 1 unit of taq polymerase; 200 µM dNTPs; 7 µl of DNA, in a final reaction volume of 20 µl);

**[0165]** Cycling conditions: (95° C. for 10 minutes; 95° C. for 15 seconds; 67° C. for 1 minute (3 cycles)); (95° C. for 15 seconds, 64° C. for 1 minute (3 cycles)); (95° C. for 15 seconds, 62° C. for 1 minute (3 cycles)); and (95° C. for 15 seconds, 60° C. for 1 minute (40 cycles)).

**[0166]** The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001).

**[0167]** Results. The mean PMR for normal samples was 0.19, with a standard deviation of 0.79. None of the normal samples was greater than 2 standard deviations about the normal mean, while 18 of 33 tumour samples reached this level of methylation. The overall difference in methylation levels between tumour and normal samples is significant in a t-test ( $p=0.002$ ). The results are shown in FIG. 1. Significantly, the tumour samples are substantially hypermethylated

relative to normal control tissue. A ROC curve of the assay was also determined. The AUC for the MSP-Methyl-Light-Assay is 0.80 (FIG. 11).

#### Example 6

**[0168]** Methylation within colon cancer was analysed using a Calcitonin-HeavyMethyl MethylLight assay. The same DNA samples were also used to analyse methylation of the Calcitonin-CpG island with a HeavyMethyl MethylLight assay using a Taqman® style probe (see Example 2).

**[0169]** The CpG island assay (methylation assay) was performed using the following primers and probes:

(SEQ ID NO: 46)  
Primer: GGATGTGAGAGTTGTTGAGGTTA;

(SEQ ID NO: 47)  
Primer: ACACACCCAAACCATTACTACTCT;

(SEQ ID NO: 83)  
Probe: ACCTCCGAATCTCTCGAACGATCGC;  
and

(SEQ ID NO: 85)  
Blocker: TGTTGAGGTTATGTGTAATTGGGTGTA.

**[0170]** The reactions were each run in triplicate on each DNA sample with the following assay conditions:

**[0171]** Reaction solution: (300 nM primers; 450 nM probe; 3.5 mM magnesium chloride; 2 units of taq polymerase; 400 µM dNTPs, 5 µM blocker; and 7 µl of DNA, in a final reaction volume of 20 µl);

**[0172]** Cycling conditions: (95° C. for 10 minutes); (95° C. for 15 seconds, 67° C. for 1 minute (3 cycles)); (95° C. for 15 seconds, 64° C. for 1 minute (3 cycles)); (95° C. for 15 seconds, 62° C. for 1 minute (3 cycles)); and (95° C. for 15 seconds, 60° C. for 1 minute (40 cycles)).

**[0173]** The corresponding control assay was performed as described above (Example 2).

**[0174]** Results. The mean PMR for normal samples was 0.13 with a standard deviation of 0.58. None of the normal samples was greater than 2 standard deviations about the normal mean, while 19 of 33 tumour samples reached this level of methylation. The overall difference in methylation levels between tumour and normal samples is significant in a t-test ( $p=0.0004$ ). The results are shown in FIG. 1. A ROC curve of the assay was also determined. The AUC for the HM-Methyl-Light-Assay is 0.84 (FIG. 12).

**[0175]** In order to estimate the sensitivity and specificity of a real time assay analysing a gene panel comprising the genes Calcitonin and EYA 4, the ROC of said assay was in silico determined by combining the ROCs of the 2 genes (as described above) using a logistics model. The AUC of said curve (FIG. 21) is 0.97.

**[0176]** The MSP and HeavyMethyl variants of the MethylLight assay were determined to be equivalent for the analysis of methylation in the gene Calcitonin, FIG. 23 shows the regression plot of the percentage methylation detected in each sample using the two methods.

#### Example 7

**[0177]** Serum analysis with Calcitonin-HM MethylLight assay. Twelve of the colon tissues analysed by real-time PCR also had paired serum taken before surgery. DNA was extracted from 1 ml of that serum using a Qiagen UltraSens®

DNA extraction kit, the DNA sample was bisulfite treated, and the HeavyMethyl-MethylLight-assay was run on those samples (see Example 3). Calcitonin was not methylated in all tumours, but of the five patients with the highest levels of methylation in the tumoursect methylation was detected in the serum of four of them. In contrast, no methylation was detected in any of the 11 serum samples taken from healthy donors.

#### Example 8

**[0178]** Identification of the methylation status of a CpG site within the Calcitonin Gene. A fragment of the upstream region of the calcitonin gene (SEQ ID NO:1) was amplified by PCR using the primers CCTTAGTCCCTACCTCTGCT (SEQ ID NO:94) and CTCATTTACACACACCCAAAC (SEQ ID NO:95). The resultant amplicate, 378 by in length, contained an informative CpG at position 165. The amplicate DNA was digested with the methylation sensitive restriction endonuclease Nar I; recognition motif GGCGCC. Hydrolysis by said endonuclease is blocked by methylation of the CpG at position 165 of the amplicate. The digest was used as a control.

**[0179]** Genomic DNA was isolated from the samples using the DNA WIZZARD® DNA isolation kit (PROMEGA™). Each sample was digested using Nar I according to manufacturer's recommendations (New England Biolabs).

**[0180]** About 10 ng of each genomic digest was then amplified using PCR primers CCTTAGTCCCTACCTCTGCT (SEQ ID NO:94) and CTCATTTACACACACCCAAAC (SEQ ID NO:95). The PCR reactions were performed using a thermocycler (Eppendorf GmbH) using 10 ng of DNA, 6 pmole of each primer, 200 µM of each dNTP, 1.5 mM MgCl<sub>2</sub> and 1 U of HOTSTART®Taq (Qiagen AG). The other conditions were as recommended by the Taq polymerase manufacturer.

**[0181]** Using the above mentioned primers, gene fragments were amplified by PCR performing a first denaturation step for 14 min at 96° C., followed by 30-45 cycles (step 2: 60 sec at 96° C., step 3: 45 sec at 52° C., step 4: 75 sec at 72° C.) and a subsequent final elongation of 10 min at 72° C. The presence of PCR products was analysed by agarose gel electrophoresis.

**[0182]** PCR products were detectable, with Nar I-hydrolyzed DNA isolated wherein the tissue in question contained upmethylated DNA, when step 2 to step 4 of the cycle program were repeated 34, 37, 39, 42 and 45 fold. In contrast, PCR products were only detectable with Nar I-hydrolysed

DNA isolated from downmethylated tissue when steps 2 to step 4 of the cycle program were repeated 42- and 45-fold.

#### Example 9

**[0183]** Analysis of methylation within colon cancer using a Versican-MSP-MethylLight Assay. The colon cancer samples described in Example 1 were also analysed using a Versican-MSP-MethylLight Assay with a TAQMAN® style probe. The sample preparation was carried out as described above (Example 1) The assay was performed using the following primers and probes:

(SEQ ID NO: 36)

Forward Primer: TGGGATTAAGATTTTCGGTTAGTTTC;

(SEQ ID NO: 37)

Reverse Primer: CACTACAACGCTACGCGACTAAA;

and

(SEQ ID NO: 79)

Probe: TCGACGTTACCCAAACGAATCACATAAAAAAC

**[0184]** The corresponding control assay was performed as described above (Example 1). The reactions were run in triplicate on each DNA sample with the following assay conditions:

**[0185]** Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM magnesium chloride; 1 units of taq polymerase; 200 µM dNTPs, 5 µM blocker; and 7 µl of DNA, in a final reaction volume of 20 µl);

**[0186]** Cycling conditions: 95° C. for 10 minutes; (95° C. for 15 seconds, 60° C. for 1 minute) 50 cycles.

**[0187]** The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001).

**[0188]** Results. The results are shown in FIG. 1. The mean PMR for normal samples was 3.93, with a standard deviation of 3.57. The mean PMR for tumour samples was 23.06, with a standard deviation of 20.23. The overall difference in methylation levels between tumour and normal samples is significant in a t-test (p=0.00003063). The ROC curve of the assay is shown in FIG. 13. The AUC is 0.84.

**[0189]** This was further confirmed using a Versican-HeavyMethyl MethylLight assay, using dual LIGHTCY-CLER™ probes.

**[0190]** Methods. The CpG island assay (methylation assay) was performed using the following primers and probes:

---

Forward Primer: TGGATAGGAGTTGGGATTAAGATTTT (SEQ ID NO: 96);

Reverse Primer: CTTATTACAATTTAAAAAATAATTCACTACAA (SEQ ID NO: 97);

Blocker: AAATTCACTACAACACTACACAATAATTCACATTAC (SEQ ID NO: 98);

Probe: TTTTCGTATTTTTTTTCGGTTATTACGTTTT-Fluor (SEQ ID NO: 99); and

Probe: LC640-ATGTGATTGGTTTGGGTAACGTCGA-Phos (SEQ ID NO: 100).

---

The reactions were each run in triplicate on each DNA sample with the following assay conditions:

- [0191] Reaction Conditions:
- [0192] 500 nM primers;
- [0193] 10 uM blocker; and
- [0194] 250 nM probes.

LightCycler FastStart Hybridization Probes Mix

- [0195] 4 mM Magnesium Chloride.

Cycling Profile:

- [0196] 95° C. denaturation for 10 minutes; and
- [0197] 50 cycles: 95° C. 10 seconds, 57° C. 30 seconds, 72° C. 20 seconds.

#### Example 10

[0198] Analysis of methylation within colon cancer using a TPEF-MSP-MethylLight Assay. The colon cancer samples described in Example 1 were also analysed using a TPEF-MSP-MethylLight Assay with a TAQMAN® style probe. The sample preparation was carried out as described above (Example 1) The assay was performed using the following primers and probes:

(SEQ ID NO: 38)

Forward Primer: TTTTTTTTCGACGTCGTTG;

(SEQ ID NO: 39)

Reverse Primer: CCTCTACATACGCCGAAT;

and

(SEQ ID NO: 80)

Probe: AATTACCGAAAACATCGACCGA.

The reactions were run in triplicate on each DNA sample with the following assay conditions:

- [0199] Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM magnesium chloride; 1 units of taq polymerase; 200 μM dNTPs, 5 μM blocker; and 7 μl of DNA, in a final reaction volume of 20
- [0200] Cycling conditions: 95° C. for 10 minutes; (95° C. for 15 seconds, 60° C. for 1 minute) 50 cycles.
- [0201] The corresponding control assay was performed as described above (Example 1). The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001).
- [0202] Results. The results are shown in FIG. 1. The mean PMR for normal samples was 3.04, with a standard deviation of 4.21. The mean PMR for tumour samples was 21.38, with a standard deviation of 24.08. The overall difference in methylation levels between tumour and normal samples is significant in a t-test (p=0.0000101973). The ROC curve of the assay is shown in FIG. 14. The AUC is 0.80.
- [0203] This was further confirmed using a TPEF-HeavyMethyl MethylLight assay (using dual labeled LIGHTCYCLER™ probes.
- [0204] Methods. The CpG island assay (methylation assay) was performed using the following primers and probes:

(SEQ ID NO: 101)

Forward Primer: GTAGGTTATTGTTGGTTAATAAAT;

-continued

(SEQ ID NO: 102)

Reverse Primer: TAAAAAAAAAAAAAAAAACTCCTCTACATAC;

(SEQ ID NO: 103)

Blocker: AACTCCTCTACATACACCACAATAAATT;

(SEQ ID NO: 104)

Probe: CGAAAACATCGACCGAACACG-Fluor;

and

(SEQ ID NO: 105)

Probe: LC640-GTCCGAAAAAAAAAAAAAAAACGAACTCC-Phos.

The reactions were each run in triplicate on each DNA sample with the following assay conditions:

Reaction Conditions:

- [0205] Forward primer: 600 nM;
- [0206] Reverse primer: 300 nM;
- [0207] Blocker: 10 uM;
- [0208] Probes: 500 nM;
- [0209] Taq polymerase: 0.1 U/ul;
- [0210] dNTPs: 0.2 mM each;
- [0211] Magnesium Chloride: 4 mM;
- [0212] BSA: 0.25 mg/ml; and
- [0213] Roche buffer with no MgCl: 1x.

Cycling Conditions:

- [0214] 95° C. denaturation for 10 minutes; and
- [0215] 50 cycles: 95° C. for 10 seconds, 57° C. for 25 seconds, 72° C. for 10 seconds.

#### Example 11

[0216] Analysis of methylation within colon cancer using a TPEF-MSP-MethylLight Assay. An additional assay for TPEF was tested on colon samples. The assay was tested on two sets of tissues, each with 12 colon adenocarcinomas and 12 normal adjacent tissue samples.

[0217] The sample preparation was carried out as described above (Example 1). The assay was performed using the following primers and probes:

(SEQ ID NO: 48)

Forward Primer: GGACGTTTTTATCGAAGCGG;

(SEQ ID NO: 49)

Reverse Primer: GCCACCCAACCGCA;

and

(SEQ ID NO: 90)

Probe: ACCCGAAATCAGCGCGAAAAA.

[0218] The reactions were run in triplicate on each DNA sample with the following assay conditions:

- [0219] Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM magnesium chloride; 1 units of taq polymerase; 200 μM dNTPs, 5 μM blocker; and 7 μl of DNA, in a final reaction volume of 20
- [0220] Cycling conditions: 95° C. for 10 minutes; (95° C. for 15 seconds, 60° C. for 1 minute) 50 cycles.
- [0221] The corresponding control assay was performed as described above (Example 1).
- [0222] The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001). In both cases, TPEF was significantly more methylated in the

cancer samples. The ROC curves of the assays are shown in FIG. 19-20. The AUC are 0.93 and 1.

#### Example 12

[0223] Analysis of methylation within colon cancer using a H Cadherin-MSP-MethyLight Assay.

[0224] The colon cancer samples described in Example 1 were also analysed using a H Cadherin-MSP-MethyLight Assay. The sample preparation was carried out as described above (Example 1). The assay was performed using the following primers and probes:

(SEQ ID NO: 42)  
Forward Primer: GACGGATTTTTTTTAAACGTTTTTTC;

(SEQ ID NO: 43)  
Reverse Primer: AAATAAATACCACCTCCGCGA;  
and

(SEQ ID NO: 82)  
Probe: GCTCCTCGCGAAATACTCACCCCG

[0225] The reactions were run in triplicate on each DNA sample with the following assay conditions:

[0226] Reaction solution: (900 nM primers; 300 nM probe; 3.5 mM magnesium chloride; 1 units of taq polymerase; 200  $\mu$ M dNTPs, 5  $\mu$ M blocker; and 7  $\mu$ l of DNA, in a final reaction volume of 20  $\mu$ l).

[0227] Cycling conditions: 95° C. for 10 minutes; (95° C. for 15 seconds, 60° C. for 1 minute) 50 cycles.

[0228] The corresponding control assay was performed as described above (Example 1). The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001).

[0229] Results. The results are shown in FIG. 1. The mean PMR for normal samples was 2.25, with a standard deviation of 2.42. The mean PMR for tumour samples was 25.67, with a standard deviation of 17.57. The overall difference in methylation levels between tumour and normal samples is significant in a t-test ( $p=0.0000000118$ ). The ROC curve of the assay is shown in FIG. 15. The AUC is 0.94.

[0230] This was further confirmed using a H Cadherin-HeavyMethyl MethyLight assay, using dual Lightcycler probes using Lightcycler style dual probe technology.

[0231] Methods. The CpG island assay (methylation assay) was performed using the following primers and probes:

(SEQ ID NO: 106)  
Forward Primer: GTTAGTTAGTTAATTTTTTAAATAGATTAGTAG;

(SEQ ID NO: 107)  
Reverse Primer: CAAAAAACAATAAATAACCACTCC;

(SEQ ID NO: 108)  
Blocker: CCTCCACAAAACCTCACTCCTCACAAAATAC;

(SEQ ID NO: 109)  
Probe: red640 TTTCTTTTGTATGGTAGATACGGGTGA-

phosphate;  
and

(SEQ ID NO: 110)  
Probe: ATTAATGGTTTTATAAGACGGATTTTTTTTAAACGT-

fluoresceine.

[0232] The reactions were each run in triplicate on each DNA sample with the following assay conditions:

Reaction Conditions:

[0233] Forward primer: 600 nM;

[0234] Reverse primer: 300 nM;

[0235] Blocker: 10  $\mu$ M;

[0236] Probes: 500 nM;

[0237] Taq polymerase: 0.1 U/ $\mu$ l;

[0238] dNTPs: 0.2 mM each;

[0239] Magnesium Chloride: 4 mM;

[0240] BSA: 0.25 mg/ml; and

[0241] Roche buffer with no MgCl: 1 $\times$ .

Cycling Conditions:

[0242] 95° C. denaturation for 10 minutes; and  
50 cycles: 95° C. for 10 seconds, 57° C. for 25 seconds, 72° C. for 10 seconds.

#### Example 13

[0243] Multiplex-PCR of colon cancer samples. In the first step the genomic DNA was isolated from the cell samples using the WIZZARD™ kit from (Promega). The isolated genomic DNA from the samples are treated using a bisulfite solution (hydrogen sulfite, disulfite). The treatment is such that all non methylated cytosines within the sample are converted to thiamine, conversely 5-methylated cytosines within the sample remain unmodified. The treated nucleic acids were then amplified using multiplex PCRs, amplifying 8 fragments per reaction with Cy5 fluorescently labelled primers. PCR primers used are described in Table 1. PCR conditions were as follows.

Reaction Solution:

[0244] 10 ng bisulfite treated DNA;

[0245] 3.5 mM MgCl<sub>2</sub>;

[0246] 400  $\mu$ M dNTPs;

[0247] 2 pmol each primer; and

[0248] 1 U Hot Star Taq (Qiagen).

[0249] Forty cycles were carried out as follows. Denaturation at 95° C. for 15 min, followed by annealing at 55° C. for 45 sec., primer elongation at 65° C. for 2 min. A final elongation at 65° C. was carried out for 10 min.

[0250] All PCR products from each individual sample were then hybridised to glass slides carrying a pair of immobilised oligonucleotides for each CpG position under analysis. Each of these detection oligonucleotides was designed to hybridise to the bisulphite converted sequence around one CpG site which was either originally unmethylated (TG) or methylated (CG). See Table 2 for further details of all hybridisation oligonucleotides used (both informative and non-informative). Hybridisation conditions were selected to allow the detection of the single nucleotide differences between the TG and CG variants.

[0251] A 5  $\mu$ l volume of each multiplex PCR product was diluted in 10 $\times$ Ssarc buffer (10 $\times$ Ssarc:230 ml 20 $\times$ SSC, 180 ml sodium lauroyl sarcosinate solution 20%, dilute to 1000 ml with dH<sub>2</sub>O). The reaction mixture was then hybridised to the detection oligonucleotides as follows. Denaturation at 95° C., cooling down to 10° C., hybridisation at 42° C. overnight followed by washing with 10 $\times$ Ssarc and dH<sub>2</sub>O at 42° C.

[0252] Fluorescent signals from each hybridised oligonucleotide were detected using genepix scanner and software.

Ratios for the two signals (from the CG oligonucleotide and the TG oligonucleotide used to analyse each CpG position) were calculated based on comparison of intensity of the fluorescent signals.

[0253] The data was then sorted into a ranked matrix (as shown in FIGS. 16 to 18) according to CpG methylation differences between the two classes of tissues, using an algorithm. The most significant CpG positions are at the bottom of the matrix with significance decreasing towards the top. Black indicates total methylation at a given CpG position, white represents no methylation at the particular position, with degrees of methylation represented in gray, from light (low proportion of methylation) to dark (high proportion of methylation). Each row represents one specific CpG position within a gene and each column shows the methylation profile for the different CpGs for one sample. On the left side a CpG and gene identifier is shown this may be cross referenced with the accompanying tables (Table 1 to 6) in order to ascertain the gene in question and the detection oligomer used. On the right side p values for the individual CpG positions are shown. The p values are the probabilities that the observed distribution occurred by chance in the data set.

[0254] For selected distinctions, we trained a learning algorithm (support vector machine, SVM). The SVM (as discussed by F. Model, P. Adorjan, A. Olek, C. Piepenbrock, Feature selection for DNA methylation based cancer classification. *Bioinformatics*. 2001 June; 17 Suppl 1:S157-64) constructs an optimal discriminant between two classes of given training samples. In this case each sample is described by the methylation patterns (CG/TG ratios) at the investigated CpG sites. The SVM was trained on a subset of samples of each class, which were presented with the diagnosis attached. Independent test samples, which were not shown to the SVM before were then presented to evaluate, if the diagnosis can be predicted correctly based on the predictor created in the training round. This procedure was repeated several times using different partitions of the samples, a method called crossvalidation. Note that all rounds are performed without using any knowledge obtained in the previous runs. The number of correct classifications was averaged over all runs, which gives a good estimate of our test accuracy (percent of correct classified samples over all rounds).

Tables

[0255]

TABLE 1

PCR primers and products		
Gene	Primers	Amplificate length
Versican (SEQ ID NO: 2)	GGATAGGAGTTGGGATTA AGAT AAATCTTTTTCAACACCA AAAT	414
EYA4 (SEQ ID NO: 3)	GGAAGAGGTGATTAAATG GAT CCCAAAAATCAACAACAA	226
H-Cadherin (SEQ ID NO: 4)	TTTGTATTAGGTGGAAGT GGT CCCAAATAAATCAACAAC AACA	286

TABLE 1-continued

PCR primers and products		
Gene	Primers	Amplificate length
TPEF (SEQ ID NO: 5)	TTGTTTGGGTTAATAAATG GA CTTCTCTCTTCTCCCTCTC	295

TABLE 2

Hybridisation oligonucleotides		
Gene	Oligomer sequence	
Versican (SEQ ID NO: 2)	AAGATTTTCGGTTAGTTT (SEQ ID NO: 88)	
Versican (SEQ ID NO: 2)	AAGATTTTCGGTTAGTTT (SEQ ID NO: 89)	
Versican (SEQ ID NO: 2)	ATGTGATTCGTTTGGGTA (SEQ ID NO: 50)	
Versican (SEQ ID NO: 2)	ATGTGATTTGTTTGGGTA (SEQ ID NO: 51)	
Versican (SEQ ID NO: 2)	GGGTAACGTCGAATTTAG (SEQ ID NO: 52)	
Versican (SEQ ID NO: 2)	GGGTAATGTTGAATTTAG (SEQ ID NO: 53)	
Versican (SEQ ID NO: 2)	AAAAATCGCGAGTTTAG (SEQ ID NO: 54)	
Versican (SEQ ID NO: 2)	AAAAATTTGTGAGTTTAG (SEQ ID NO: 55)	
EYA4 (SEQ ID NO: 3)	TATATATACGTGTTGGGTA (SEQ ID NO: 56)	
EYA4 (SEQ ID NO: 3)	TATATATATGTGTTGGGTA (SEQ ID NO: 57)	
EYA4 (SEQ ID NO: 3)	AGTGTATGCGTAGAAGGT (SEQ ID NO: 58)	
EYA4 (SEQ ID NO: 3)	AGTGTATGTGTAGAAGGT (SEQ ID NO: 59)	
EYA4 (SEQ ID NO: 3)	TTTAGATACGAAATGTTA (SEQ ID NO: 60)	
EYA4 (SEQ ID NO: 3)	TTTAGATATGAAATGTTA (SEQ ID NO: 61)	
EYA4 (SEQ ID NO: 3)	AAGTAAGTCGTTGTTGTT (SEQ ID NO: 62)	
EYA4 (SEQ ID NO: 3)	AAGTAAGTTGTTGTTGTT (SEQ ID NO: 63)	
H-Cadherin (SEQ ID NO: 4)	GAAGTGGTCGTTAGTTTT (SEQ ID NO: 64)	
H-Cadherin (SEQ ID NO: 4)	GAAGTGGTTGTTAGTTTT (SEQ ID NO: 65)	
H-Cadherin (SEQ ID NO: 4)	TTGTTTACGCTGATTTGT (SEQ ID NO: 66)	

TABLE 2-continued

Hybridisation oligonucleotides	
Gene	Oligomer sequence
H-Cadherin (SEQ ID NO: 4)	TTGTTTAGTGTGATTTGT (SEQ ID NO: 67)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTCGTTTTGTAA (SEQ ID NO: 68)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTTGTTTTGTAA (SEQ ID NO: 69)
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTCGTGATGTTG (SEQ ID NO: 70)
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTGTGATGTTG (SEQ ID NO: 71)
TPEF (SEQ ID NO: 5)	ATTTGTTTCGATTAATTT (SEQ ID NO: 72)
TPEF (SEQ ID NO: 5)	ATTTGTTTTCGATTAATTT (SEQ ID NO: 73)
TPEF (SEQ ID NO: 5)	ATAGGTTACGGTTGGAG (SEQ ID NO: 74)
TPEF (SEQ ID NO: 5)	ATAGGTTATGGTTGGAG (SEQ ID NO: 75)
TPEF (SEQ ID NO: 5)	AATTTGCGAACGTTTGGG
TPEF (SEQ ID NO: 5)	AATTTGTGAATGTTTGGG

TABLE 3

Oligonucleotides used in differentiation between colon adenomas or carcinoma tissue and healthy colon tissue.	
Gene	Oligo:
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTCGTGATGTTG (SEQ ID NO: 70)
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTGTGATGTTG (SEQ ID NO: 71)
TPEF (SEQ ID NO: 5)	AATTTGCGAACGTTTGGG (SEQ ID NO: 76)
TPEF (SEQ ID NO: 5)	AATTTGTGAATGTTTGGG (SEQ ID NO: 77)
Versican (SEQ ID NO: 2)	GGGTAACGTCGAATTTAG (SEQ ID NO: 52)
Versican (SEQ ID NO: 2)	GGGTAATGTTGAATTTAG (SEQ ID NO: 53)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTCGTTTTGTAA (SEQ ID NO: 68)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTTGTTTTGTAA (SEQ ID NO: 69)
TPEF (SEQ ID NO: 5)	ATAGGTTACGGTTGGAG (SEQ ID NO: 74)
TPEF (SEQ ID NO: 5)	ATAGGTTATGGTTGGAG (SEQ ID NO: 75)

TABLE 3-continued

Oligonucleotides used in differentiation between colon adenomas or carcinoma tissue and healthy colon tissue.	
Gene	Oligo:
EYA4 (SEQ ID NO: 3)	AAGTAAGTCGTTGTTGTT (SEQ ID NO: 62)
EYA4 (SEQ ID NO: 3)	AAGTAAGTTGTTGTTGTT (SEQ ID NO: 63)
EYA4 (SEQ ID NO: 3)	AGTGTATGCGTAGAAGGT (SEQ ID NO: 58)
EYA4 (SEQ ID NO: 3)	AGTGTATGTTGTAAGGT (SEQ ID NO: 59)
Versican (SEQ ID NO: 2)	AAAAATTCGCGAGTTTAG (SEQ ID NO: 54)
Versican (SEQ ID NO: 2)	AAAAATTTGTGAGTTTAG (SEQ ID NO: 55)
Versican (SEQ ID NO: 2)	AAGATTTTCGGTTAGTTT (SEQ ID NO: 88)
Versican (SEQ ID NO: 2)	AAGATTTTGGTTAGTTT (SEQ ID NO: 89)
TPEF (SEQ ID NO: 5)	ATTTGTTTCGATTAATTT (SEQ ID NO: 72)
TPEF (SEQ ID NO: 5)	ATTTGTTTTCGATTAATTT (SEQ ID NO: 73)

TABLE 4

OLIGONUCLEOTIDES USED IN DIFFERENTIATION BETWEEN COLON CARCINOMA TISSUE AND HEALTHY COLON TISSUE.	
Gene	Oligo:
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTCGTGATGTTG (SEQ ID NO: 70)
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTGTGATGTTG (SEQ ID NO: 71)
TPEF (SEQ ID NO: 5)	AATTTGCGAACGTTTGGG (SEQ ID NO: 76)
TPEF (SEQ ID NO: 5)	AATTTGTGAATGTTTGGG (SEQ ID NO: 77)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTCGTTTTGTAA (SEQ ID NO: 68)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTTGTTTTGTAA (SEQ ID NO: 69)
Versican (SEQ ID NO: 2)	GGGTAACGTCGAATTTAG (SEQ ID NO: 52)
Versican (SEQ ID NO: 2)	GGGTAATGTTGAATTTAG (SEQ ID NO: 53)
EYA4 (SEQ ID NO: 3)	AGTGTATGCGTAGAAGGT (SEQ ID NO: 58)

TABLE 4-continued

OLIGONUCLEOTIDES USED IN DIFFERENTIATION BETWEEN COLON CARCINOMA TISSUE AND HEALTHY COLON TISSUE.	
Gene	Oligo:
EYA4 (SEQ ID NO: 3)	AGTGTATGTGTAGAAGGT (SEQ ID NO: 59)
EYA4 (SEQ ID NO: 3)	AAGTAAGTCGTTGTTGTT (SEQ ID NO: 62)
EYA4 (SEQ ID NO: 3)	AAGTAAGTTGTTGTTGTT (SEQ ID NO: 63)
TPEF (SEQ ID NO: 5)	ATAGGTTACGGGTTGGAG (SEQ ID NO: 74)
TPEF (SEQ ID NO: 5)	ATAGGTTATGGGTTGGAG (SEQ ID NO: 75)

TABLE 5

Oligonucleotides used in differentiation between colon adenoma tissue and healthy colon tissue.	
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTCGTGATGTTG (SEQ ID NO: 70)
H-Cadherin (SEQ ID NO: 4)	AATGTTTTTCGTGATGTTG (SEQ ID NO: 71)
TPEF (SEQ ID NO: 5)	AATTCGCGAACGTTTGGG (SEQ ID NO: 76)
TPEF (SEQ ID NO: 5)	AATTTGTGAATGTTTGGG (SEQ ID NO: 77)
TPEF (SEQ ID NO: 5)	ATAGGTTACGGGTTGGAG (SEQ ID NO: 74)
TPEF (SEQ ID NO: 5)	ATAGGTTATGGGTTGGAG (SEQ ID NO: 75)
Versican (SEQ ID NO: 2)	GGGTAACGTCGAATTTAG (SEQ ID NO: 52)
Versican (SEQ ID NO: 2)	GGGTAATGTTGAATTTAG (SEQ ID NO: 53)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTCGTTTTGTAA (SEQ ID NO: 68)
H-Cadherin (SEQ ID NO: 4)	AAGGAATTTGTTTTGTAA (SEQ ID NO: 69)
EYA4 (SEQ ID NO: 3)	AAGTAAGTCGTTGTTGTT (SEQ ID NO: 62)
EYA4 (SEQ ID NO: 3)	AAGTAAGTTGTTGTTGTT (SEQ ID NO: 63)
EYA4 (SEQ ID NO: 3)	AGTGTATGCGTAGAAGGT (SEQ ID NO: 58)
EYA4 (SEQ ID NO: 3)	AGTGTATGTGTAGAAGGT (SEQ ID NO: 59)
Versican (SEQ ID NO: 2)	AAAAATTCGCGAGTTTAG (SEQ ID NO: 54)

TABLE 5-continued

Oligonucleotides used in differentiation between colon adenoma tissue and healthy colon tissue.	
Versican (SEQ ID NO: 2)	AAAAATTTGTGAGTTTAG (SEQ ID NO: 55)

TABLE 6

Genes analysed according to FIGS. 16 to 18	
Number in Figures	Gene name
Healthy vs Non-Healthy	
50-D	H-CADHERIN
20-C	CD44
54-C	TPEF (=TMEFF2; =HPP1)
21-C	VERSICAN
50-C	H-CADHERIN
25-B	GSTP1
43-C	TGFBR2
36-B	N33
49-A	CAV1
52-C	PTGS2
46-A	TP73
54-B	TPEF (=TMEFF2; =HPP1)
20-A	CD44
24-D	EYA4
24-B	EYA4
26-B	GTBP/MSH6
4-C	EGR4
15-E	CDH1
23-E	EGFR
30-B	LKB1
22-D	DAPK1
29-D	IGF2
10-A	HLA-F
29-C	IGF2
36-C	N33
21-D	VERSICAN
39-D	PTEN
32-B	MLH1
26-A	GTBP/MSH6
14-C	CALCA
22-C	DAPK1
39-C	PTEN
9-D	WT1
23-A	EGFR
21-A	VERSICAN
30-A	LKB1
9-C	WT1
60-E	ESR1
12-A	APC
29-A	IGF2
8-D	MYOD1
36-A	N33
54-A	TPEF (=TMEFF2; =HPP1)
18-E	CDKN2a
15-D	CDH1
12-C	APC
Healthy vs Carcinoma	
50-D	H-CADHERIN
54-C	TPEF (=TMEFF2; =HPP1)
50-C	H-CADHERIN
21-C	VERSICAN
20-C	CD44
24-B	EYA4
12-A	APC
52-C	PTGS2
24-D	EYA4
39-B	PGR
25-B	GSTP1

TABLE 6-continued

Genes analysed according to FIGS. 16 to 18	
Number in Figures	Gene name
49-A	CAV1
23-E	EGFR
36-B	N33
29-C	IGF2
10-D	HLA-F
54-B	TPEF (=TMEFF2; =HPP1)
46-A	TP73
Healthy vs Adenoma	
20-C	CD44
10-A	HLA-F
43-C	TGFBR2
26-A	GTBP/MSH6
26-B	GTBP/MSH6
30-B	LKB1
20-A	CD44
36-C	N33
50-D	H-CADHERIN
46-A	TP73
39-D	PTEN
36-B	N33
54-C	TPEF (=TMEFF2; =HPP1)
25-B	GSTP1
23-A	EGFR
40-A	RARB
36-D	N33
49-A	CAV1
54-B	TPEF (=TMEFF2; =HPP1)
18-E	CDKN2a
36-A	N33
32-B	MLH1
12-C	APC
21-C	VERSICAN
15-E	CDH1
52-C	PTGS2
62-D	RASSF1
9-C	WT1

TABLE 6-continued

Genes analysed according to FIGS. 16 to 18	
Number in Figures	Gene name
18-D	CDKN2a
60-E	ESR1
29-D	IGF2
8-D	MYOD1
50-C	H-CADHERIN
4-C	EGR4
42-C	S100A2
22-D	DAPK1
31-E	MGMT
24-D	EYA4
56-A	CEA
9-D	WT1
7-E	GPIb beta
14-C	CALCA
52-D	PTGS2
8-B	MYOD1
24-B	EYA4
21-D	VERSICAN
38-C	PGR
58-A	PCNA
34-D	MSH3
9-B	WT1
35-B	MYC
27-C	HIC-1
52-B	PTGS2
23-E	EGFR
30-A	LKB1
29-C	IGF2
39-C	PTEN
13-D	BCL2
5-B	AR
15-D	CDH1
Carcinoma vs Adenoma	
18-B	CDKN2a
7-E	GPIb beta

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 110

<210> SEQ ID NO 1

<211> LENGTH: 965

<212> TYPE: DNA

<213> ORGANISM: Homo Sapiens

<400> SEQUENCE: 1

```

gatcaattaa gggcatctta gaagttagcg gttcccgctg cctcctttga gcacggaggc      60
caccaacccc tagggggaag agatgtagcg cgaggcaggg gtgtcgtgct aagaaatttc      120
gacgcttctg gggactgagg acaaagggtc ggacacgacc ccggggatcc tggagttccg      180
tgactcgcgc cacggaagcgc acacctaggg gctaatttct gctctgcctc aaagaacctc      240
aagctagagt ccttgccctc gccacagacc ccgggatgcc gctgctgcgc tcaccgcaca      300
ggcagcgcgc ggaaccgctg cagcagatcg cgcgctgcgc gttccaccgg gagatggtgg      360
agacgctgaa aagcttcttt cttgccactc tggacgctgt gggcggcaag cgccttagtc      420
cctacctctg ctgagctgaa cgctcaggca cagtggaact gaaaccgggt tccctacctc      480
tgctgagctg aacgctcagg cacagtggaa ctgaaacctg gttctgcggg atgtgagagc      540

```

-continued

---

tgttgaggtc acgcgtaatt ggggtgatg gagggcgct gttcgtgatg tgtgcaggtt	600
tgatgcaagc aggtcatcgt cgtgcgagtg tgtggatgcg accgcccag agactcggag	660
gcaggcttgg gacacgtttg agtgaacacc tcaggatact cttctggcca gtatctgttt	720
tttagtgctc gtgattcaga gtgggcacat gttgggagac agtaaatgggt ttgggtgtgt	780
gtaaatgagt gtgaccgaa gcgagtgtga gcttgatcta ggcagggacc acacagcact	840
gtcacacctg cctgctcttt agtagaggac tgaagtgcgg ggggtgggggt acggggccgg	900
aatagaatgt ctctgggaca tcttggcaaa cagcagccgg aagcaaaggg gcagctgtgc	960
aaacg	965

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 16579

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo Sapiens

&lt;400&gt; SEQUENCE: 2

tgttttttaa aatgatgtct agctgaacca gccaatggag acaaaaatat ttatatataa	60
tttccctcta agttcactag tcattttata gatcagttat aactttggct actccaatga	120
gttaatttct ttctagagac attataaact ccttaaagat ggaaaataat tgttactttt	180
ttgaataatg cctactatct aaaacaatgc acataaaagc tgcacaataa acttttggtg	240
aattgagaga aaatataaaa aattgaaagc agtctgtttg taataagtgt ttgctaagtc	300
ttacagatct tctctttttt ttggaattat tgcaacaaat atttgacctt tgattttag	360
gattaaaatg tgatttctca tattttaata ggagcaagca ttagtctctgc aattctttt	420
gtcaatagaa taaaaatgga atgtatgctt gctgcttcta catgccggtt gatacccatg	480
ctgaacactt ttcttacaac gcagaggttt ccttgtttct tgtcttgcgt tctaccttg	540
ctcaccacag agacattaat gatgccatgt gtctttttgc accatcagtg cagcagcttt	600
cattatgctc cttgtctcgg tcccactgta gtctaattaa ataaatggaa acgggtagtc	660
actgaaaata aacagcaact gcaaaaagtt tgctcttaa gaaaccttct gcattgtatt	720
ccgccctgcc tcattgattc taagatctct accgcttctt tegtctcaga aaccagactc	780
ttcagtttgt cttgcttcag tcagacaaat gctttataaa aatgacctat attttccaga	840
ctgtggggag agaagaaaaa gactttgaga aagtgactaa ggaggagta agagtgatac	900
aatctttgat tcaatttagt tggactgtgt tctttatgcc ttcactttt tctttataaa	960
aaagatagcg tttttgatgc tctgctccag gcagcctccc aggggtgagtg cagatgtgct	1020
gagccatttg gctgccttga ggaccggctc agagagcctc tcttctcact aatgaagtca	1080
tctctctctg gcagcccggc ctctcccatc actccatggt ggaacttggc tttggcttgc	1140
atagttctgt cgtgatgtag atggctcttc agttaagttc tagagaaatc tgctcaacca	1200
tgaacaggac atttacagca caggagaaaa aaaaatttat taaaaaaat attttactgt	1260
aacattttga tttgggggtt ggagagcaaa aactcaatat ctgatgtctg ggaaaaatct	1320
aatgttttat ggctttctaa ggaaatttgt ctgtgtttaa tttatcagag caatttata	1380
acttcgggtt aaagcacaat gagattgcca ccatcacttt gccttttgc agatttcttt	1440
tttaagaata aaaagaaggc aataaagacc ttatatttct gcctttccta gaaagccaaa	1500
gagaattcaa agtgcctgat gaaatgtctg tccttaattt gcagtgtttt atttaaaaaa	1560

-continued

---

aaaattatga cacaatctta gaggataggc ttcttttttc cagggctctt gaaagttaa	1620
agtacatttt taactctgtag cactactgta tcatccagtg aggataattt gcttagttgt	1680
gcctaggacc agtcagatca aggagttgaa actaagctct ttctgaattt gaatgggact	1740
ggaatcttgt catccctaga gcagcaatag atgatgccat gtctaagaac acttcacctg	1800
atcaataaaa aaggagaagt ctgaaatgta atcctaaga acctcagaag gaggtttaga	1860
tgaagataat aaaattcctt ggagacacat ctgccatcta cacctttcct ctccaagtca	1920
ttggcttgct gccctgagc tcagccaaat tcccaaagac aagatcaaag caaggaagca	1980
aggcacatag cttctggctt atggtttctt ttagacacaa gcattgtgaa tgaagatctt	2040
gctttattta tacctgggat cttggagttt tcacaagaa tatccccagt acaaggatgt	2100
gatttaaagg gggactcaca gaataatatt ttaattactg tttgattcat acttgagtgt	2160
aatggacaaa taaatacaat tttagttata aaacctgaaa aatatatatt ttcccattca	2220
aaaactcttg atagttttgt caggcattgg agaactcagaa ataatgctc tctcattttc	2280
tttctcttga tgaagaaaag tcattaaaaa cagttacata gaaataataa tacaactgta	2340
gactctattc taaatatagc tggatgtttt tcctttttct tttttatcca ttctccattt	2400
tagacagcta gaggggaaaag tctgccctag agatagaaa aagatatcca tggggataaa	2460
tttcatTTta gttatctgat agcctgaaga agtatttaat tttctgaata ataattgtac	2520
aagtcaaaga agagaaaaga caggcagttt aaagcttttc aacattgtgc atattcta	2580
tccttaagtc agcctgaga aagtgctttt gtctctgtaga agtatttggc caagtgttgt	2640
gaagtctatt ggcattaaaa gaaactaaaa attttggagg attagagaat caaatttggc	2700
aacataaatg aaaacatctc agaagttaa aagtatctaa cccatcactt tcaatcatgt	2760
tccattatat acagatggca ttatttatag tgaatgttaa ctattaacat ttccctaaaa	2820
atgggcatta atctgctatt agttgcttgt ttctgatcta gttctccttt gttgaaaggt	2880
tttacttaac tgtatcataa agggacaaaa gggtaaaaga taataagctt ttcagactgg	2940
cagatctgtc acttacaaaa tggccttggg caggttactt gaactcttg aaactcaact	3000
agatttactc aaatttaaga atgagaatac aaatccacat cttgaagtgt ttcacagaaa	3060
ggctctctt aatgtctgga gtatatatt caatgaacat tcattttatt ttatttctct	3120
ccattcttga atcaagcaat cttgaatcta aagttgctat gattagcact gaaaagacca	3180
ctggactatt aattgtgtga ctttgggaca gtaactttct gcacottagt ttgtttacat	3240
gttatacatg aaggttgaag tctgattctg ctctgtgact atcattctaa acatctgatg	3300
aaatcaaatt tcagtgtttg gaatggtagt acaataaatt tactaagaat aaataattca	3360
ctgcaaaaac acattgattt ccaaatgatg taactgacag ttatattact gcagagggct	3420
gataaataac aaaagaaatg aaagatgcac atgggtgagaa ctgaaattat cctgacaagt	3480
cttctacttg tttatcactt aaaatcaatg accatgctga atgcctacaa attacaaaat	3540
ataaaagaaa tcttataaat gcgcatgtac aggagtctaa gttactaaaa gttttaagc	3600
ataagtttaa accaaactaa tcaaagaagt tgagaggaaa aattggcttt catctttaat	3660
cactactgtt ttgaggtcct atgtttaata taattttcta agtagaggct tcagagagaa	3720
gagttgtgag gatactttca tatttgtgta gaagaaaag tttgccatcc attctagtat	3780
ccctagtgtt atactgatgt gcaccttggg tttattttgt tcctattgta taaactcata	3840

---

-continued

---

cttgacttca aagaaaagga aaatocaaag tcctctttt ctaaggggac agaaatcctt 3900  
tgtgtcaact gtttgacctt tttctctgta aggtcctatt ggaaatcttt tgtaacacaa 3960  
tgcaggggac tcttccatgt gttgatgctg tttacacagt ggggtgggac tgactgaaga 4020  
aaaaaaatcg catatacgca tgaagatta tggctcttatt tccggaaagc atgaaaggtg 4080  
attgatactt ccaagaagtc cctgttactc aggaaaatta tcaaatattc tactcagaga 4140  
tacttgaaa gactgaagga aaggaagaac gaagaaagca gaatctagac ttatgtgggg 4200  
agagatttgt ggcagaggaa aagtattctc tttgaatccg acaagggtt tgctggggg 4260  
aatttctgt ccagcctttt attaccaggg tcttttgaag cgggctccc cattgggcag 4320  
ttcctggga gtgcagtggg gaattcttac actttccctc taggtccccg aaggatctcg 4380  
ttttctcagt gtctctttca ggttggcagg agccttgagc ctgacacttc cctttgatgg 4440  
gacaggcaag ctctgtgggc gcgtaaacac gctgtaacca agttctttgc tgattttaca 4500  
gttttgtgtg ctcccagaaa gaagtgatcg tactcaattg tctattgctg gctgcccc 4560  
taagagcctg ggggctcctt tcccctaacc cagaactagc tgcacggggg gcggggaagt 4620  
gggggtgggg aaggagtggg agggcagtgg tttccgagc cagagcgtg ttactgagtg 4680  
agtcctgaa tggggagcgc tgctgtcccc aagccgattg gtacttcttg tcaggaagaa 4740  
acgccaagag gtgggagtc ctggggaggg aggcaggcgg tccctaccgc aggcgcgggg 4800  
agetgccttt ccgcccctc cctgctttc caagcctgga ctcttaggag tggctgaagc 4860  
tgcggagcgc ttttgagcc tgtgaatgaa ccctcctcct ctcccctc cttcttctcg 4920  
ctgagtctcc tctcggctc tgacggtaca gtgatataat gatgatgggt gtcacaacc 4980  
gcatttgaac ttgcaggcga gctgccccga gcctttctgg ggaagaactc caggcgtgcg 5040  
gacgcaacag ccgagaacat tagtggtgt ggacaggagc tgggaccaag atcttcggcc 5100  
agcccgcct cctccgcct cttccagcac cgtccgcac cctccgcac cttcccggg 5160  
ccaccacgct tctatgtga cccgctggg caacgccgaa cccagtcgag cagcgtgca 5220  
gtgaatttc ccccaaaact gcaataagcc gccttccaag gtaatcacgt ttctttgtt 5280  
cccccttaa aaaacaaaa caaaaaactt atagaaaaa acccgcgagc ttagaaaaa 5340  
gaagcaattg gtagaaggct ttaattaagg caaagagctg taaggcgaag ttaagaaaat 5400  
gtaggcactt aaaaaatgca ggtaactttc ataaaggctt ttggggagag gcatacagag 5460  
ggacctggt gttgaaaaag attcagacaa aagaaacca ggggtgggtg ggggtaaaat 5520  
gactaacgga attgggggaa gggagggaat aaattgtaa gaaatcatag aaaagtgtg 5580  
ggttcttgag ctggagagaa gagagggacc tttggcactt tgattttttt gttgtgtgtg 5640  
ttcttaacac gctcagggca aaagtttgaa tggggactac caagactgc cacagacaag 5700  
tccccgaagc cgccttggtg caggccacct ggtttccag cccctggtgt gtggtcagtg 5760  
cctggtgttc ctggaagcc actcccggc agctcctgac agtgcgacc gccgccaag 5820  
cagcctggga ccttgccgag acctgacccc ttcagaccgc aggcagtctg ggaggaggtc 5880  
cggccggggg aggtgcagga tcccgcctgt gtctctttga cgacttgggg actgtcacgg 5940  
ttctctccc ggcgccctgt gttctttgt cctgcacgag gtgcgaaggg gccagcaggg 6000  
aaggagcaga ggatggggg tgggggtgtt ggagccccgc ggaggtctgg gaggccctg 6060  
ggcgggaaaa gctgttctg aatcggcagg gatgtgcaac aattttctc acctgaaga 6120

---

-continued

---

gtgaaatagg gtctgtcgct cccatctcaa caagcaaacc ggcacccaga ggcactgca 6180  
gacaaagggt gctcggggac ccgaatcagg ggctctggg tcagtcctct cgcccagact 6240  
acaggagtcc tccgtttcct tacatattcc ccaccctctc ctagtctctc gctctagttg 6300  
agcaacttta ctggcaggtc tagcggcagc ggcgcgcgcg tgtgccggga gccccggggg 6360  
acgtcctcgg ctggagcggc ccaccgtcct gcagaggcgt gggcgctgta gggcgacatc 6420  
cctggtgctg gcagacctag ggcaccggg ttgttctggc ccgcggctctg tgtcactggg 6480  
cgtggagcgg tctgggtgtt aggcagagga gagcggggta gaaaaatagt gtcatagcct 6540  
aaactatttg ctctccaatc caactgcgtc tcggcaagtc ctgcatgtcc ctgggagatg 6600  
ctgcgggaag ggggagaaat ctacacggcg cctggagagt tgtcctgccg cgcgcacacc 6660  
cgcggcagag tctctctggg tcgcgttctc catctttgtg tctccctctg tctccatctt 6720  
ctctgectcc ctccccctct cccagcctct gttctctctc tctagcttct gegtccccctc 6780  
ccccacagct gacaaatgaa tggctagtgt gaaatccctg cctttccctg gctccaaggc 6840  
agcaggaggg gagagcggag gggaggcggg cgctccttcg gatctgcccc cagttcagct 6900  
cacaagactt gcagaaccta gatgtctagg aattgggagt tttgcggcgg gtgtggggcg 6960  
cccctgatgg agaagtcccc cacagggcga gaaaaacaag cccccagac caagcgtgca 7020  
tctttcacia ccttgtgttc aaggatggaa ggccttagtt ttcccttaag tcattcattt 7080  
tgcactccac aatctcggc gtacatctag gagtttgag gacctgaaa aaaggtcttg 7140  
gtctgtgcaa agtgacagaa tgccttctcg cgagtccgct ggaacgcacg cggcgccctg 7200  
ggtccagttc gcccttagt ggacgaacgc acatggccgg ggtggagcca ggtttcggag 7260  
gtctagacgc gccaccctgg gcgtctttaa gaaagataat acacataccg tgctctcaa 7320  
accatgacc acttgaacct aacggtaggg cccccccca ggatattgca aaagaagggc 7380  
tcctaacca aaacttaaac ttctcaact tcaggcggg cgctcggaagg gaaagtggag 7440  
agaaggcggg cagtgggagg aaaaaagaaa gggaaaggaag ggaagtggaa ggaagagga 7500  
ccgtggaggg gaggaagagg agaaaggaac ggagatgggg cagatggcat tcaaacagcc 7560  
aggaacggac cggaggggtt cgtttcgggg tttgggtctg aatttcgaga ctcgtttag 7620  
gatactcttt tccctttccc agcggcaaca aatgtgactg cagaggcgtg ccgggatgga 7680  
aggggtggaa agaactagac aagcggaggt ggtccagccc cgcccaaagg gctgggtct 7740  
ggggcccgcc cccactcgtc cacctcgcgc caaggaggga ccgcacaacgg gcgctccccg 7800  
cagcggctgg gttagccagc cgccacgtgg gttggcgtt cggtcccag ccgagacggg 7860  
tgtgacttcc tgggctcaga gactgcagta cagccgcaca ggcggccggc gcagccgagg 7920  
ccccaccaca ccgagccaaa tggggggcgg ggcgggaagg cagggcccag cttaaccgag 7980  
agagccgagt gcctcgcgcc ccgggattcg gggccggctg cttctggagc tgaccagagt 8040  
tccatccaga cagcgtgaa cactcttatt aaaaggctat ttaagttta ttttatttta 8100  
ttttattttt gcttctctc tgggagtggt aggaggggca tcgtgctttt gttttgaggc 8160  
aggggagagg ttaaaattta ctccaagcaa actttttgtc agaaaaatga gactttatac 8220  
cgtattccat ttgttgactt tttaaataca ttcaggtcac tctaaagccc attcttagtc 8280  
aattaacaga ttaatttcgt tccagactgt cataaagagt tcgctggaa tggcaaatc 8340  
tgccattcga caagactaat agaatttcat gggggcaaac atggaagggc tgcgagtga 8400

-continued

---

actgaggttc	ctgtgectca	atgaggaacc	ttagcgggga	cctggcgggg	actggaagct	8460
cagcacctat	ttagggagtg	actgagctgg	aaacctggag	aagcttggtg	cagacgagac	8520
ggtgcttctt	agagetgcat	ttaagtctct	cccattttgc	tgcattttgg	aagaactgga	8580
ttttaagat	cttttccatc	acctcagatt	ctgtagctaa	cagccacgtc	tttttctttt	8640
tgtctgctga	aagctgtgac	ctaattgttt	taagcacttc	ttgacaattt	cgccaaaccc	8700
atcctgaaaa	tattcacata	aagtcctaaa	ataggatggt	gtgagtgtaa	tataaaaaata	8760
ggaaggtacc	aatggaagca	acaattagtg	ttctgagata	gaagcgggtc	ccagtttga	8820
acttgaaat	gttaagagca	gtacatcagg	gagaaggtct	ccaagagtct	gacttcttgt	8880
tctgagctct	gagtaaagaa	agaaatagtg	tagaattgta	ttgtctgccc	gcagtaatat	8940
tgtaaaggg	tccaatggat	tccaatgtga	ttgagcaatt	acttctctca	gcaccaagtc	9000
cacagtctct	aatcttcttg	tgagaagatt	aggtaggaag	cgggatgaga	ttcaagacaa	9060
tcatttagga	tttatgtcag	aattcaaagt	tttgattctt	ttggtagata	aaaacagagt	9120
tggttttaa	ttgtatttcc	tgctttcaaa	aaaacgaggt	ttgaaaggaa	tccccctaga	9180
tttacatagg	tttacagatg	taagtttctt	ctttttccct	gtgtttgcaa	actggaggga	9240
gatgcctttg	atattccaga	atacctgtat	ctatacactt	gtaaattctg	ctaggaaagg	9300
acctaaacag	ctaaacaaa	ggtgaccatg	attaatgtta	aagatactac	ccattgttct	9360
ctctttgtt	gtattattag	gcttaaatac	ttggagttta	atgattggtc	acttaattga	9420
gtaacagagc	ctttgagagc	gcaaagatca	actaactctg	tctcccaaag	tctggcaact	9480
aagccctct	ttctggtgta	gatttaaaca	tttttgttga	cgtcattcat	ggacaaacat	9540
cggtaaacac	tgtattttaa	atgggatatt	gcctctgtat	ggagcagaac	ttctcaagac	9600
acatggttct	tttgaagaa	agaattaatg	actgtagtaa	ggattttttt	caccacccac	9660
gtagagaaac	cagttaggac	tgacaaacac	tcattgttgg	gagttacttg	ctctggtctg	9720
agaaaatag	ctgattgtta	cataggctctg	tattttacca	gaaaaacaaa	agcatgtact	9780
ctccctgatt	tggaaatcaga	aaagggggat	tagatatgga	tgtcttgccc	aagctccatt	9840
tcaattagtg	aatttttggc	agtttctctt	tgtagcctag	acttggtaaa	ttacatata	9900
ataaagtctg	tgtgtgtgtg	tggggtgcgt	gtgtgtgtgt	atgtatacag	ggttttttaa	9960
cactgctaag	ttcttaccat	tgaatgtggt	agaggtggtt	gtatttcaga	ggtcagtaca	10020
aagaatcttt	gtttaatgtg	tcagttctgt	ttcctgaaac	acacotctgg	gtagattttac	10080
caactataaa	aggatacaaa	tgtatatact	ttccacatct	gtatacatcc	tacctcattc	10140
ttctaaggat	ttaagaggt	gtttatgtct	ataaaatgag	attacatatt	ttacaataaa	10200
catatttgat	aaattataat	ctttttatta	ttaaattatt	gatttgatat	tacttgaggt	10260
atagtggaat	gagcactaga	cttggaaatca	gaaagtttga	gttggactca	ctagctgaga	10320
aaaccaagca	aagatattac	ttttggacct	ttagtgtcat	aatctattaa	atgaatatgc	10380
agttataaag	atcaaaagag	taagtacttt	gaaaaactat	acaaaaaagg	gatgataata	10440
tattatgtct	tgggattcat	attgatctaa	ttctataatt	ttaataataa	ctgcattacc	10500
caagccatgt	agtttctact	gaatctataa	aataaatcag	agaggggcat	attgttttaa	10560
aaattgttta	gtttcattaa	tcagaagaca	gattacagaa	attgactcca	atgaagtgtg	10620
acattcagtg	gaaaggaaga	attattcaat	aaatggttcc	aggacaattg	atgggtccat	10680

-continued

---

gcaaaataag aatcagttct tacctgattt ctcatcctaa aagaaatata agatgaatca 10740  
aatatgaaat gtagggaata actttttaca aaatgataaa agaaaatcat gggattaaaa 10800  
aaaataagtt aggaatagag aaacctgttc taagcaaaac ataaaacca gtgccattct 10860  
aggaaatact gataggtctg actaagtaaa aattacacaa aagcaataa tatatttgga 10920  
aaacatctac gacatttgag agacaaaaag ctaataactc ttgatattcc atggactttt 10980  
atgaataact ttgaaaagtc cacatctatc aaaatatgag caaaaattat gtacaggtag 11040  
tggcagaaag agacatagag ttttaaaatt atagaaaaaa tcgctcagcc tcactctgtaa 11100  
ttaaataaat agtaattaa acaccaatga tatgctacac cttttttca ccaatcagag 11160  
gagtaacaaa aaaaaataat ttttttctt ttttttttt ttgagatgga gtctgttact 11220  
caggttgag tgcagtggca cgatctcagc tcaactgcaac ctccgccttc caggttcaag 11280  
caattctcct gtgtcagcct ccctagtagc tgggattatg ggcacccacc accatgcttg 11340  
gctaattttc gtgtttttag tagagtggg tttcaccatg ttggccatgg cagttgatct 11400  
tgaactcctg acctcaagtg attcaccac ctcagcctcc caaagtgctg ggattacagg 11460  
catgagccac tgtgcctggc aaaaatgaaa ttaaaaagtt tgatgcaagt ttgtgttggc 11520  
caggttgagg agaaactggg tctctcatat actgttgtat gaaagtaaat aggttcctct 11580  
actcaaaaa ataatttggc aatcactttt ttttgattct ttatcaaaag tttttaaag 11640  
atgtatatac tctttgattc agcaattoca cttctcgaca ttattatag gatggatatt 11700  
ggtagcaata tcaccaaga caaatggaca aggatgttct ttgcagcatt gcttctaaga 11760  
gcaaatatct gaaacagtg taacaggctc atcataggga accattttaa taaaagttgc 11820  
tgtatccacc taatggaaaa ctatgcagct gtcaaaaata acatggaaaa attgtatttg 11880  
ctgatttggg acattctcca aaatatattg tattctaagt gaaaaaattg atagataaaa 11940  
ctgagtatat attatacatc taattatgta aaatgggtgt atattcattc ttttaacaaa 12000  
tatttactga gtgttctagg cattgtggat acaggagtga acaaaacatc aaagactgct 12060  
ctgttggaac ttttactatg gtaggtagag aaaggcaatt cccaatcaca tatgcaaaat 12120  
ggtagatagt gctaggatga aaatagagca tgggaaaggg aatagagtac agtatgtctt 12180  
aggtagagta atgagaagg agttcaccag gcagagaaca gaggagtgtc aaggggagac 12240  
aatgtgaaaa agcttagcca agtccaggtt caggaattca catgatgcta ttacttgtat 12300  
gaagtaaat catgaaattt cgcacctaaa tcactgattt tctctctgct gtacaatgat 12360  
tctttctgtg agaaagaacc agtgagcaat atgatatgtg atgctcatta catccactt 12420  
gtagctgaag taccctagaa agttgggtgt ctagggtacc taattagcag aatgggtggc 12480  
aaaaccaccg tgtctctgtc ctttaagtcac cccttgcag ctgagcttcc ccaaaatcaa 12540  
gggagaatac taaatatcta ttttaatagaa tcttgcttt ctctggagg ctgcccattc 12600  
ctggtgtgtc ccttatctc cctgcagaca tttctcatgg ataggagtcc cattgacctt 12660  
gtcttttatg tctcccaggc aattaaacag tctcagtacc ctgtacctgg attgttcaat 12720  
gactttctc ccttctcatt gacttcactg gagtagaagg ctaaataatag gcattcactc 12780  
tcatactctc caggtcttcc catgcagaca gagttgttat cactagtaga cggtgcttat 12840  
aagagagaca tggaaaagtg gaagaatgga aatagggtat tggagtgggtg ctaaaaacaa 12900  
agaaagaggc ttctgaaagc tttccattta atatgatcaa acatagaaga caaaagatgt 12960

-continued

---

aacttaatga ccaaggatag ggacatagcc ttgagaaaat tacagaagtg gaatgaacgt 13020  
cttgactcc aaagaagagg gaatggaaat gcagttctga ggcagctaaa gtcaataaag 13080  
ccctctgata tatttactgt agaaataaat gatcatcttt taattcaagg tttgcaaaact 13140  
taggtgacta caggggagag gcatggagtg gaaatgggta agagactaag tgggggcagt 13200  
aatggccttg agcacacaga ctttaagggga gagactgctg cttgggtaat gctcatgatc 13260  
agccagctga aatatagact tagtggtgct agacctctg attttttgaa agaagttaga 13320  
gattttatat ttagtgtaaa cagtctaatt tttaaattga gtaacttatt caaacatatt 13380  
tatgcaactt attcaaatgt atttagagca tgtctgatct cctccttgta ttcttaccta 13440  
gtcttagggg taactttaag gaaagtgatg cgtatacttg ttgtacctat gtagatggag 13500  
gaccattctc ctctcccaaa caaaagacac ttttataagg attgagtgtg gcacaaagag 13560  
ttcgattttt agtattttaa ttgagttaga aactcaaata cagaagatcc tttcctgggt 13620  
cacaccttg gtttttttaa atgagatttg actctctaga tattgacaaa catcatatta 13680  
ttgggtcatg gtaaatcttc aagtggactg attcaaaagt gttcaaaagt atttcagggt 13740  
ttaaggaatt gttggcacgg agtttcaaag gtgtttctgc ataagaaata gaggcgttga 13800  
tcatgtatth gatctacaaa cattcctcaa ggatatacco tgaacctggc aatgaagaaa 13860  
tattgtggga attattaagt tgaataaagt actagctctg ccttcagaa gcctactggt 13920  
ctacagagga gattagacat gtttacagat aatttgaagt acatagcaat aaattatagg 13980  
atgatttatt ccaagcataa agtgtatatt gctactatag tggtaaccta ggttataatt 14040  
ttcaaagtga agagcccaaa atactattct ttttactctt agtgataga aactatggtta 14100  
gaacatctgc agaaagtctg aattagtaat ttaataata aatattgagt actgatata 14160  
tgtatatcac tgagatttat tttatataca ttattttatc aattoctatg atcaactcgg 14220  
taagtactaa ttttcacaa taagtagtga cagaaaatga ggtttagaga gatatagtat 14280  
tttacccaaa gtcacagaat ataagtagag aatgcagaac ttgcattctg ggtatttga 14340  
cagtgagacc cagactttta aaattaggct tataaagatt tcctgaaaa aaagaaaaat 14400  
aaaaataga gaacaaatth tcctgaagaa aagtaggaaa ttattattta taagtataat 14460  
taaagagaga tctttaatth cttttttttt tttttttttg agacggagtc tcgctctgct 14520  
gcccaggcgc aatctcagct cactgcaacc tccacctccc gggttcatgc cataaagaaa 14580  
gtttctatth gattggtccg acctgattgt ctactagaac ataggcttth cagtaggggt 14640  
aacaagtcath aacagttctg gtgatataag ataaatacaa actaataata aggttaatth 14700  
agaattatca acatatgtth ttcattgtga attatcagtt ggtctataga tggaaaatat 14760  
aagatatata atgtattagc ttaaaaaatg ctcttcagag agtgtctaga attatagaat 14820  
ttaaaaaatth tgttttgaaa aggtctaatt cttatttcaa aatatataaa gaaatcaaat 14880  
gcaaaaaaat taagagaaaa ggtcttaatt aattaaatgc tgtacaactc caagccctat 14940  
ttatgttcaa gatcagttth gtaatttatg actaatttca gcattagctt gttttcaatc 15000  
taggtttgct aaataatata gaatttctth tccaatgaag tttttcaggg tcccagctga 15060  
aaaatatata ccagtttaga tgctttcaac tgtaaggaaac aaaaagtcac gattttattg 15120  
ggtttaaaaca ataatgacat ttataatctc atataaaatg aagttcaaag gcatggcca 15180  
ctctaggtct aatccttaag gacttacgcc ttgcctttgc tgattcttth tcagcttcat 15240

-continued

---

```

ttccagaggt atctcactta ctcatggatg caaaatccct acaacagttc cagcagaaaag 15300
attccagcat gataatgtcc agggaagaaa actgaccacc tcttcctttg tctctttctc 15360
aggatgaatc tatctttccc agggaccott cactggagga agagatcccc tcacatctca 15420
ctggccaaaa ctatattata tactgttoca aatccaatca ttggaaaagg gaatgggtacc 15480
tctataattg acttaggata tcaggatcat cccttgggct ggggtttgac ccaggattct 15540
ccaaagtatg tgactgaggg taggcactag aatggaacca gagttgtgtt agaaaggaga 15600
aacgcatggc tattatagaa tagttctaaa tgctaccgag gagggcacia ctgtaaaaaac 15660
caaataatct tttgccattc ttttgaagtt ggcactttga tttctagatg gttccccaac 15720
acaggttcct ctctctctca tcattatagc tgcttgaaa tttgagctgg aagggaacat 15780
tctgagaccc agattgttaa atgtcttttc caaagtcag caataaatta aatggcaaag 15840
ccagggcagt tcttgactc agtacaggtt atttcttttc attctttact cttgagactt 15900
tagaactggt ggtactgctt taaaattcat ggcaagaact ggctactttt gtaattaaca 15960
cctccttata atacatttgt tttgtttgct tagccagcta gaaactacat ggagtctgtg 16020
ctttaaaaag cctgccgaag tccttattct ctgttttggt attatgtgca tgaaccacca 16080
attggttctc tctcacctta cacttgatga agatgccttt ctttcaacat ctttctctat 16140
tgctcccat cttctcttgc tctatttatg atcagctgct tgtttctaaa tagactttgt 16200
ggtcaccat tctttttgt gccagctctc atccactagt tacttgaact gtggtctcta 16260
ttgttctca tacagctac tcacttggg ttgtcacaca catccacatt gttatatgtt 16320
cccaaactgt attctggaat gattttgga atggctgcat tggatgagat tcaaaactaat 16380
aattaaagca ttgagatagc ttctgctatt acaagtttac tctgttttt atagtccaag 16440
aggagctact tctctactc ttattactta atgcttaata ctacccttat tacatacaat 16500
gcacaaaaag catgtgattt atgaccact ttaaactga atgcttgtga tttactttgt 16560
gtttttcttc atttctagg 16579

```

&lt;210&gt; SEQ ID NO 3

&lt;211&gt; LENGTH: 3000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo Sapiens

&lt;400&gt; SEQUENCE: 3

```

acatttaaat tgcataattg gcttgaata tattacatat ttaaaagaat tactattttt 60
ctgaatacca ttaactacta taatagcacc acagtgacta agagtggact taaataaagg 120
ctggcttgaa actcaggtct gccatttatt agcaagcttc taaaaattct gagcccttag 180
ttttctcaca tgtgaaatgg agaaaataaa tctgcaaaat taattaaatc tgtccattca 240
tatttttctc tccagtgtat attaactggc attcctcgtt aggccagaat gtgctctcaa 300
ccatgctcca aatccgcttt gtgccaaccc cactgccaga accctttcta ccttgagaac 360
cagaaaagga aacattatgc ctggcaatgc ctacaccctc caaaataaat ctgcaggaaa 420
gaacaccocag taagtatgta gacagcaac gactgccttt atcattttaa atttacaaca 480
ccaccttttc tagagcctct taagcattgt agataattcc ccactcatta aaaaataaat 540
tgtaaccata agtattcagg gttgatactg cttttgaatt agacagtgct catatcagtt 600
gcataagacc aacctaaagt agaggatgaa atcttttttc tgaacctttt tcagaacgta 660

```

-continued

---

acttagtgaa tatattaaaa ctaaactttc tttgaatggg agtaatttct acggattaat	720
ctgtaatctc ttagaccaca cctaaggtaa tgtagagggt gttgtatcat aggcctttgtg	780
attagagacc actggatttg tctttgaaa agtcactcat gattctttgg gctttggttt	840
cctcatcttt tataccggct taataatgac caccgtagag ttgtcatgga agttaaata	900
actcatgtac cataagtgac caatacaatg attgacactc agtgatttat caataaatta	960
aaacatttat tatcaatatg acagagaagg tgccgctaaa atagacaata ggtttttgga	1020
agaggtgatt aaatggatgc aaaatttatg gattgtttat tccgtctacc tttgctgtgt	1080
ccctggttg tggcatacac acgtgtgggt ataaaatcgt aaatcctatg tagtcgcgta	1140
gtgcatgctc agaaggctta gacacgaaat gtcatttcag caatgtgcct agagaagctc	1200
tgacgccgcc ttggaagtaa gtcgttgctg cctgaccttt gggcgtctgg gacggatgcc	1260
tatacctgca cccagcagca ctggaagggg cccaggccct tcgcagcaca gcctatcccc	1320
agaccgctta gtcttcata acatatactc ccacgaaaa gggattttcc tcccgtcaga	1380
aaaagcggcc cagtctggtc tgggttggtt tttatttcac gttgttgcaa gtaggcgaag	1440
tcccttctgt ctctccctt ggggtaagtg gaaaggagtc cggcaggggg cccgcagtgg	1500
cctgcacagg ggaactgggt agcgagagag ttccaggcaa ttccgggggc tgccccacag	1560
aagcagtggt ggatcgacag tggctctccg gccacgggag gagagcgcgg tcgcgggtcc	1620
ctcccctcag cctggaggct gcagccgctc gagtcggccc ggggtggggc ggggtggggg	1680
cggcgcggag ggcacggaga ttacggcggc gccacccggg acatccaggg ccccgaggcc	1740
ctgggcggtc cccacgcgag atcgcaaaac atgacaatag gcagtcaccc gaggtcaaat	1800
aaaaacggag tgggtcccc gcgcgcggcc gcccccgcg tccctggcgg cctccccga	1860
ggccccggc ggcctcagca gcccgagta gccggtggcg acgtcgcccc cgccccacct	1920
ccctgcgcaa gtgcgaggct gccggcagcg cggcgcaegc tccggccggt cccggttcc	1980
gcgcaaaact tccatctctg ccacgtgaag ttgtcgtctc cttagagagg gggaaagagc	2040
tgcgggaaaa gccggggagt gacgactgcg gcggctgggc gcgctctctc attttcttt	2100
cttctcttt cccccctgct gcagtcggga gttttggctc ctctccttc ctctcccc	2160
tccgagccgg cttctccctc cgccccgctt ctccccgct tgtgtacgct atttgtgtg	2220
gggtggccga aggggatgct ctgttttcac cagaggcaca gcgcgaagg gaaacttcga	2280
cactggaagg aacgagaata aatacttaat tacggacgca ctgaaccgcg gctgggacag	2340
acacttcggg aacccgagcg ggaccggggc acgaggtgag tgacccttc ttccaacccc	2400
cgccccaggg ctcccggggg agcctgagtt gagagaaccc ccaaactttc cgggaaagtg	2460
cgcgaggctc cgccggggc gccgagcgtt gggactgag gacgcgcagc tggacgggtc	2520
gtgggcgcct gctccccgg ggggcgcttg gaggcgggtt gccccacgcc tgagggcccg	2580
ggcgcctcgg accgcagcgg tgcctctctc cctagaagac gtcccccaagc cccaagggtc	2640
ccttcgagc ctgcctgtcc cttccggggc cggcgcgag cctgcgcgta acggagtcca	2700
tccagcagtc cagcgcgcgg cttctacctg cccccgctt ccacctggca gaggcgcgag	2760
catcggggtc tccccacat ctttcttatg acgtgtatta cttctgatg accccctaga	2820
tggtcaggc gcgaggatgc tgaccagag tccttcggag ggtcacaggc gcctgggctt	2880
tcccgggtgcc ggggtcgtgt gtaactttaa ggctcgcgtt ctaactctcca ggcactgatc	2940

-continued

---

 gggcttttca actgcggcga tcccacttta atagttttta tgtggcgtgg actgaatgtc 3000

<210> SEQ ID NO 4  
 <211> LENGTH: 1984  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo Sapiens  
 <220> FEATURE:  
 <221> NAME/KEY: unsure  
 <222> LOCATION: (9, 17, 18, 33, 41, 49, 50, 52, 54, 62, 80, 84, 127,  
 157, 159)  
 <223> OTHER INFORMATION: unknown base  
 <220> FEATURE:  
 <221> NAME/KEY: unsure  
 <222> LOCATION: (175, 207..209, 214, 1181)  
 <223> OTHER INFORMATION: unknown base

<400> SEQUENCE: 4

agcttcagnc atcaactnnc tctttatcaa tgnccctggg nccacaggnn gngnccttct 60  
 gnccaaacag cagccaatgn aggnatctta aaaggcacat cactttatat ttattcagtg 120  
 gttatgncat cacttaatcc tcaccacaac cttgtgngnt cagtacaatt tgtgntttca 180  
 ttttatctgt gaacaaatgg aggcagnnnt ttgngtaatt tacaagtaag tggcagatcc 240  
 gggggctgtg cctagacaat gtggctccag actccttgc cctccctgcc ttctgetgca 300  
 tcttgattga gtctagccat ctatttcttt aacacacact aattgctgtc tccatgttcc 360  
 aggccctgtg ccagatgcta gaaatacatc cccatccact ggatcttggc atggccatt 420  
 gctcagtgtc ctgctatgtt acgtgttaca taatgcattt attccggat ttagcccatg 480  
 gaaaaaatg ctgaaagaca ttgtatgcat gtttctacaa caaaactatg taacttatgt 540  
 ttcttttatt gctttggcta tcaggaagct cttatccaaa tcagagcaaa tacattagaa 600  
 tttgggcttg tcatttcagt ttgctgaact tttccttctg gccagattt tctattttgg 660  
 tcataaatt ctattgcaca aatgtccttt attgtaaac acctaaatt ctttctaagg 720  
 gaaggctgca tggaaatgat acagtaaggt cttccctgca ttttcttaga ttctattag 780  
 ggaaggcaga cctagatgtc cctcttacc tcagtcccaa agcccccat ttataaaatc 840  
 ccttaagcag tgactactgc tgttctgagt acctggaggt agttcagagc ttctgaaagg 900  
 taaaaatccat atacaaaaag aagtctcttc cgatccagat cccagctttg gtgccagatg 960  
 cacatttgag gagtaggtgg ctagtcagac tctcacctga gcagttaata aaatctattg 1020  
 ccccttaatg gaatttttct tgcaagctcg aattgatctg tcatctttgt gattttgtgag 1080  
 atggcagggg agcaccaaa accatcatga cttgggccac agtgggtggg aaaaaggaaa 1140  
 aaagaaaaaa aaaatccact gccaaagcct gccaggcgta naaagggctg gaactgctgg 1200  
 ggccatttta tctgatttat tggaaataga gtggatctta ttaacatttt aataaagaga 1260  
 atctttttgc actaggtctg aagtggccgc cagtcccccg tgcaattcca ttctctggaa 1320  
 aagtggaatc agctggcatt gccagcgtg atttgtgagg ctgagcccca acagtccaaa 1380  
 gaagcaaatg gtagccacc tccgcggggc tcgctcctcg cgaggtgctc accccgtatc 1440  
 tgccatgcaa aacgagggag cgttaggaag gaatccgtct tgtaaagcca ttggtcctgg 1500  
 tcatcagcct ctacccaatg ctttctgtat gctgctgctg atctatttgg gaagttggct 1560  
 ggctggcgag gcagagcctc tcctcaaaag ctggtcccca cggaaaatat gctcagtgca 1620  
 gccgcgtgca tgaatgaaaa gcgcccggg cgcttctagt cggacaaaat gcagccgaga 1680

-continued

---

actccgctcg ttctgtgctg tctcctgtcc caggtaggga agaggggctg ccgggcgcgc	1740
tctgcgcccc gtttctgcat tcggatcgcc cggcacgggc aggggtgaggg ggctttcggg	1800
gggtcggggc ctccggtcgc ggcggcgaag acagatcggg gctcggtagg gaggtcattc	1860
cgagcccaga gatcctaggc acccccaca cacaggctcc cactctggcg tcegtgcgtg	1920
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtacgtt cgttaacggg	1980
agga	1984

<210> SEQ ID NO 5  
 <211> LENGTH: 7833  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo Sapiens

<400> SEQUENCE: 5

gtctttgggtg agatatgtgt tttacaagtt ttaatggaga aaaatgtaag tattttacct	60
cctgaaactt ggctatttga gtaatgagaa aatagtcact ttccccagga cagtggttct	120
caatcatggc tatgtgtttc tccaggaaaa ctttaaaaat atatataac caatgcttct	180
gtgtcacttc tagggattcc aagtccttga atacgaactc tgcacagta ttctttaatt	240
atccaggatg tttgtatgtg aaatcatgac tgagccccac tgctotaaga tgaataaac	300
tttctcagc actgaaatca caaacttaaa ctacccaaaat taattaaggc catgggaatc	360
aataaggcat agggaagctt ttacattata aaattatttc tttaaatcac agctcattgt	420
ttatatgtta tttgccattg tagaaaaggg tgaaaaata gcaaattaa ttactctcag	480
tttgaanaat tatccgaaaa tgaagatgac gactctgaaa cattgtcaat atcatttgac	540
ctataaataa tgttctaata catttactac aactgatag atacttttcc atatgaatat	600
tatacattaa aactaaggca ataatgcatt tagaacattc tatctatata tatgtatctt	660
aagtaggcta gaaattaaga tatgagttat taagtatgag atgttaaggc gtggggttag	720
aaattatact gtacttcatt atcaataatc aacatatact tcaatatcac atacatttaa	780
ctttaatttg tacatcttta actattttta attatgtgta taaatataag tacacacatc	840
tttatgtatt tatttattca tacctccatt cacttattta tataggggat cccccaaat	900
ccactaccat taaaccatac atttttattt taatcttttag aacaagccca ggaggcaggt	960
attgttatta ctcacatttt acaaatgagg aaattgtcta cagtcacaaa gttactgtgt	1020
cagacatatt agaagcttaa tacatatttg gtgaacatat gcataaaaac agagagacag	1080
acatgtacaa cagctcatct ttacactgag taaaagcttt taacctgtct cagaaacctc	1140
tctgtgaaaa ctgagcaaaa atcgaggat cctttcattt gtcatatagg tataggtggt	1200
accttacttc tccaacaagg atgaatattg aaatgtggat cccaaggccc aactccagat	1260
tttctgaatc cctgatagtg ggacttggaa tttgtctatt gtttcaaagt ttctcaagga	1320
atccatatga tcaaccaggc tcagaaatca ctggatctta ttgcogaagt ttgagaatta	1380
aagtttgggc ctactgcgg ctccacagaa agggcaaatg aagtatcatg gacagaactg	1440
atcgttccc agttagtttc ccctctcaga agctaacagg cagcaatata gcagaaatta	1500
gtgacttatg tcttgtgctc tgaagtcagg cagaatttca cagagtccca gcagtgtcac	1560
tgacgagatt tgtttcttgg ggcaagttgc ctgatgcttt caaagccata ttctttttat	1620
ataaaatgag ataattattc ttgtctcata ggggtgtttt aaagattaaa taaaaataac	1680

-continued

---

atgttctatc ctacatggca caatgcctga cacctaagaa gcaaaggata catcttacct	1740
ttattgaagc aatcagaaaag tatgaaatca tgaaggagat aagagttctg attggcagtg	1800
tatcttattt tcccagggtc atttatztat cttaaactat tcttggtgga gaataactcc	1860
caagccccct acttaagctg tgagtaatct cacactttat aatgatgttc tttccatgag	1920
aaaaaaaaat gttcttaagt tttctggaga aaatatatct gcactatttc tactgaaaaa	1980
tctaacaact ggactctgct cctctgcac aattctagag tgotatgcc acaataaag	2040
tgttctagct caagaagatt gaaagtaa atggtatagt attttaaaat aagaattttg	2100
caaatacatg gtatgattgt gtcatattac tagcaatcat atgatacgca atgcaaagta	2160
cagttcatag acttaaaatt aattctaata agtaaactga ttttgccctg ctggggaaaa	2220
gttaaagcac taatccaatt gctaagtcag tcttgtctac ttctttggta cctagtgaca	2280
agtctaaata atgtatata ttttatttac atattcagta atacaattct ctgctcaatg	2340
agtgatgttc ttctgccact tgggtgtgct tgccagtttc agaatttgtt tcttggtggc	2400
actataacac taagtacaga gtaagtgcaa caaaattgca gcattcccat tgaagggt	2460
ttgcttcaaa ctgtttaata atttaaagga cctctgtgga agcaaccgca tttgttaacc	2520
agttacaacc agtaattaac tcctttggag ttttaactta cttttggcaa aacgtcttag	2580
gaagagcata tattattaga aagtatgcca aaaatttact tagcagaaaa ttcaaaaaa	2640
gttttctct gctaagaggt tctctaaaat tctacttaca tagccaaact ctgaaatcct	2700
agcaggctct gtttcattat cataattact gcataaacac ttttaaggac tttgccttta	2760
gtttcaagca tgacttattt tcataagcct gattagttac cacaccagcc ttgctatgga	2820
aatgacatg ttctcattct ctgctgtaga gttgttaaat cttgatctat atttatggtg	2880
ccttctctgc tgaagcctg tagcgaaga aatttcta atcttgtttt gcaatattag	2940
ttggcagctc tatctaaggt gtattctggt tccttaaaga atttagctgc tctgtctaga	3000
agccgatttt ctgatgcctc caacgtctgg tctaattgat ctgtttta atggagtcttcg	3060
tcggtgagga gcgagatgcc accgactaga atgctgggat ctgctgctta attgccagga	3120
gtgagagaca ctgagattca gaaatctttg gaggtgggag gggagaggga cagtctcgga	3180
cggaggcggg gatgtaagat aaagggatgg atttcacaca ggaaaaaaaaa aaagatttcg	3240
ttgaggcact gaggtgctgc acgatcacat ctctcaaagg agaagttaa aagcaaggaa	3300
gtgggaggag gttggagggt aaagtactta aaagattac tcgggtacaa tttgttttc	3360
tgctggtgct tgcaaaggat agatagtccc gttttcaaag tatatgaatg cctcttttaa	3420
gtgattggga atggacacta attgcctggt aaatgttatc aaatgctctc ctaaatcag	3480
gggacacaga aagaggggca caaaaggaga atttaaatag aaaaaggag gatccggagg	3540
cttttgaag cggggggaga agaaggagga gggataacag agaggaatag agaaggagag	3600
cggagagaag ataaacaaaa caaaaaacag gaatcactga ataatcacac accaaaaaga	3660
aagctcttcc ctatggggca tccaaaacac tgagactgca atagtgaccc cggatcatgga	3720
agaaaatgt tctctccc ccttgtcccc gaaagctctt ggtcccgtta ctggcgacta	3780
aatteccatt aggtctaaaga gtgtgtctaa ctgcctgaag aatgcagcag acggaaggcg	3840
ggtcccgtc tgccgtttgc ccttcccgtt ggagagaatg aaagaaacgc gcagagccag	3900
agactcctgc cgagtttagc cttctctcgt cgtcccaggt caccggccat ccggcaaaaga	3960

-continued

---

cccagtaag gaacgcagg tcaactgcctg ggccaacaaa tggagccgc tctcccctc	4020
ccggacgccc ctgcccggcc gatgctccc gcaaccacc cgcggcgat gcagaggagc	4080
ctttctctt ctctcagacc acttgctccc accaatctga ccttccaaac acatctgacc	4140
gcacctccca ggtggacaca ctaataggct acgggctgga gaggagcggg tgatgaggag	4200
agggattcaa acctgcgaac gcttgggctg ggtcggagct gcggggggcc tgggaggaga	4260
gaggggagaa gagagaagga aggagagcgc ctgccgggat ggctgagctg cctcggcag	4320
cagccttggg gttgcacgct cttgtggag atgctgctgt tgcttccagg tcggcaagag	4380
cggttctaac accatcgctt ctcaccctct ttctgtaaa tccctagaga aacgtccctg	4440
gcctctccgc cgcgacatcc ccagcctgca tcccctaca gcctaggcgg cgcgctccc	4500
cacgctggag cgcggctgc cagcaggacg cctctcccgc cgcgactcg cccctctctg	4560
ccctgctgct gctgctctc tgacacctcc gccccacca tctccagctc ggagagacgc	4620
caccagccc cggcccgcac tcgcccggc gggtcacgc cgggaagagg gcgctagtcc	4680
ggaccccgcc ttcggtaggg ggcgtctctg agcggagagt gaggcgaatg gtatatgagt	4740
gtcgggtag cccaccctga agcccagct tctcatttga gccatgccc gcctagcccc	4800
actegggcca gcgctggcg agcagaccca tctgtgctt ccgcccgcgc ctctctctg	4860
catccttga ctaactcgtc gaccctccc tcccgggacc tgcatcctgc tccaccaatc	4920
agagcccgac tgctcttcc cactgaccc cgggcgggct gaggacctgc tgcttccaa	4980
acgcccagag gatgcggggc gcagagctcg agaggcggct gccgggctgc ggggcgcctt	5040
gactctccct ccacctgccc tctcgggct ccaactcgtt gccctggac tcccgtctcc	5100
tctgtctct cggcttcca gagctccctc cttatggcag cagcttccc cgtctccggc	5160
gcagcttctc agcggacgac cctctcctc cggggctgag cccagctccc ggatgttct	5220
gaaactctcg agatcatgcg cgggtttggc tgctgcttcc ccgcccggtg ccaactgccc	5280
cgcccggccc tctgctgccc ccgtcccggc gatgctcagt agcccctgc cggccccgc	5340
cgatcctgtg ttctcggaa gccgtttgct gctgcagagt tgcacgaact agtcatggtg	5400
ctgtgggagt ccccggcga gtgcagcagc tggacacttt gcgagggctt ttgctggctg	5460
ctgctgctgc ccgtcatgct actcatcgta gcccccggc tgaagctgc tgetttccct	5520
acctcctaa gtgactgcca aacgcccacc ggctggaatt gctctggtaa gtccagaacc	5580
cccgtcccc accctttaa tcccagaag aacacgcgta tccagcacag accagcctac	5640
cctagcgcgc ctctcagcc cctcacctcc tactgcctca gaccctaat accaccacc	5700
tctatccaga gaaacaagg gaactgttc aggccggggc gtgaggggtg gttctgggat	5760
gggcagaaag tgcagggtga gcaggaaacc tttgcatgct tgcgcttaca ttggagctgc	5820
gaggattttg agaaatatta aacgggatgg tttctgggt tcaactgttt gaaagagcac	5880
caatcctagg gaaacactg aaacagaagc tttgtcatca ttaaagaaaa aagtcttact	5940
aggatgagga agaaataact ttatgagaaa gaatgagcga gaaagcaata aatcaaatgg	6000
tgactgcagg ggaatcgctg attcctggca aagggtccat gaggtcgcac tggctcccgc	6060
ttgaagacca ggtcacacag attctagagg agctgggttt caatagaatt tctctctctc	6120
tctctctctc tctctctctc tctctctctc tctctctctc tctctctctc tctctctctc	6180
tcattccctt ctctcctagg cggcaaaaga cattggtttt gcagtccaga tatgccctc	6240

-continued

---

```

tctttgcttc cctaagcttc aaggtagtac aggggagttg agaaaaagaa cactttgcgg 6300
gtctcccagg ccggagtggg catgactgag gctggtcagg ctccatgtag gcgagccgag 6360
ggcggaaccg acttcagtg ggcgtgactc ctccatttct ggacaggctt ctgtggagtg 6420
ggtcaggcac tcttcttctt cgctcgggtt ccttcagatt ctgacggcga acgcttggca 6480
ggcttcgctc tgctgaagct tcccaattaa ataggccag aggatgggag ttgctgcaact 6540
cctagctggc atagcattcg gtttgacagc ctgtagtata gggtagatgt aatttttcat 6600
cttctgtgaa tataattttg ctgtagttaa atctggctct gaataaagtg tctttcaaag 6660
atgtatataa gctgaagtgt atgtaacttt agagaggagg gaatgaccaa ctgtaactca 6720
gggtgaaagc ctgtatagtt cctagtattt actgatgtaa atgccccaaag gaaaattatt 6780
atgcatcatt ctaatttatc ctttacaag acaagttgag atatgcaacc ctattagatt 6840
tgggtcaata gattgttctc ttttttgga gtttctaaat ttggcatttt aataaaactc 6900
aacatgtttc tataacttct tgattcatgc gtacatgtgt gttgtttttg aaagaataag 6960
tttcaacttg ctattgccta atcacttttt agatgcttta ttatggtaat aattatgagc 7020
ctgcaaaaac aatttttgga aatgttgatg gctttgtagt ccaacacaga ctggtttctt 7080
tcattcctag ccottgcatt gttttaggaa ataactaact taaatgtgaa gttgacattt 7140
gcaatcaaga aattacatat ttaccagata ttttaaaggg gactgcataa actaaagaga 7200
ataaactggt tttgcagata ggttgcaag aacttgccac cccgcttcca ccctgttaa 7260
cttagaggtg atcaatcttc atttgagcca aacagaccat cacagaaaac actgtgcctg 7320
tttatcttta ttattgagcc tttgttctct ctttctctgg atacatttca aataaggggt 7380
tgtttcagtc gttgaagcaa aagaacaatt aaagatgggg aaatggtaaa agggatttca 7440
gagatcatca ctagctcttt tccaaaatgt ggagttttgt ggtcataaat attgtccacc 7500
taatgagcaa aaaataaaaa taaaaaaaaa acaggaagca aatgtaagc tttcattcac 7560
cactgtcagt attaacgcaa gctttaaaaa atagcactat cagaaaagga tactaaagga 7620
gaattgacta gaaaagaatt gtggaaaatg gaaacgaata ttgatcactt aactagattt 7680
tgaggttatc agtagacagt gaccttgca g tacagctata gttgttggat ttaaaattha 7740
ggacaagtat ttaaaagctt caaagtatg cttttttttg ttaaaaatct gtaagatgtt 7800
ttaatgactg gagtgttctc tttgaatttg agg 7833

```

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 965

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 6

```

gattaattaa gggattttta gaagttaggc gttttcgttg ttttttttga gtacggaggt 60
tattaathtt tagggggaag agatgtagcg cgaggtaggg gtgtcgtggt aagaaatthc 120
gacgtttttg gggattgagg ataaaggtgc ggatacgatt tcggggatt ttgagtttctg 180
tgatttcgct tacggacggt atatttaggg gtttaattht gttttgtttt aaagaattht 240
aagttagagt ttttgttttc gtttatagtt tcgggatgtc gttgttgcgt ttatcgtata 300
ggtagcgttc ggatcggttg tagtagatcg cgcgttgcgc gttttatcgg gagatgggtg 360

```

-continued

---

```

agacgttgaa aagttttttt tttgttattt tggacgttgt gggcggtaag cgttttagtt 420
tttatttttg ttgagttgaa cgtttaggta tagtgaatt gaaattcggg tttttatttt 480
tgttgagttg aacgtttagg tatagtgga tggaaattcg gttttgcggg atgtgagagt 540
tgttgaggtt acgcgtaatt ggggtgatg gagggcgttt gttcgtgatg tgtgtaggtt 600
tgatgtaagt aggttatcgt cgtgcgagtg tgtggatcgc atcgttcgag agattcggag 660
gtaggtttgg gatacgtttg agtgaatatt ttaggatatt tttttggtta gtatttgttt 720
tttagtgttt gtgatttaga gtgggtatat gttgggagat agtaatgggt ttgggtgtgt 780
gtaaatgagt gtgacgggaa cgcagtgatg gtttgattta ggtagggatt atatagtatt 840
gttatatttg tttgtttttt agtagaggat tgaagtgcgg ggggtggggg acggggtcgg 900
aatagaatgt ttttgggata ttttggtaaa tagtagtcgg aagtaaaggg gtagttgtgt 960
aaacg 965

```

```

<210> SEQ ID NO 7
<211> LENGTH: 965
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

&lt;400&gt; SEQUENCE: 7

```

cgtttgata gttgtttttt tgttttcggg tgttgtttgt taagatgttt tagagatatt 60
ttatttcggg ttcgtatatt tattttcgta ttttagtttt ttattaaaga gtaggtaggt 120
gtgatagtgt tgtgtggttt ttgtttagat taagtttata ttcgttttcg gttatatatta 180
tttatatata tttaaatatta ttattgtttt ttaatatgtg tttattttga attatagata 240
ttaaaaaata gatattgggt agaagagtat tttgaggtgt ttatttaaac gtgttttaag 300
ttgttttcg agtttttcgg gcggtcgtat ttatatattc gtacgacgat gatttgtttg 360
tattaaattt gatatatta cgaataggcg tttttatta tatttaatta cgcgtgattt 420
taatagtttt tatatttcgt agaatcgggt tttagtttta ttgtgtttga gcgtttagtt 480
tagtagaggt agggaatcgg gtttttagtt tatttgtttt gacggttttag tttagtagag 540
gtagggatta aggcgttttg cgtttatagc gtttagagtg gtaagaaaga agttttttag 600
cgtttttatt atttttcggg ggaacgcgta gcgcgcgatt tgttgtagtc ggttcgggcg 660
ttgtttgtgc ggtgagcgta gtagcgggat ttcggggttg tgggaggagg taaggatttt 720
agtttgaggt ttttgaggt agagtagaaa ttagttttta ggtgtgtcgt tcgtggcgcg 780
agttacggaa ttttaggtat ttcggggtcg tgttcgtatt tttgttttta gtttttagaa 840
gcgtcgaaat ttttagtac gatatttttg tttcgcggtta ttttttttt ttttaggggt 900
tgggtggttt cgtgtttaaa ggaggtagcg ggaacgttta attttaaga tgtttttaat 960
tgatt 965

```

```

<210> SEQ ID NO 8
<211> LENGTH: 16579
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

&lt;400&gt; SEQUENCE: 8

-continued

---

tgttttttaa aatgatgttt agttgaatta gttaatggag ataaaaatat ttatatataa	60
ttttttttaa agtttattag ttattttata gattagttat aattttggtt attttaatga	120
gttaattttt ttttagagat attataaatt ttttaagat ggaaaataat tgttattttt	180
ttgaataatg tttattattt aaaataatgt atataaaagt tgtttaataa atttttgttg	240
aattgagaga aaatataaaa aattgaaagt agtttgttg taataagtgt ttgttaagtt	300
ttatagattt tttttttttt ttggaattat tgtaataaat atttgatttt tgattttag	360
gattaaaatg tgatttttta tattttaata ggagtaagta ttagttttgt aatttttttt	420
gttaatagaa taaaaatgga atgtatgttt gttgttttta tatgtcggtt gatatttatg	480
ttgaatattt tttttataac gtagaggttt ttttgttttt tgttttgcgt tttatttttg	540
tttattatag agatattaat gatgttatgt gtttttttgt attattagtg tagtagtttt	600
tattatgttt tttgtttcgg ttttattgta gtttaattaa ataaatggaa acgggtagtt	660
attgaaaata aatagtaatt gtaaaagttt tgttttttaa gaaatttttt gtattgtatt	720
tcgttttggt ttattgattt taagattttt atcgtttttt tegtgtttaga aattagattt	780
tttagtttgt tttgttttag ttagataaat gttttataaa aatgatttat attttttaga	840
ttgtggggag agaagaaaaa gattttgaga aagtgattaa ggagggagta agagtgatat	900
aatttttgat ttaatttagt tggatttgtt tttttatggt tttatttttt tttttataaa	960
aaagatagcg tttttgatgt tttgttttag gtagtttttt aggggtgagtg tagatgtgtt	1020
gagttatttg gttgttttga ggatcggttt agagagtttt tttttttatt aatgaagtta	1080
ttttttttcg gtagttcggg tttttttatt attttatggt ggaatttggg tttggtttgt	1140
atagttttgt cgtgatgtag atggtttttt agttaagttt tagagaaatt tgtttaatta	1200
tgaataggat atttatagta taggagaaaa aaaaatttat taaaaataat attttattgt	1260
aatattttga tttggggggt ggagagttaa aatttaatat ttgatgtttg ggaaaaattt	1320
aatgttttat ggttttttaa ggaatttgtt ttgtgtttta tttattagag taatttatat	1380
atttcggggt aaagtataat gagattgtta ttattatttt gttttttgta agattttttt	1440
tttaagaata aaaagaaggt aataaagatt ttatattttt gtttttttta gaaagttaaa	1500
gagaatttaa agtgtttggt gaaatgtttg tttttaattt gtagtgtttt atttaaaaaa	1560
aaaattatga tataatttta gaggataggt tttttttttt tagggttttt gaaagttaaa	1620
agtatatttt taattttagt tattattgta ttatttagtg aggataattt gtttagttgt	1680
gtttaggatt agttagatta aggagttgaa attaagtttt ttttgaattt gaatgggatt	1740
ggaattttgt tattttttaga gtagtaatag atgatgttat gtttaagaat attttatttg	1800
attaataaaa aaggagaagt ttgaaatgta attttaaaga attttagaag gaggtttaga	1860
tgaagataat aaaatttttt ggagatatat ttgttattta tttttttttt ttttaagtta	1920
ttggtttgggt gtttttgagt ttagttaaat ttttaagat aagattaaag taaggaagta	1980
aggtatatag tttttgggtt atgggttttt ttagatataa gtattgtgaa tgaagatttt	2040
gttttattta tatttgggat tttggagttt ttataaagaa tatttttagt ataaggatgt	2100
gatttaaagg gggatttata gaataatatt ttaattattg tttgatttat atttgagttg	2160
aatggataaa taaatataat tttagttata aaatttgaaa aatatatatt tttttattta	2220
aaaatttttg atagtttttg taggtatttg agaattagaa ataaatgttt ttttattttt	2280

-continued

---

tttttttga tgaaggaaag ttattaaaa tagttatata gaaataataa tataaattgta	2340
gattttatatt taaatatagat tggatatgttt tttttttttt tttttattta ttttttattt	2400
tagatagtta gaggggaaag tttgttttag agatagaaag aagatattta tggggataaa	2460
ttttatttta gttatttgat agtttgaaga agtatttaaat tttttgaata ataattgtat	2520
aagttaaaga agagaaaaa taggtagttt aaagtttttt aatattgtgt atattttaat	2580
tttttaagtt agttttgaga aagtgttttt gttttgtaga agtatttggg taagtgttgt	2640
gaagtttatt ggtattaaaa gaaattaaaa attttggagg attagagaat taaatttggg	2700
aatataaatg aaaatatttt agaattgtaa aagtatttaa tttattttt ttaattatgt	2760
tttattatat atagatggta ttatttatag tgaatgttaa ttattaatat ttttttaaaa	2820
atgggtatta atttgttatt agttgtttgt ttttgattta gttttttttt gttgaaaggt	2880
tttatttaaat tgtattataa agggataaaa gggtaaaaga taataagttt tttagattgg	2940
tagatttgtt atttataaaa tggttttggg taggttattt gaattttttg aaatttaatt	3000
agattttatt aaatttaaga atgagaatat aaatttatat tttgaagtgt tttatagaaa	3060
ggttttttt aatgtttgga gtatatattt taatgaatat ttattttatt ttattttttt	3120
ttatttttga attaagtaat tttgaattta aagttgttat gattagtatt gaaaagatta	3180
ttggattatt aattgtgtga ttttgggata gtaatttttt gtatttttagt ttgtttatat	3240
gttatatatg aaggttgaag tttgattttg ttttgtgatt attattttaa atatttgatg	3300
aaattaaatt ttagtgtttg gaatggtagt ataataaatt tattaagaat aaataattta	3360
ttgtaaaaat atattgattt ttaaatgatg taattgatag ttatattatt gtagagggtt	3420
gataaataat aaaagaaatg aaagatgtat atgggtgagaa ttgaaattat tttgataagt	3480
ttttattttg tttattattt aaaattaatg attatgttga atgtttataa attataaaat	3540
ataaaagaaa ttttataaat gcgtatgtat aggagttaa gttattaaaa gttttaaagt	3600
ataagtttaa attaaattaa ttaagaagt tgagaggaaa aattggtttt tatttttaat	3660
tattattgtt ttgaggtttt atgttttaata taatttttta agtagagggtt ttagagagaa	3720
gagttgtgag gatattttta tatttgtgta gaaggaaaag tttgttattt attttagtat	3780
ttttagtgtt atattgatgt gtatttttga tttattttgt ttttattgta taaatttata	3840
tttgatttta aagaaaagga aaatttaaag tttttttttt ttaaggggat agaaattttt	3900
tgtgttaatt gtttgatttt tttttttgta aggttttatt ggaaattttt tghtaatataa	3960
tgtaggggat ttttttatgt gttgatgttg tttatatagt ggggtgggtt tgattgaaga	4020
aaaaaaatcg tatatacgtg tgaagatta tggttttatt ttcggaaagt atgaaaggtg	4080
attgatattt ttaagaagtt tttgttattt aggaaaatta ttaaatattt tatttagaga	4140
tatttggaag gattgaagga aaggaagaac gaagaaagta gaatttagat ttatgtgggg	4200
agagatttgt ggtagaggaa aagtattttt tttgaattcg ataaggatt tgtttggggg	4260
aattttttgt ttagtttttt attattaggg ttttttgaag tgggtttttt tattgggtag	4320
ttttttggga gtgtagtggg gaatttttat attttttttt taggttttcg aaggatttcg	4380
tttttttagt gtttttttta ggttggtagg agttttgagt ttgatatttt tttttgatgg	4440
gataggttaag ttttggggc gcgtaaatac gttgtaatta agttttttgt tgattttata	4500
gttttgtgtg ttttcgagaa gaagtgatcg tatttaattg tttattgttg gtttgtttt	4560

-continued

---

taagagtttg ggggtttttt tttttaatt tagaattagt tgtacggggg gcggggaagt	4620
gggggtgggg aaggagtggg agggtagtgg ttttcgcgag tagagcgatg ttattgagtg	4680
agtttttgaa tggggagcgt tgttgttttt aagtcgattg gtattttttg ttaggaagaa	4740
acgttaagag gtgggagtgt ttggggaggg aggtaggcgg tttttatcgt aggcgcgggg	4800
agttgttttt tcgttttttc gtttgttttt taagtttgga tttttaggag tggttgaagt	4860
tgcggagcgt ttttgagttt tgtgaatgaa tttttttttt tttttttttt ttttttttcg	4920
ttgagttttt ttttcggttt tgacgggata gtgatataat gatgatgggt gttataattc	4980
gtatttgaat ttgtaggcga gttgtttcga gtttttttgg ggaagaattt taggcgtgcg	5040
gacgtaatag tcgagaatat tagtggttgt ggataggagt tgggattaag attttcggtt	5100
agtttcgtat tttttcgtat tttttagtat cgtttcgtat ttttcgtatt ttttttcggg	5160
ttattacgtt ttttatgtga ttcgtttggg taacgtcgaa tttagtcgcg tagcgttga	5220
gtgaattttt ttttaaaatt gtaataagtc gttttttaag gtaattacgt tttttttgtt	5280
ttttttttaa aaaaaaaaaa taaaaaatat atagaaaaaa attcgcgagt ttagaaaaaa	5340
gaagtaattg gtagaagggt ttaattaagg taaagagttg taaggcgaag ttaagaaaat	5400
gtaggtattt aaaaaatgta ggtaattttt ataaaggttt ttggggagag gtatatagag	5460
ggatttttgt gttgaaaaag atttagataa aagaaattha ggggtgggtg ggggtaaaat	5520
gattaacgga attgggggaa gggagggaat aaattgtaa gaaattatag aaaagtggtg	5580
ggtttttgag ttggagagaa gagagggatt tttggtattt tgattttttt gttggtggtg	5640
tttttaatac gttcaggtga aaagtttgaa tggggattat taagatttgt tatagataag	5700
ttttcgaagt cgttttggtg taggttattt ggttttttag tttttggtgt gtggttagtg	5760
tttggtggtt ttggaagtt attttcgggt agtttttgat agtgcgattc ggcgtttaag	5820
tagtttggga ttttcgcgcg atttgatatt tttagatcgt aggtagtttg ggaggaggtt	5880
cggtcggggg aggtgtagga ttttcgctgt gtttttttga cgatttgggg attggtacgg	5940
tttttttttc gcgtttttgg gtttttttgt tttgtacgcg gtgcgaaggg gttagtaggg	6000
aaggagtaga ggatgggggg tggggttggt ggagtttcgc ggaggtttgg gaggtttttg	6060
ggcgggaaaa gtttgttttg aatcggtagg gatgtgtaat aatttttttt attttgaaga	6120
gtgaaatagg gtttgcgtt tttattttaa taagtaaadc ggtatttaga gcgtattgta	6180
gataaagggt gttcggggat tcgaattagg ggttttttgg ttagtttttt cgttttagatt	6240
ataggagttt ttcgtttttt tataatattt ttattttttt ttagtttttc gtttttagttg	6300
agtaatttta ttggtagggt tagcggtagc ggcgcgcgcg tgtgtcggga gtttcggggg	6360
acgttttcgg ttggagcgtt ttatcgtttt gtagaggcgt gggcgttga gggcgatatt	6420
tttggtcgtg tagatattag ggtattcggg ttgttttggt tcgcggtttg tgttattggg	6480
cgtggagcgg tttgggtgtt aggtagagga gagcggggta gaaaaatagt gttatagttt	6540
aaattatttg tttttaatt taattgcgtt tcggtaagtt ttgtatgttt ttgggagatg	6600
ttcgggaag ggggagaaat ttatacggcg tttggagagt tgttttgcg cgcgtatatt	6660
cgcggtagag tttttttggg tcgcgttttt tttttttgtg tttttttttg tttttatatt	6720
ttttgttttt tttttttttt tttagttttt gttttttttt tttagttttt gcgttttttt	6780
ttttatagtt gataaatgaa tggttagtgt gaaatttttg tttttttcgt gttttaagtt	6840

-continued

---

agtagggagg gaggagcgag ggagggcgtg cgttttttcg gatttgttt tagtttagtt	6900
tataagattt gtagaattta gatgttagg aattgggagt tttgcggcgg gtgtgggcgt	6960
tttttgatgg agaagtttcg tataggcggg gaaaaataag ttttttagat taagcgtgta	7020
ttttttataa ttttggtttt aaggatggaa ggtttttagt tttttttaag ttatttattt	7080
tgtattttat aatttcgggc gtatatttag gagtttgag gattttgaaa aaaggttttg	7140
gtttgtgtaa agtgtagaga tgttttttcg cgagtcggtt ggaacgtacg cggcgttttg	7200
ggtttagttc gtcgttttagt ggacgaacgt atatggcgg ggtggagtta ggtttcggag	7260
gtttagacgc gttatttttg cgcgttttaa gaaagataat atatataatc gttttttaa	7320
atttatgatt atttgaattt aacggtaggg tttcgtttta ggatattgta aaagaagggt	7380
ttttaattta aaatttaaat ttttttaatt ttagggcggg cgtcggacgg gaaagtggag	7440
agaagggcgg tagtgggagg aaaaaagaaa ggaaggaag ggaagtggaa ggaagagga	7500
tcgtggaggg gaggaaaagg agaaaggaac ggagatgggg tagatggtat ttaaatagtt	7560
aggaacggat cggaaaggggt cgtttcgggg tttgggttg aatttcgaga ttcgttttag	7620
gatatttttt tttttttttt agcggtaata aatgtgattg tagaggcgtg tcgggatgga	7680
aggggtggaa agaattagat aagcggaggt ggttttagtt cgtttaaagg ggttgggttt	7740
gggttcggtt tttattcgtt tatttcgctt taaggagga ttcgtaacgg gcgttttttcg	7800
tagcggttgg gttagttagt cgttacgtgg gttggcgttt cggtttttag tcgagacggg	7860
tgtgattttt tgggttttaga gattgtagta tagtcgtata ggcggtcggc gtagtcgagg	7920
ttttatttta tcgaggttaa tgggggcggg gcggggaagg tagggtttag tttaatcgag	7980
agagtcgagt ggtttcgttt tcgggattcg ggttcggttg tttttggagt tgattcgagt	8040
tttatttaga tagcgtttaa ttttttatt aaaaggttat ttaagtttta tttattttta	8100
ttttattttt gtttgttttt tgggagtggt aggaggggta tcgtgttttt gttttgaggt	8160
aggggagagg taaaaattta ttttaagtaa atttttgtt agaaaaatga gattttatat	8220
cgtattttat ttgttgattt tttaaatata tttaggttat tttaaagttt atttttagtt	8280
aattaataga ttaatttcgt tttagattgt tataaagagt tcgtttgaa tggtaaatat	8340
tgttattcga taagattaat agaattttat ggggttaaat atggaagggt tgcgagtgta	8400
attgaggttt ttgtgtttta atgaggaatt ttagcgggga tttggcgggg attggaagtt	8460
tagtatttat ttagggagtg attgagttgg aaatttgag aagtttggtg tagacgagac	8520
ggtgtttttt agagttgat ttttaagttt tttattttgt tgtatattgg aagaattgga	8580
ttttaaagat tttttttatt attttagatt ttgtagttaa tagttacgtt tttttttttt	8640
tgtttgttga aagttgtgat ttaattgttt taagtatttt ttgataattt cgttaaattt	8700
atttgaaaa tatttatata aagttttaaa ataggatgtt gtgagtgtaa tataaaaaata	8760
ggaaggtatt aatggaagta ataattagtg ttttgagata gaagcggttt ttagtttgta	8820
atttgaaat gttaagagta gtatattagg gagaaggttt ttaagagttt gattttttgt	8880
tttgagtttt gagtaagaa agaaatatgt tagaattgta ttgtttgttc gtagtaatat	8940
tgtaaagggt tttaatggat tttaatgtga ttgagtaatt atttttttta gtattaagtt	9000
tatagttttt aatttttttg tgagaagatt aggtaggaag cgggatgaga ttttaagataa	9060
ttatttagga tttatgttag aatttaaagt tttgattttt ttggtagata aaaatagagt	9120

-continued

---

tggttttaaa ttgtattttt tgtttttaaa aaaacgaggt ttgaaaggaa tttttttaga	9180
tttatatagg tttatagatg taagtttttt tttttttttt gtgtttgtaa attggaggga	9240
gatgtttttg atatttttaga atatttgtat ttatatattt gtaaaatttg ttaggaaagg	9300
atttaaatag ttaaattaaa ggtgattatg attaatgtta aagatattat ttattgtttt	9360
ttttttgtt gtattattag gtttaaatat tttggagtta atgattgggt atttaattga	9420
gtaatagagt ttttgagagc gtaaagatta attaatattg ttttttaaag tttggtaatt	9480
aagttttttt tttttggtga gatttaata tttttgttga cgttattat ggataaata	9540
cggtaaatat tgtatttaaa atgggatatt gttttgtgat ggagtagaat tttttaagat	9600
ataggtttt tttgaaagaa agaattaatg attgtagtaa ggtattttt tattatttac	9660
gtagagaaat tagttaggat tgataaatat ttattgttgg gagttattg ttttggttg	9720
agaaaatag ttgattgtta tataggtttg tttttatta gaaaataaa agtatgtatt	9780
tttttgatt tggattaga aaaggggat tagatatgga tgttttggt aagttttatt	9840
ttaattagtg aatttttgtt agttttttt tgtagtttag atttggtaa ttatatata	9900
ataaagtttg tgtgtgtgtg tgggggtcgt gtgtgtgtgt atgtatag ggtttttaa	9960
tattgttaag tttttattat tgaatgtgtt agagggtgtt gtattttaga ggttagtata	10020
aagaattttt gtttaagtgt ttagttttt tttttgaaat atatttttg gtagatttat	10080
taattataaa aggatataaa tgtatatatt ttttatattt gtatatattt tattttattt	10140
ttttaaggat ttaagaggtt gtttatgttt ataaaatgag attatatatt ttataataa	10200
tatatattgat aaattataat tttttatta ttaaattatt gatttgatat tatttgaggt	10260
atagtgaat gagtattaga tttggaatta gaaagtttga gttggattta ttagttgaga	10320
aaattaagta aagatattat ttttgattt ttagtgttat aatttattaa atgaatatgt	10380
agttataaag attaaaagag taagtatttt gaaaaattat ataaaaagg gatgataata	10440
tattatgttt tgggatttat attgatttaa ttttataatt ttaataataa ttgtattatt	10500
taagttatgt agtttttatt gaatttataa aataaattag agaggggtat attgttttaa	10560
aatgtttta gttttattaa ttagaagata gattatagaa attgatttta atgaagtgg	10620
atatttagtg gaaaggaaga attatttaat aatggtttt aggataattg atgggtttat	10680
gtaaaataag aattagtttt tatttgattt tttattttaa aagaaatata agatgaatta	10740
aatatgaaat gtagggaata attttttata aatgataaa agaaaattat gggattaaaa	10800
aaaataagtt aggaatagag aaatttgttt taagtaaaat ataaaattta gtgttatttt	10860
aggaaatatt gataggtttg attaagtaaa aattatataa aagtaataa tatatttgga	10920
aatattttac gatatttgag agataaaaaa ttaataattt ttgatatttt atggattttt	10980
atgaataatt ttgaaaagt tatarattatt aaaatagag taaaaattat gtataggtag	11040
tggtagaaag agatatagag ttttaaaatt atagaaaaa tegttagtt ttatttgtaa	11100
ttaataaat agtaattaaa atattaatga tatgttatat tttttttta ttaattagag	11160
gagtaataaa aaaaaataat tttttttttt tttttttttt ttgagatgga gtttgttatt	11220
taggttgagg tgtagtggta cgatttttagt ttattgtaat tttcgttttt taggtttaag	11280
taattttttt gtgttagttt ttttagtagt tgggattatg ggtatttatt attatgtttg	11340
gtaattttc gtgtttttag tagagtgggg ttttattatg ttggttatgg tagttgattt	11400

-continued

---

tgaatttttg attttaagtg atttattat tttagttttt taaagtgttg ggattatag	11460
tatgagttat tgtgtttggt aaaaatgaaa ttaaaaagtt tgatgtaagt ttgtgttgg	11520
taggttggg agaaattggg ttttttatat attggtgtat gaaagtaaat aggttttttt	11580
attttaaaa ataatttggg aattattttt ttttgatttt ttattaaaag tttttaaaag	11640
atgtatatat tttttgattt agtaatttta tttttcgata ttattatag gatggatatt	11700
ggtattaata ttatttaaga taaatggata aggatgtttt ttgtagtatt gtttttaaga	11760
gtaaatattt gaaaatagtg taatagggtt attataggga attatttaa taaaagtgt	11820
tgtatttatt taatggaaaa ttatgtagtt gttaaaaata atatggaaaa attgtattg	11880
ttgatttga atatttttta aaatatattg tattttaagt gaaaaaattg atagataaaa	11940
ttgagtatat attatatatt taattatgta aaatgggtgt atatttttt ttttaataaa	12000
tattttatga gtgttttagg tattgtggat ataggagtga ataaaatatt aaagattgt	12060
ttgttgaat ttttattatg gtaggtagag aaaggaatt ttaattata tatgtaaaat	12120
ggtgatatgt gtaggatga aaatagagta tgggaaaggg aatagagtat agtatgttt	12180
aggtagagta atgagaaagg agtttattag gtagagaata gaggagtgtt aagggcgat	12240
aatgtgaaaa agtttagtta agtttaggt taggaattta tatgatgta ttatttggat	12300
gaagtaaat tatgaaattt cgtattttaa ttacgtgatt tttttttgtt gtataatgat	12360
ttttttgtg agaaagaatt agtgagtaat atgatatgtg atgtttatta tattttttt	12420
gtagttgaag tatttttaga agttgggtgt ttaggtatt taattagtag aatggtggat	12480
aaaattatcg tgtttttgtt ttaagttat tttttgtag ttgagttttt ttaaaattaa	12540
gggagaatat taaatattta ttaatagaa ttttggtttt ttttcggagg ttgtttattt	12600
ttggtgtgtt ttttttttt tttgtagata ttttttatgg ataggagttt tattgatttt	12660
gtttttatg ttttttaggt aattaaatag ttttagtatt ttgtatttgg attgtttaat	12720
gatttttttt tttttttatt gattttattg gagtagaagg ttaaatatag gtattttatt	12780
ttatattttt taggtttttt tatgtagata gagttgttat tattagtaga cgggtgttat	12840
aagagagata tggaaaagtg gaagaatgga aatagggtatt tggagtgttg ttaaaaaata	12900
agaagaggt ttttgaaagt tttttattta atatgattaa atatagaaga taaaagatgt	12960
aatttaataa ttaaggatag ggatatagtt ttgagaaaat tatagaagtg gaatgaacgt	13020
tttgattttt aaagaagag gaatggaaat gtagttttga ggtagttaa gtttaataag	13080
ttttttgata tattttattg agaaataaat gattttttt taatttaagg tttgtaaat	13140
taggtgatta tagggcgag gtatggagt gaaatgggta agagattaag tgggggtagt	13200
aatggttttg agtatataga ttaagggga gagattgttg tttggttaat gtttatgatt	13260
agttagtga aatatagatt tagtgtgtt agattttttg attttttga agaagttaga	13320
gattttatat ttatgttaa tagtttaatt tttaaattga gtaatttatt taatatatt	13380
tatgtaattt atttaaatgt attttagata tgtttgattt tttttttgta tttttattta	13440
gttttagggg taatttttaag gaaagtgatg cgtatatttg ttgtatttat gtagatggag	13500
gattattttt ttttttttaa taaaagatat ttttataagg attgagtgtg gtataaagag	13560
ttcgattttt agtattttaa ttgagttaga aatttaata tagaagattt ttttttgggt	13620
tatatttttg gtttttttaa atgagatttg attttttaga tattgataaa tattatatta	13680

-continued

---

ttgggttatg gtaaattttt aagtgattg atttaaaagt gtttaaaagt attttaggtg	13740
ttaaggaatt gttggtacgg agttttaaag gtgtttttgt ataagaaata gaggcgttga	13800
ttatgtatgt gatttataaa tattttttta ggatatatgt tgaatttggg aatgaagaaa	13860
tattgtggga attattaagt tgaataaagt attagttttg ttttttagaa gtttattgtt	13920
ttatagagga gattagatat gtttatagat aatttgaagt atatagtaat aaattatagg	13980
atgatttatt ttaagtataa agtgtatatt gttattatag tggtaattta ggttataatt	14040
tttaaagtga agagttttta atattatgtt ttttattttt agtgatatga aattatgtta	14100
gaatatttgt agaaagtttg aattagtaat ttaataata aatattgagt attgatatat	14160
tgtatattat tgagatttat tttatatata ttattttatt aatttttatg attatttcgg	14220
taagtattaa tttttataat taagttagta tagaaaatga ggtttagaga gatatagtat	14280
tttattttaa gttatagaat ataagtagag aatgtagaat ttgtattttg ggttatttga	14340
tagtggaggt tagattttta aaattaggtt tataaagatt ttttgaaaa aaagaaaaat	14400
aaaaaataga gaataaattt ttttgaagaa aagtaggaaa ttattattta taagtataat	14460
taaagagaga tttttaattt tttttttttt tttttttttg agacggaggt tegttttgc	14520
gtttagcgt aattttagtt tattgtaatt tttatttttc ggtttatgt tataaagaaa	14580
gttttattt gattggttcg atttgattgt ttattagaat ataggttttt tagtagggtt	14640
aataagttat aatagtttta gtgatataag ataaataaa attaataata aggttaattt	14700
agaattatta atatatgttt tttattgtga attattagtt ggtttataga tggaaaatat	14760
aagatatata atgtattagt ttaaaaaatg ttttttagag agtgtttaga attatagaat	14820
ttaaaaaatt tgttttgaag aggtttaatt tttattttta aatataaaa gaaattaaat	14880
gtaaaaaaat taagagaaaa ggttttaatt aattaaatgt tgtataattt taagttttat	14940
ttatgtttta gattagtttt gtaatttatg attaatttta gtattagttt gtttttaatt	15000
taggtttgtt aaataatata gaattttttt tttaatgaag ttttttaggg ttttagttga	15060
aaaatatata ttagttagga tgtttttaat tgaaggaat aaaaagtac gattttattg	15120
ggtttaaata ataatgatat ttataatgtt atataaaatg aagtttaaag gtatggttta	15180
ttttaggtt aatttttaag gatttacggt ttgtttttgt tgattttttt ttagttttat	15240
tttttagagt attttattta tttatggatg taaaatgtt ataatagttt tagtagaaag	15300
attttagtat gataatgttt agggaaagaa attgattatt tttttttttg tttttttttt	15360
aggatgaatt tatttttttt agggattttt tattggagga agagattttt ttatatttta	15420
ttggttaaaa ttatattata tattgtttta aatttaatta ttgaaaaagg gaatggtatt	15480
tttataattg atttaggata ttaggattat tttttgggtt ggggtttgat ttaggatttt	15540
ttaaagtatg tgattgaggg taggtattag aatggaatta gagttgtgtt agaaaggaga	15600
aacgtatggt tattatagaa tagtttttaa tgttatcgag gagggataa ttgtaaaaat	15660
taaataattt tttgtatttt ttttgaaggt ggtattttga ttttttagatg gttttttaat	15720
ataggttttt ttttttttta ttattatagt tgttttgaag ttttaggttg aagggaatat	15780
tttgagattt agattgttaa atgttttttt taaagttatg taataaatta aatggtaaag	15840
ttaggtagt tttttgattt agtatagggt attttttttt attttttatt tttgagattt	15900
tagaattgtt ggtattgttt taaaatttat ggtaagaatt ggttattttt gtaattaata	15960

-continued

---

```

ttttttata atatattgt tttgtttgt tagtttagta gaaattatat ggagtttgtg 16020
ttttaaaaag tttgtcgaag tttttatttt ttgttttggg attatgtgta tgaattatta 16080
attgggtttt tttttattta tatttgatga agatgttttt tttttaatat ttttttttat 16140
tgttttttat ttttttttgt tttatttatg attagtgttt tgtttttaaa tagattttgt 16200
ggttatttat ttttttttgt gttagttttt atttattagt tatttgaatt gtgggtttta 16260
ttgtttttta tatagtttat ttatttgtgg ttgttatata tatttatatt gttatatgtt 16320
tttaaattgt attttggaaat gattttggta atgggtgtat tggatgagat ttaaattaat 16380
aattaaagta ttgagatagt ttttgttatt ataagtttat ttttgttttt atagtttaag 16440
aggagtatt ttttttattt ttattattta atgtttaata ttatttttat tataataat 16500
gtataaaaag tatgtgattt atgattttat ttaaatttga atgtttgtga tttattttgt 16560
gttttttttt attttttagg 16579

```

&lt;210&gt; SEQ ID NO 9

&lt;211&gt; LENGTH: 16579

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 9

```

tttagaaatg aagaaaaata taaagtaaat tataagtatt taggtttaaa gtgggttata 60
aattatatgt tttttgtgta ttgtatgtaa taagggtagt attaagtatt aagtaataag 120
agtagaagaa gtagtttttt ttggattata aaaataggag taaatttga atagtagaag 180
ttattttaat gttttaatta ttagtttgaa ttttatttaa tgtagtatt attaaaatta 240
ttttagaata tagtttggga atatataata atgtggatgt gtgtgataat tataagtgag 300
taggttgtat gaagaataat agagattata gtttaagtaa ttagtggata ggagttggt 360
taaaaagaaa tgggtgatta taaagtttat ttagaatag atagttgatt ataaatagag 420
taagagaaga tggggagtaa tagagaaga tgttgaaga aaggtatttt tattaagtgt 480
aaggtagaag ggaattaatt ggtggtttat gtatataata ttaaaataga gaataaggat 540
ttcgttaggt tttttaaagt atagatttta tgtagttttt agttggttaa gtaataaaaa 600
taaatgtatt ataaggaggt gtttaattata aaagtgatta gtttttgta tgaatttta 660
agtagatta atagttttaa agtttaaga gtaaagaatg aaagaaaata ttttgtattg 720
agttaagaaa ttgttttggg tttgttattt aatttattgt atgattttgg aaaagatatt 780
taataatttg ggttttagaa tgtttttttt tagttttaat ttaaggtag ttataatgat 840
gaggaggaga agaatttgtg ttggggaatt atttagaaat taaagtgtta attttaaaag 900
aatggtaaaa gattatttgg tttttatagt tgtgtttttt tggtagtat ttagaattat 960
tttataatag ttatgcgttt ttttttttta atataatttt ggttttattt tagtgtttat 1020
ttttagttat atattttgga gaattttggg ttaaatttta gtttaaggga tgattttgat 1080
attttaagtt aattatagag gtattatttt tttttttaat gattggattt ggaatagtat 1140
ataatatagt tttggttagt gagatgtgag gggatttttt tttttagtga agggtttttg 1200
ggaaagatag atttattttg agaagagaat aaaggaagag gtggttagtt ttttttttg 1260
gatattatta tgttgaat tttttgtgg aattgttga gggattttgt atttatgagt 1320

```

-continued

---

aagtgagata tttttgaaa tgaagttgag aaagaattag taaaggtgag gcgtaagttt	1380
ttaaggatta gatttagagt ggattatggt tttgaatgtt attttatatg agattataaa	1440
tgttattatt gtttaaat	1500
tttaattggt atataatgtt tagttgggat tttgaaaaat tttattggaa aagaaatgtt	1560
atattattta gtaaatttag attgaaaaa agttaatggt gaaattagtt ataaattata	1620
aaattgattt tgaatataaa tagggttgg agttgtatag tatttaatta attaagattt	1680
tttttttaa ttttttgta tttgatgtt ttatatattt tgaataaga attagatttt	1740
tttaaaataa aatttttaa ttttataatt ttagatattt tttgaagagt attttttaag	1800
ttaatatatt atataatgtt tttttttat ttatagatta attgataatt tataatgaaa	1860
aatataggtt gataatgtt aattaatgtt attattagtt tttattttat ttatattatt	1920
agaattggtt tgattgtt attttttg aaagtttatg ttttagtaga taattaggtc	1980
ggattaatta aatagaaatt tttttatgg tatgaattcg ggaggtggag gttgtagtga	2040
gttgagattg cgtttggcg atagagcgag atttcgtttt aaaaaaaaa aaaaaaaga	2100
aattaaagat tttttttaa ttatatattt aaataataat tttttatttt tttttaggaa	2160
aattgtttt ttattttta ttttttttt tttttaggaa atttttataa gtttaatttt	2220
aaaagtttg gttttattg taaatagttt agaatgtaag tttgtattt tttatttata	2280
ttttgtgatt ttgggtaaaa tattatattt ttttaattt tttttttgt tattatttag	2340
ttgtgaaat tagtatttt cgaagtgatt ataggaattg ataaataat gtatataaaa	2400
taaattttag tgatataaa tatattagta ttaaatattt attattaaa ttattaattt	2460
agattttttg tagatgtttt aatatagttt tatattatta agagtaaaaa gaatagttt	2520
tggtgtttt atattgaaa attataattt aggttattat tatagtagta atatatattt	2580
tatgtttgga ataaattatt ttataattta ttgttatgta ttttaatta tttgaaata	2640
tgtttaattt tttttgtag atagtaggtt tttggaaggt agagtagta tttattttaa	2700
tttaataatt tttataat ttttttttg ttaggtttag ggtatatttt tgaggaatgt	2760
ttgtagatta aatatatgat taacgttttt attttttatg tagaaatatt tttgaaattt	2820
cgtgttaata attttttaat atttgaaata tttttgaata tttttgaatt agtttatgtg	2880
aagatttatt atgatttaat aatagatgt ttgttaatat ttagagagtt aaattttatt	2940
taaaaaatt aaaggtgtga tttaggaaag gattttttgt atttgagttt tttatttaat	3000
ttaaataatta aaaatcgaat tttttgtgtt atatttaatt tttataaaag tgtttttgt	3060
tggttgagagga agaattgttt tttatttata taggtataat aagtatacgt attatttttt	3120
ttaagattat ttttaagatt aggtaagaat ataaggagga gattagatat gttttaaata	3180
tatttgaata agttgtataa atatgtttga ataagttatt taatttaaaa attagattgt	3240
ttatattaaa tataaaatgtt ttaatttttt ttaaaaaatt agaaggttta gtaatattaa	3300
gtttatatt tagttgggtg attatgagta ttaattaagt agtagttttt ttttttaagt	3360
ttgtgtgttt aaggttatta ttgtttttat ttagtttttt atttattttt attttatgtt	3420
tcgtttttgt agttatttaa gtttgtaaat tttgaattaa aagatgatta tttattttta	3480
tagtaaatat attagaggtt tttattgatt ttagttgttt tagaattgta tttttatttt	3540
ttttttttg gagtttaaga cgtttatttt atttttgtaa ttttttaag gttatgtttt	3600

-continued

---

tatttttggg	tattaagtta	tattttttgt	tttttatggt	tgattatatt	aatggaaag	3660
tttttagaag	tttttttttt	tgtttttagt	attattttaa	atttttat	ttattttttt	3720
atttttttat	gtttttttta	taagtatcgt	ttattagtga	taataat	gtttgtatgg	3780
aaagatttga	gagatatgag	agtgaatggt	tatatttagt	ttttat	agtgaagtta	3840
atgagaaggg	aagaaagtta	ttgaataat	taggtatagg	gtattgagat	tgtttaattg	3900
tttgggagat	ataaaagata	aggttaatgg	aatttttatt	tatgagaaat	gtttgtaggg	3960
aggataaggg	atatattagg	gatgggtagt	tttcgagaga	aagttaagat	ttattaaat	4020
agatatttag	tatttttttt	tgattttggg	gaagtttagt	tgataagggg	tgatttaagg	4080
atagagatac	ggtggttttg	tttattat	tgtaattag	gtattttaga	taattaat	4140
tttaggatat	tttagttata	aagtggatgt	aatgagtatt	atatattata	ttgtttattg	4200
gttttttttt	atagaagaa	tattgtata	gtagagagaa	attacgtgat	ttaggtgcca	4260
aattttatga	atttatttta	tataagtaat	agtattatgt	gaat	at	4320
ggtaagttt	ttttatattg	tcgtttttta	gtattttttt	gtttttgtt	tggtgaat	4380
tttttttatt	atttatttta	agatatattg	tattttat	ttttttat	gtttat	4440
tatttttagta	tatattatta	ttttgtat	gtgattggga	attgtttttt	ttattttatt	4500
atagtaaaag	ttttaataga	gtagtttttg	atgtttgtt	tatttttga	ttataatgt	4560
ttagaatatt	tagtaaatat	ttgttaaag	aatgaatata	tattttttt	atataattag	4620
atgtataata	tatatttagt	tttatttatt	aat	tttagaatat	aatatatttt	4680
ggagaatggt	ttaaatttagt	aaatataat	tttttatggt	at	gtt	4740
ttttatttag	gtggatatag	taat	ttaaatggtt	ttttatgatg	agttgttat	4800
attgttttta	gatatttgtt	tttagaagta	atgttgtaaa	gaat	gtttatttgt	4860
tttgggtgat	attggtatta	atatttattt	atataaat	gtc	ggaattggtg	4920
aattaaagag	tatatatatt	tttaaaaaat	ttttgataaa	gaattaaaaa	aaagtgattg	4980
ttaaattatt	ttttgaagta	gaggaattta	tttattttta	tataatagta	tatgagagat	5040
ttagtttttt	tataatttgg	ttaatatata	tttgatttaa	at	ttttttttg	5100
ttaggatag	tggtttatgt	ttgtaatttt	agtattttgg	gaggttgagg	tggtgaatt	5160
at	gtgaggtt	aggagtttaa	gattaattgt	tatggtta	atggtgaaat	5220
t	aaaaat	acg	aaaattagtt	aggtatgggtg	gtgggtgttt	5280
gtt	at	gtat	ttgaatttgg	aaggcggagg	ttgtagt	5340
g	tt	at	ttt	ta	tt	5400
tt	at	tt	tt	tt	tt	5460
tt	at	tt	tt	tt	tt	5520
tt	at	tt	tt	tt	tt	5580
at	tt	tt	tt	tt	tt	5640
tt	at	tt	tt	tt	tt	5700
ag	at	tt	tt	tt	tt	5760
tt	at	tt	tt	tt	tt	5820
at	tt	tt	tt	tt	tt	5880

-continued

---

agattgattt ttattttga tggatttatt aattgttttg aaattattta ttgaataatt	5940
tttttttttt attgaatggt taattttatt ggagttaatt tttgtaattt gttttttgat	6000
taatgaaatt aataatsttt taaaataata tgtttttttt tgatttattt tatagattta	6060
gtagaaatta tatggtttgg gtaatgtagt tattattaaa attatagaat tagattaata	6120
tgaattttta gatataatat attattattt tttttttgta tagtttttta aagtatttat	6180
ttttttgatt tttataattg tatatttatt taatagatta tgatattaaa ggtttaaaag	6240
taatattttt gtttggtttt tttagttagt gagtttaatt taaatttttt gattttaagt	6300
ttagtgttta ttttattata ttttaagtaa tattaaatta ataatttaaat aataaaaaga	6360
ttataattta ttaaatatgt ttattgtaaa atatgtaatt ttattttata gatataaata	6420
ttttttaaaa tttttagaag aatgaggtag gatgtatata gatgtggaaa gtatatatat	6480
ttgtattttt ttatagttgg taaatttatt tagaggtgtg ttttaggaaa tagaattgat	6540
atattaaata aagatttttt gtattgattt ttgaaatata attattttta atatattta	6600
tggttaagaat tttagtagtg ttaaaaattt tgtatatata tatatatata cgtattttat	6660
atatatatat agattttata tatatgtaat ttattaagtt taggttataa gaggaaattg	6720
gtaaaaattt attaatgaa atggagtttg gtttaagatat ttatattta ttttttttt	6780
ttgattttta attagggaga gtatatgttt tgtgtttttt ggtaaaatat agatttatgt	6840
aataattagt atatttttt agattagagt aagtaatttt taataatgag tgtttgtag	6900
ttttaattgg tttttttacg tgggtggtga aaaaatattt tattatagtt attaattttt	6960
ttttttaaaa gaattatgtg ttttgagaag ttttgtttta ttatagggta atattttatt	7020
ttaaatatag tgtttatcga tgtttgttta tgaatgacgt taataaaaat gtttaaaatt	7080
tattaggaaa ggaggggtta gttgttagat tttgggagat agagttagtt gatttttgcg	7140
tttttaaagg ttttgttatt taattaagtg attaattatt aattttaaag tatttaagtt	7200
taataatata ataaaagaga gaataatggg tagtattttt aatattaatt atggttattt	7260
ttggtttagt tgtttaggtt tttttttagt agaatttata agtgtataga tataggattt	7320
ttggaatatt aaaggtattt ttttttagtt tgtaaatata gggaaaaaga agaaatttat	7380
atttgtaaat ttatgtaaat tttaggggat tttttttaaa tttcgttttt ttgaaagtag	7440
gaaatataat ttaaaattaa ttttgttttt atttattaaa agaattaaa tttgaaatt	7500
tgatataaat tttaaatgat tgttttgaat tttatttcgt tttttattta atttttttat	7560
aagaagatta aggattgtgg atttggtggt ggaggaagta attgtttaat tatattggaa	7620
tttattggaa ttttttataa tattattgcg ggtagataat ataattttta tatatttttt	7680
ttttatttta gagtttagaa taagaagtta gatttttggg gatttttttt ttgatgtatt	7740
gtttttaata tttttaagtt ataaattggg aatcgttttt attttagaat attaattggt	7800
gtttttattg gtattttttt atttttatat tatatttata atattttatt ttaggatttt	7860
atgtgaatat ttttaggatg ggtttggcga aattgttaag aagtgtttta aataattagg	7920
ttatagtttt tagtagataa aaagaaaaag acgtggttgt tagttataga atttgaggtg	7980
atgaaaaaga tttttaaaa ttagtttttt taatatgtag taaaatggga ggatttaaaa	8040
tgtagtttta agaagtatcg tttcgtttgt attaagtttt tttaggtttt tagtttagtt	8100
attttttaaa taggtgttga gtttttagtt ttcgttaggt tttcgttaag gttttttatt	8160

-continued

---

gaggtatagg aatTTtagtt atattcgtag tttttttatg tttgttttta tgaatttta	8220
ttagttttgt cgaatggtag tatttggtat ttttagcgaa tttttatga tagtttgaa	8280
cgaattaat ttgttaattg attaagaatg ggttttagag tgatttgaat gtatttaaaa	8340
agttaataaa tggaaacgg tataaagttt ttttttttg ataaaaagt tgtttggagt	8400
aatTTtaat ttttttttg ttttaaaata aaagtacgat gttttttta ttattttta	8460
gaagtaagta aaaataaaat aaaataaaat aaaatttaa tagttttta ataagagtgt	8520
ttagcgttgt ttggatggaa ttcgggttag ttttagaagt agtcggtttc gaatttcggg	8580
ggcgaggta ttcggttttt tcggttaagt tgggttttgt ttttttcggt tcgtttttat	8640
tgttttcggt ggggtggggg ttcggttgcg tcggtcgttt gtgcggttgt attgtagttt	8700
ttgagtttag gaagtatat tcgtttcggg tgggagtcga agcgttaatt tacgtggcgg	8760
ttggttaatt tagtcgttcg ggggagcgtt cgttgccggg ttttttttg cgcgaggtga	8820
gcgagtgggg gcgggtttta gatttagttt ttttggcggg ggttgatta ttttcgttg	8880
tttagttttt ttttattttt ttatttcggg acgtttttgt agttatattt gttgctggtg	8940
ggaaaggaa aagagtattt taaaacgagt ttcgaaattt agatttaaat ttcgaaacga	9000
atTTtttcgg ttcggttttg gttgttgaa tgttatttgt tttattttcg ttttttttt	9060
ttttttttt tttttacggg tttttttttt tttatttttt tttttttttt ttttttttt	9120
ttttattgt tcgttttttt tttatttttt cgttcgacgt tcgttttgaa gttgaggaag	9180
tttaagtttt ggattaggag tttttttttt gtaatatattt ggggcggggg tttatcgttg	9240
ggtttaagtg gttatgggtt ttgagagtac ggtatgtgta ttattttttt taaagacgtt	9300
tagggtggcg cgtttagatt ttcgaaattt ggttttattt cggttatgtg cgttcgttta	9360
ttaagcggcg aattggattt agggcgtcgc gtgcgtttta gtcgattcgc gagaaggat	9420
ttttgtattt tgtatagatt aagatttttt tttagggttt tttaaatttt tagatgtacg	9480
tcgtagattg tggagtgtaa aatgaatgat ttaaggaaa attagggttt tttatttttg	9540
aatataaggt tgtgaaagat gtacgtttg tttgggggtt ttgtttttt tcgtttggtc	9600
gggatttttt tattaggggg cgtttatatt cgtcgtaaaa ttttaattt ttagatattt	9660
aggttttgta agttttgtga gttgaattgg gggtagattc gaaggagcgt atcgtttttt	9720
tcgttttttt tttttgttg ttttgagta cgggaaaggg agggatttta tattagttat	9780
ttatttgta gttgtggggg aggggacgta gaagttagag agagagaata gaggttggga	9840
gaggggaagg gaggtagaga agatggagat agaggagat ataaagatga ggaacgcgat	9900
ttagagagat tttgtcgcgg gtgtgcgcgc ggtaggataa ttttttaggc gtcgtgtaga	9960
ttttttttt ttttcgtagt attttttagg gatatgtagg atttgtcgag acgtagttgg	10020
attggagagt aaatagtta ggttatgata ttattttttt atttcgtttt tttttgtta	10080
atatttagat cgttttacgt tttagtatat agatcgcggg ttagaataat tcggatggtt	10140
taggtttgta cgtattaggg atgtcgtttt atagcgttta cgtttttgta ggacggtggg	10200
gcgttttagt cgaggacgtt ttcgggggtt ttcggtatac gcgcgcgtcg ttgtcgttag	10260
atTTgttagt aaagtgtttt aattagagcg gagaattagg agagggtggg gaatatgtaa	10320
ggaaacggag gatttttgta gtttggcga gaggattgat ttagagggtt ttgattcggg	10380
ttttcgagtt atTTttgttt gtagtgcgtt ttgggtgtcg gttgtttgt tgagatggga	10440

-continued

---

gcgatagatt	ttatatttatt	ttttagggtg	agaaaaattg	ttgtatattt	ttgtcgattt	10500
agaatagggt	tttttcggtt	aggggttttt	tagattttcg	cggggtttta	ataatattat	10560
ttttattttt	ttgttttttt	ttgtttggtt	ttttcgtatc	gcggttagga	taaaagaatt	10620
taggggctc	gggagagaat	cgtgatagtt	tttaagtcgt	taaagagata	cggcgggat	10680
tttgattttt	tttcggtcgg	attttttttt	agattgtttg	cggtttgaag	gggttaggtt	10740
cgcgtaaggt	tttaggttgt	ttgggcgtcg	ggtcgtattg	ttaggagttg	ttcgggagtg	10800
gttttttagg	aatattaggt	attgattata	tattaggggt	tgggaaatta	ggtggtttgt	10860
attaagcgg	tttcgggat	ttgtttgtgg	taagttttgg	tagtttttat	ttaaattttt	10920
gtttcgagcg	tgtaagaat	aataataata	aaaaaattaa	agtgttaaag	gttttttttt	10980
tttttttagt	ttaagaattt	attatatttt	tatgattttt	ttataattta	tttttttttt	11040
tttttaatt	tcgttagtta	ttttattttt	attttatttt	gggttttttt	tgtttgaatt	11100
tttttaata	ttaaggtttt	ttgtatggtt	tttttttaaa	agtttttatg	aaagttattt	11160
gtatttttta	agtgtttata	tttttttaat	ttcgttttat	agttttttgt	ttaatttaa	11220
gttttttatt	aattgttttt	tttttttaag	ttcgcggggt	tttttttata	agttttttgt	11280
ttttgttttt	taagggggga	ataaaagaaa	cgtgattatt	ttggaagcgc	gtttattgta	11340
gtttgggggg	aaaattttat	gtagcgttgc	gcgattgggt	tcggcgttgt	ttaggcgggt	11400
tatataggaa	gcgtggtggt	tcggggaagg	atgcggaggg	tcggggacgg	tgttgaaga	11460
tcggggagga	tcggggggtt	gtcgaagatt	ttggtttttag	tttttgttta	taatatttaa	11520
tgttttcgg	tgttgcgttc	gtacgttttg	agtttttttt	tagaaagggt	cgggtagtt	11580
cgtttgtaag	tttaaatgcg	ggttggtgata	tttattatta	ttatattatt	gtatcgttag	11640
agtcgaggag	gagatttagc	gagaagaagg	aggagggaga	ggaggagggt	ttatttatag	11700
gttttaaaag	cgtttcgtag	ttttagttat	ttttaagagt	ttaggttttg	aaagtagcgc	11760
gagggcgga	aaggtagttt	ttcgcgtttg	cggtagggat	cgtttgtttt	tttttttagg	11820
tatttttatt	ttttggcgtt	tttttttgat	aagaagtatt	aatcggtttg	gggatagtag	11880
cgttttttat	ttagggattt	atttagtaat	atcgttttgt	tcgcggaaat	tattgttttt	11940
ttattttttt	tttattttta	ttttttcgtt	tttcgtgtag	ttagtttttg	gtaggggaa	12000
aggagttttt	aggtttttag	gggttaggtt	agtaatagat	aattgagtac	gattattttt	12060
tttcgggagt	atataaaatt	gtaaaattag	taaagaattt	ggttatagcg	tgtttacgcg	12120
tttatagagt	ttgtttggtt	tattaaagg	aagtgttagg	tttaaggttt	ttgttaattt	12180
gaaagagata	ttgagaaaac	gagatttttc	ggggattttg	agggaaagtg	taagaatttt	12240
ttattgtatt	tttagggaat	tgtttaaatg	ggagttcgg	tttaaaagat	tttgtaata	12300
aaaggttga	taggaaattt	ttttaggtaa	attttttgc	ggatttaaag	agaatatttt	12360
ttttttgtta	taattttttt	tttatataag	tttagatttt	gtttttttcg	tttttttttt	12420
tttttagttt	ttttaagtat	ttttgagtag	aatatttgat	aatttttttg	agtaatagg	12480
attttttga	agtatttaatt	attttttatg	tttttcggaa	ataagattat	aattttttat	12540
gcgtatatgc	gatttttttt	ttttagttag	gtttatttta	ttgtgtaaat	agtattaata	12600
tatggaagag	ttttttgtat	tggtttataa	aagattttta	ataggatttt	atagagaaaa	12660
gggttaaata	gttgatataa	aggatttttg	tttttttaga	aaagagggat	tttgattttt	12720

-continued

---

ttttttttt gaagttaagt atgagtttat ataataggaa taaaataaat ttaaggtgta	12780
tattagtata atattagga tattagaatg gatggtaaat ttttttttt atataaatat	12840
gaaagtattt ttataatttt tttttttgaa gtttttattt agaaaattat attaaatata	12900
ggattttaaa atagtagtga ttaaagatga aagttaattt ttttttttaa tttttttgat	12960
tagtttggtt taaatttatg ttttaaaatt ttttagtaatt tagatttttg tatatgcgta	13020
tttataagat tttttttata ttttgtaatt tgtaggtatt tagtatgggt attgatttta	13080
agtataaat aggtagaaga tttgttagga taattttagt ttttattatg tgtatttttt	13140
attttttttg ttatttatta gttttttgta gtaataaat tgttagttat attatttgga	13200
aattaatggt tttttgtagt gaattattta ttttttagtaa atttattgta ttattatttt	13260
aaatattgaa atttgatttt attagatggt tagaatgata gttatagagt agaattagat	13320
tttaattttt atgtataata tgtaataaaa ttaaggtgta gaaagttatt gttttaaagt	13380
tatataatta atagtttagt ggttttttta gtgtaatta tagtaatttt agatttaaga	13440
ttgtttgatt taggaatgga gagaataaaa ataaaatgaa tgtttattga aatataat	13500
ttagatatta agatagattt ttttgtgaaa ttttttaaga tgggatttg tttttttatt	13560
tttaattttg agtaaaatta gttgagtttt aaagagtta agtaattgt ttaaggttat	13620
tttgtaagtg atagattttg tagtttgaaa agtttattat tttttatttt tttgtttttt	13680
tatgatatag ttaagtaaaa ttttttaata aaggagaatt agattagaaa taagtaatta	13740
atagtagatt aatgtttatt tttagggaaa tgttaatagt taatattat tataaataat	13800
gttatttga tataatggaa tatgattgaa agtgatgggt tagatatttt ttatattttg	13860
agatgttttt atttatggtt ttaaatttga ttttttaatt ttttaaaatt ttttagtttt	13920
tttaattgta atagatttta taatatttga ttaaatattt ttataggata aaggattttt	13980
tttagggttg atttaaggaa ttagaatag tataatggtg aaaagtttta aattgttgt	14040
tttttttttt ttttgatttg taatattatt atttagaaaa ttaaatattt ttttaggta	14100
ttagataatt aaaatgaaat tatattttat ggatattttt tttttatttt tagggtagat	14160
tttttttttt agttgtttta aatggagaat ggataaaaa gaaaaaggaa aatatttag	14220
ttatatttag aatagagttt atagttgtat tattattttt atgtaattgt ttttaatgat	14280
ttttttttat taagagaaag aaaatgagag agtatttatt tttgattttt taatgtttga	14340
taaaattatt agaagttttt gaatgggaaa atatatattt tttaggtttt ataataaaa	14400
ttgtatttat ttgtttatta tatttaagta tgaattaaat agtaataaaa atattatttt	14460
gtgagttttt ttttaaatga tatttttga ttgggatat tttttgtgaa aattttaaga	14520
ttttaggtat aaataaagta agatttttat ttataatggt tgtgtttaaa agaaattata	14580
agttagaagt tatgtgtttt gttttttgt tttgattttg tttttgggaa tttggttgag	14640
tttaggggta gtaagttaat gatttgaga ggaaaggtgt agatggtaga tgtgttttta	14700
gggaatttta ttatttttat ttaaattttt ttttaggtt ttttaggatt atattttaga	14760
tttttttttt tttattgatt aggtgaagtg tttttagata tggattattt tattgttgtt	14820
ttaggatga taagatttta gttttattta aatttagaaa gagtttagtt ttaatttttt	14880
gatttgattg gttttaggtta taattaagta aattattttt attggatgat atagtaatgt	14940
tatagattaa aaatgtattt ttaattttta aagatttttg aaaaaggaag tttatttttt	15000

-continued

---

```

aagattgtgt tataatTTTT ttttaaaata aaatattgta aattagggat agatatttta 15060
tttagtattt tgaatTTTT ttggtTTTT aggaaaggta gaaatataag gtttttattg 15120
tttttttttt atttttaaaa aagaaatTTT gtaaaaggta aagtgatggt ggtaatttta 15180
ttgtgtttta attcgaagta tataaattgt ttgataaat taaatataga taaatTTTT 15240
tagaaagtta taaaatatta gatttttttt agatattaga tattgagttt ttgtttttta 15300
atTTTaaat taaaatgTTa tagtaaaata tttgtTTTaa taaatTTTT tttttttgt 15360
gttgtaaatg ttttgtttat ggttgagtag atttttttag aatttaattg aaagattatt 15420
tatattacga tagaattatg taagttaaag ttaagtttta atatggagtg atgggagagg 15480
tcgggtgtgc gagaggagat gattttatta gtgagaagag aggttttttg agtcggtttt 15540
taaggtagtt aaatgTTta gtatatttgt atttattttg ggaggttggt tggagtagag 15600
tattaaaaac gttattTTTT ttataaaaga aaagatgaag gtataaaaga tatagtTaa 15660
ttaaattgaa ttaaagattg tattattttt attttttttt tagttatttt tttaaagttt 15720
tttttttttt tttttatagt ttgaaaaata taggttattt ttataaaaga tttgtttgat 15780
tgaagtaaga taaattgaag agtttggttt ttgaatataa ggaagcggta gagattttag 15840
aattaatgag gtaggcggga atataatgta gaaggTTTT taagaggtaa aatttttgta 15900
gttggtggtt attttttagt attattcgtt tttatttatt taattagatt atagtgggat 15960
cgagataagg agtataatga aagttgtgtt attgatggtg taaaaagata tatggtatta 16020
ttaatgTTTT tgtggtgagt aaaggtagaa cgtaagatag aaaataagga aatttttgcg 16080
ttgtaagaaa agtgtttagt atgggtatta atcggtatgt agaagtagta agtatatatt 16140
ttatttttat tttattgata aaaagaattg taggattaat gtttgttttt attaaaaat 16200
gagaaattat attttaattt tataaattaa aggttaaata tttgttgtaa taattttaaa 16260
aaaaagagaa gatttgtaag atttagtaaa tatttattat aaatagattg tttttaattt 16320
tttatatttt ttttaattt aataaaagtt tattgagtag ttttatgtg tattgtttta 16380
gatagtaggt attatttaaa aaagtaataa ttatttttta ttttaagga gtttataatg 16440
tttttagaaa gaaattaatt tattggagta gttaaagtta taattgattt ataaaatgat 16500
tagtgaattt agaggaaaat tatatataaa ttttttggtt tttattgggt ggtttagtta 16560
gatattattt taaaaaata 16579

```

&lt;210&gt; SEQ ID NO 10

&lt;211&gt; LENGTH: 3000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 10

```

atatttaaT tgtatattTt gtttgaata tattatataT ttaaagaat tattattttt 60
ttgaatatta ttaattattaa taatagtatt atagtgatta agagtggatt taaataaagg 120
ttggtttgaa atttaggttt gttatttatt agtaagtttt taaaaatttt gagtttttag 180
tttttttata tgtgaaatgg agaaaaataa tttgtaaaat taattaaatt tgtttattta 240
tatttttttt tttagtgtat attaaattgg atttttcggt aggttagaat gtgtttttta 300
ttatgtttta aattcgtttt gtgttaattt tattgttaga atttttttta ttttgagaat 360

```

-continued

---

tagaaaagga aatattatgt ttgtaatgt ttatatTTTT taaaataaat ttgtaggaaa	420
gaatatttag taagtgatga gagtagaac gattgTTTT attatttaa attataata	480
ttattTTTT tagagTTTT taagtattgt agataatTTT ttatttatta aaaaataaat	540
tgtaattata agtatttagg gttgatattg tttttgaatt agatagtgtt tatattagtt	600
gtataagatt aatttaaagt agaggatgaa atttttTTTT tgaattTTTT ttagaacgta	660
atttagtgaa tatattaaaa ttaaattTTTT tttgaatggg agtaattTTTT acggattaat	720
ttgtaatttt ttagattata ttttaaggtaa thtagagggt gttgtattat aggttttTg	780
attagagatt attggatttg tttttggaaa agttatttat gattttttgg gttttggttt	840
ttttattTTTT tatatcggtt taataatgat tatcgtagag ttgttatgga agttaaata	900
atttatgtat tataagtgat taatataatg attgatattt agtgatttat taataaatta	960
aaatatttat tattaatatg atagagaagg tgcgttaaa atagataata ggtttttgga	1020
agaggtgatt aaatggatgt aaaatttatg gattgtttat ttcgtttatt tttgttgtgt	1080
tttttggttg tggatatat acgtgtgggt ataaaatcgt aaattttatg tagtcgcgta	1140
gtgtatcgt agaaggttta gatacgaat gttatttttag taatgtgttt agagaagttt	1200
tgacgtcgtt ttggaagtaa gtcgttggtt tttgattttt gggcgtttgg gacggatgtt	1260
tatatTTgta tttagtagta ttggaagggg tttaggtttt tcgtagtata gtttattttt	1320
agatcgttta gttttttata atatataTTT ttacggaaaa gggatttttt ttcgttaga	1380
aaaagcgttt tagtttggtt tgggttggtt tttattttac gttgttgtaa gtaggcgaag	1440
tttttttTg tttttttttt ggggtaagtg gaaaggagt cggtaggggg ttcgtagtgg	1500
tttTatagg ggaattgggt agcgagagag ttttaggtaa tttcgggggt tgttttatag	1560
aagtaggtgg ggatcgatag tggtttttcg gtttagggag gagagcggg tcgcggttt	1620
tttttttag tttggaggtt gtagtcgttc gagtcggttc gggtaggggc ggggtggggg	1680
cggcgcggag ggtacggaga ttacggcggc gttattcggg atatttaggg ttcgaggtt	1740
ttgggcggtt tttacgcgag atcgtaaatt atgataatag gtagttattc gaggttaaat	1800
aaaaacggag tgggtttttc gcgcgtcgtc gtttttcgcg tttttggcgg tttttttcga	1860
gggtttcggc ggttttacga gttcgtagta gtcggtggcg acgtcgtttt cgttttattt	1920
ttttgcgtaa gtgcgaggtt gtcggtagcg cggcgtacgt ttcggtcgtt ttcggttttc	1980
gcgtaaaatt tttattttgt ttacgtgaag ttgctcgtgt tttagagagg gggaaagagt	2040
tgcgggaaaa gtcggggagt gacgattgcg gcggttgggc gcgttttttt attttttttt	2100
ttttttttt tttttttgtc gtagttcggga gttttggttt tttttttttt tttttttttt	2160
tcggagtcgg tttttttttt cgtttcgttt ttttttcggt tgtgtacgtt attgtttgtg	2220
gggtggtcga aggggatgtt ttgtttttat tagaggtata gcgcgaaggg gaaatttcga	2280
tattggaagg aacgagaata aatatttaat tacggacgta ttgaatcgcg gttgggatag	2340
atatttcggg aattcgaggc ggcgcgggc acgaggtgag tgattttttt ttttaatttt	2400
cgttttaggg ttttcggggg agtttgagtt gagagaattt ttaaattttt cgggaaagtg	2460
cgcgaggttt cgtcggggac gtcgagcgtt gggatttgag gacgcgtagt tggacggtgc	2520
gtgggcgttt gcgttttcgg ggggcgtttg gaggtcgggt gttttacgtt tgagggttcg	2580
ggtcgttcgg atcgtagcgg tgtttttgtt tttagaagac gtttttaagt ttaaggggt	2640

-continued

---

```

tttttcgagt ttgtttgttt ttttcggggt cggcgcggag tttgcgcgta acggagttta 2700
tttagtagtt tagcgcgcgg tttttatttg tatttcgttt ttatttgga gaggcgcgag 2760
tategggggt tttttatata ttttttatg acgtgtatta tttttgatg attttttaga 2820
tggttttagc gcgaggatgt tgatttagag ttttcggag ggttatagc gtttggttt 2880
tttcggtgtc gggcgcgtgt gtattttaa ggttcgcgtt ttaatttta ggtattgatc 2940
gggttttta attgcggca ttttatttta atagttttta tgtggcgtgg attgaatgtt 3000

```

```

<210> SEQ ID NO 11
<211> LENGTH: 3000
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

<400> SEQUENCE: 11

```

```

gatatttagt ttacgttata taaaattat taaagtggga tcgctcgtagt tgaaggttc 60
gattagtgtt tggagattag aacgcgagtt tttaaagtat atacgtattc ggtatcggga 120
aagtttaggc gtttgtgatt tttcgaagga ttttgggta gtattttcgc gtttgatta 180
tttagggggt tattagaaag taatatacgt tataagaaag atgtggggga gatttcgatg 240
ttcgcgtttt tgttaggtgg aggcgggggt taggtagaag tcgcgcgttg gattgttga 300
tgaattcgt tacgcgtagg tttcgcgtcg atttcggaag ggataggtag gttcgaagg 360
gatttttggg gtttggggac gtttttagg gtagagagta tcgttcggt tcgagcgtt 420
cgggttttta ggcgtggggt attcgtttt taagcgtttt tcggggacgt agcgtttac 480
gtatcgttta gttgcgcgtt tttagtatt agcgttcgac gttttcggcg gagtttcgag 540
tattttttcg gaaagtttg gggtttttt aatttaggtt ttttcgggag ttttggggcg 600
ggggttgaa gaaggggta tttatttcgt cgttcggttc gtttcgggt ttcgaagtgt 660
ttgttttagt cgcggttag tgcgttcgta attaagtatt tattttcgtt ttttttagtg 720
tcgaagtttt ttttcgcgt tgtgtttttg gtgaaaatag gatatttttt tcggttattt 780
tataataaat agcgtatata agcgggggag aagcggggcg gagggagaag tcggtttcga 840
gggggaggag gaaaggagag gagttaaatt ttcggttgc gatagggggg aaaggagaag 900
aaaagaaat gagagagcgc gtttagtcgt cgtagtcgtt atttttcgtt tttttcgtta 960
gttttttttt ttttttaag gtagcgataa ttttactggt ataggatgga agttttgcgc 1020
ggaagtcggg aacggtcggg gcgtgcgtcg cgttgcgtt agtttcgtat ttgcgtaggg 1080
aggtggggcg ggggcgacgt cgttatcgtt tattgcgggt tcgtgaggtc gtcgggggtt 1140
tcgggggagg tcgttaggga cgcggggggc gcgggcgcgc gggggattta tttcgttttt 1200
atttgatttc ggggtattgt ttattgttat ggtttcggat ttcgcgtggg gatcgttttag 1260
ggtttcgggg ttttgatgt ttcgggtgac gtcgtcgtaa ttttcgtgtt tttcgcgtcg 1320
ttttatttc gttttattc gggtcgattc gagcgttgtt agtttttagg ttgaggggag 1380
ggattcgcga tcgcgttttt tttttgggt cggagagtta ttgcgattt ttatttgttt 1440
ttgtgggta gtttcggaa ttgtttggaa tttttcgtt atttagtttt tttgttagg 1500
ttattcggg tttttgtcg gatttttttt tattttttt aagggaggag atagaaggga 1560
tttcgtttat ttgtaataac gtgaaataaa aattaattta gattagattg gggcgttttt 1620

```

-continued

---

```

tttgacggga gaaatattt ttttcgtgg agatatatgt tatgaaggat taagcggttt 1680
gggataggt tgtgttgcga agggtttggg ttttttttag tgtgttggg tgtaggata 1740
ggatttcgtt ttgacggtt aaaggttagg tagtaacgat ttatttttaa ggcggcggtta 1800
gagttttttt aggtatattg ttgaaatgat atttcgtggt taagtttttt gcgtagtat 1860
tacgcgatta tataggattt acgattttat atttatacgt gtgtatgtta taattagggg 1920
atatagtaaa ggtagacgga ataaataatt tataaatttt gtattttttt aattattttt 1980
tttaaaaatt tattgtttat tttagcggta ttttttttgt tatattgata ataaatggtt 2040
taatttattg ataaattatt gagtgtaaat tattgtattg gttatttatg gtatatgagt 2100
ttatttaatt tttatgataa ttttacgggt gttattatta agtcgggata aaagatgagg 2160
aaattaaagt ttaaagaatt atgagtgatt tttttaaaga taaatttagt ggtttttaat 2220
tataaagttt atgatataat aatttttata ttattttagg tgtggtttaa gagattatag 2280
attaattcgt agaaattatt tttattttaa gaaagtttag ttttaataata tttattaagt 2340
tacgttttga aaaaggttta gaaaaaagat tttatttttt attttagggt ggttttatgt 2400
aattgatatg agtattgttt aatttaaaag tagtattaat tttgaatatt tatggttata 2460
atttattttt taatgagtgg ggaattattt ataatgttta agaggtttta gaaaagggtgg 2520
tgttgtaaat ttaaaatgat aaaggtagtc gttgttgttt ttattattta ttgggtggtt 2580
ttttttgtag atttattttg gaggggtgtag gtattgttag gtataatggt ttttttttg 2640
gtttttaagg tagaaagggt tttggtagtg gggttggtat aaagcggatt tggagtatgg 2700
ttgagagtat attttgggtt aacgaggaat gttagttaat atatattgga gagaaaaata 2760
tgaatggata gatttaatta atttttaga tttatttttt ttattttata tgtgagaaaa 2820
ttaagggttt agaattttta gaagtttgtt aataaatggt agatttgagt ttttaagttag 2880
ttttatttta agtttatttt tagttattgt gatgttatta ttagtattaa tggattttag 2940
aaaaatagta atttttttaa atatgtaata tattataagt taaatagta atttaaatgt 3000

```

```

<210> SEQ ID NO 12
<211> LENGTH: 1984
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (9, 17, 18, 33, 41, 49, 50, 52, 54, 62, 80, 84, 127,
157, 159)
<223> OTHER INFORMATION: unknown base
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (175, 207..209, 214, 1181)
<223> OTHER INFORMATION: unknown base

<400> SEQUENCE: 12

```

```

agttttagnt attattnnta tttttattaa tgnntttggg nttataggnn gngntttttt 60
gnntaaatag tagttaatgn aggnatttta aaaggatat tattttatat ttatttagtg 120
gttatgntat tatttaattt ttattataat tttgtgngnt tagtataatt tgtgntttta 180
ttttatttgt gaataaatgg aggtagnnnt ttgngtaatt tataagtaag tggtagattc 240
gggggttggtg tttagataat gtggtttttag attttttggt tttttttggt tttgttgta 300
ttttgattga gtttagttat ttattttttt aatatatatt aattggtggt tttatgtttt 360

```

-continued

---

```

aggttttgtg ttagatgta gaaatatatt tttatttatt ggattttggt atggtttatt 420
gtttagtgtt ttgttatgtt acgtgttata taatgtattt atttcggat ttagtttatg 480
gaaaataatg ttgaaagata ttgtatgtat gtttttataa taaaattatg taatttatgt 540
tttttttatt gttttggta ttaggaaggt tttatttaa ttagagtaa tatattagaa 600
tttgggttg tttatttagt ttgttgaatt ttttttttg gtttagattt tttattttgg 660
tttataaatt tttattgata aatgtttttt attgtaaaat attttaaatt ttttttaagg 720
gaaggttga tggaaatgat atagtaaggt tttttttgta tttttttaga tttttattag 780
ggaaggtaga ttagatgtt ttttttattt ttagtttaa agtttttat ttataaaatt 840
ttttaagtag tgattattgt tgttttgagt atttgaggt agtttagagt tttgaaagg 900
taaaatttat atataaaaag aagttttttt cgatttagat ttagttttg gtgtagatg 960
tatatttgag gagtaggtgg ttagtttagat ttttatttga tagttaata aaatttattg 1020
ttttttaatg gaattttttt tgtaagtctg aattgatttg ttatttttgt gatttgtgag 1080
atggtagggg agtattaaat attattatga tttgggttat agtgggggga aaaaaggaaa 1140
aaagaaaaa aaaatttatt gtttaagttt gtttagcgtg naaaggggtg gaattgttg 1200
ggttatttta tttgatttat tggaaataga gtggatttta ttaatathtt aataaagaga 1260
atttttttgt attaggttg aagtggctgt tagttttctg tgtaatttta ttttttgaa 1320
aagtgaatt agttgttatt gtttagcgtg atttgtgagg ttgagttta atagtttaa 1380
gaagtaaatg ggatgttatt ttcgctgggt tctgtttctg cgaggtggtt atttcgtatt 1440
tgttatgtaa aacgagggag cgttaggaag gaattcgttt tgtaaagttt ttggttttg 1500
ttattagttt ttatttaagt tttctgtgat gttgtgttg atttatttg gaagttggtt 1560
ggttggcgag gtagagtttt tttttaaagt ttggttttta cggaaaatat gtttagtgta 1620
gtcgcgtgta tgaatgaaaa cgtcgtcggg cgttttttagt cggataaaat gtagtcgaga 1680
atttcgttctg ttttgtcgt tttttgttt taggtagggg agaggggttg tctggcgcgt 1740
tttgcgttctg gttttgtat tctgatcgtt cgttacgggt aggggtgagg ggtttctggg 1800
gggtcgggtt tttcgtctgc ggcggcgaag atagatcggg gttcggtagg gaggttattt 1860
cgagtttaga gattttaggt atttttata tataggtttt tttttggcg tctgtcgtg 1920
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtacgtt cgttaacggg 1980
agga 1984

```

```

<210> SEQ ID NO 13
<211> LENGTH: 1984
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (804, 1771, 1776..1778, 1810, 1826, 1828, 1858, 1901,
1905)
<223> OTHER INFORMATION: unknown base
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (1923, 1931, 1933, 1935, 1936, 1944, 1952, 1967, 1968,
1976)
<223> OTHER INFORMATION: unknown base
<400> SEQUENCE: 13

```

-continued

---

```

tttttcggt aacgaacgta tatatatata tatatatata tatatatata tatatatata    60
tatataccta cgtacgtag agtgggagtt tgtgtgtggg ggggtgttag gatttttggg    120
ttcggaaatga tttttttatc gagtttcgat ttgttttcgt cgtcgcgatc ggaggtttcg    180
atttttcgaa agttttttta ttttggtcgt gtcgggcatc tcgaatgtag aaacggggcg    240
tagagcgcgt tcggtagttt tttttttta tttgggatag gagaacgtat agaacgagcg    300
gagttttcgg ttgatttttg ttcgattaga agcgttcggc ggcgttttta tttatgtacg    360
cggttgtatt gagtatattt ttcgtgggag ttaggttttg aggagagggt ttgtttcgtt    420
agttagtaa ttttttaaat agattagtag tagtattacg aaagtattgg gtagagggtg    480
atgattagga ttaagtgttt tataagacgg attttttttt aacgtttttt cgttttctat    540
ggtagatacg gggtagtat ttcgcgagga gcgagtttcg cggagggtgt attttatttg    600
tttttttga ttgtgggggt ttagttttat aaattacggt gggtaagtgt agttgatttt    660
atttttttag agaatggaat tgtacggggg attggcgggt attttttagt tagtgtaaaa    720
agattttttt tattaataatg ttaataagat ttattttatt ttaataaat tagataaat    780
ggtttttagta gtttttagtt tttntacggt tggttaagggt tggtagtga ttttttttt    840
tttttttttt ttttttttat tattgtgggt taagtattga tgggtgttgg tgtttttttg    900
ttattttata aattataaag atgatatagtt aattcagatt tgtagaaaaa attttattaa    960
ggggaatag attttattaa ttgttttagt gagagtttga ttagttattt attttttaaa   1020
tgtgtatttg gtattaaagt tgggatttgg atcgggaagga attttttttt gtatatggat   1080
tttatttttt agaagttttg aattattttt aggtattttg aatagtagta gttattgttt   1140
aagggatttt ataaatgggg gggtttggga ttgagggtaa gagggatatt taggtttgtt   1200
ttttttaata ggaatttaag aaaatgtagg gaagatttta ttgtattatt tttatgtagt   1260
ttttttttag aaagaattta aggtgtttta taataaagga tatttgtgta atagaattta   1320
tgaattaaaa tagaaaattt gggttagaag gaaaagttta gtaaattgaa atgataagtt   1380
taaattttta tgtattttgt ttgatttggg taagagtttt ttgatagtta aagtaataaa   1440
agaaatataa gttatatagt tttgtttagt aaatagtat ataagttttt ttagtattat   1500
tttttatggg ttaaataatc gaataaatgt attatgtaac acgtaataa gtaggatatt   1560
gagtaatgga ttatgttaag atttagtgga tggggatgta ttttttagtat ttggtagatg   1620
gtttggaata tggagatagt aattagtgtg tgttaagaa atagatgggt agatttaatt   1680
aagatgtagt agaaggttag gaggggtaag gagtttgag ttatattgtt taggtatagt   1740
tttcggattt gttattttat tgtaaattat ntaaanntt gtttttattt gtttatagat   1800
aaaatgaaan tataaattgt attgantnta taaggttgtg gtgaggatta agtgatgnta   1860
taattattga ataaatataa agtgatgtgt tttttaagat nttntattg gttgtgtgtt   1920
ggntagaagg ntntnntttg tggntttaag gntattgata aagatgnnag tgatgnttga   1980
agtt                                                                 1984

```

&lt;210&gt; SEQ ID NO 14

&lt;211&gt; LENGTH: 7833

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

---

-continued

---

<400> SEQUENCE: 14

gtttttggtg agatatgtgt tttataagtt ttaatggaga aaaatgtaag tattttattt 60  
tttgaaattt ggttatttga gtaatgagaa aatagttatt ttttttagga tagtggtttt 120  
taattatggt tatgtgtttt tttaggaaaa ttttaaaaat atatataat taatgttttt 180  
gtgttatttt tagggatttt aagtttttga atacgaattt tgtattagta ttttttaatt 240  
atttagtgga ttgtgatgtg aaattatgat tgagttttat tgttttaaga tgaataaaat 300  
tttttttagt attgaaatta taaattttaa ttattaaaat taattaaggg tatgggaatt 360  
aataaggtat agggaagttt ttatattata aaattatttt ttttaattat agtttattgt 420  
ttatatgtta ttgttatttg tagaaaaggg tgaaaaaata gtaaatttaa ttatttttag 480  
tttgaaaaat tatttagaaa tgaagatgac gattttgaaa tattgttaat attatttgat 540  
ttataaataa tgttttaata tattttattat atattgatag atattttttt atatgaatat 600  
tatatattaa aattaaggtta ataagtatt tagaatattt tatttatatt tatgtatttt 660  
aagtaggtta gaaattaaga tatgagttat taagtatgag atgttaaggt gtggggtag 720  
aaattatatt gtattttatt attaataatt aatataatatt ttaattatatt atatttttaa 780  
ttttaatttg tatattttta attattttta attatgtgta taaatataag tatataatatt 840  
tttatgtatt tattttatta tatttttatt tattttatta tataggggat ttttttaaat 900  
ttattattat taaattatat atttttattt taatttttag aataagtta ggaggtaggt 960  
attgttatta tttatatttt ataaatgagg aaattgttta tagttataaa gttattgtgt 1020  
tagatatatt agaagtttaa tatatatttg gtgaatatat gtataaaaat agagagatag 1080  
atatgtataa tagtttattt ttatattgag taaaagtttt taatttgttt tagaaaattt 1140  
tttgtagaaa ttgagtaaaa atcgaggat ttttttattt gttatatagg tataggtggt 1200  
attttatttt ttttaataagg atgaatattg aaatgtggat ttttaagttt aatttttagat 1260  
tttttgaatt tttgatagtg ggatttgga tttgtttatt gttttaaagt tttttaagga 1320  
atttatatga ttaattaggt tttagaaatta ttggatttta ttgtcgaagt ttgagaatta 1380  
aagtttgggt tttattgagg ttttatagaa agggtaaatg aagtattatg gatagaattg 1440  
atcgtttttt agttagtttt tttttttaga agttaatagg tagtaataa gtagaaatta 1500  
gtgatttatg ttttgtgttt tgaagtttag tagaatttta tagagtttta gtagtgttat 1560  
tgacgagatt tgttttttgg ggtaagttgt ttgatgtttt taaagttata ttttttttat 1620  
ataaatgag ataatttttt ttgttttata ggggtgtttt aaagattaaa taaaaataat 1680  
atgttttatt ttatatggta taatgtttga tatttaagaa gtaaaggata tattttattt 1740  
ttattgaagt aattagaaag tatgaaatta tgaaggagat aagagttttg attggtagtg 1800  
tattttattt ttttaggttt atttatttat tttaaattat tttgttgga gaataatttt 1860  
taagtttttt atttaagttg tgagtaattt tatattttat aatgatgttt tttttatgag 1920  
aaaaaaaaat gtttttaagt tttttggaga aaatataatt gtattatttt tattgaaaaa 1980  
tttaataatt ggattttggt tttttgtatt aatttttagag tgtatatgtt ataaataaag 2040  
tgttttagtt taagaagatt gaaagtaaat atggtatagt attttaaaat aagaattttg 2100  
taaatatag gtatgattgt gttatattat tagtaattat atgatacgtat atgtaaagta 2160  
tagtttatag attttaattt aattttaata agtaaattga ttttgttttg ttggggaaaa 2220

-continued

---

gttaaagtat taatttaatt gttaatgtag ttttgttat tttttggta tttagtata 2280  
agtttaaata atgtatatat ttttattat atatttagta atataatfff ttgtttaatg 2340  
agtgatgttt ttttgttatt tgggtggtt tgtagtttt agaatttgtt ttttggtggt 2400  
attataatat taagataga gtaagtgtaa taaaattgta gtatttttat tgaagagggt 2460  
ttgttttaaa ttgtttaata atttaaagga tttttgtgga agtaacgta tttgttaatt 2520  
agttataatt agtaattaat tttttggag ttttaattta tttttggtaa aacgttttag 2580  
gaagagtata tattattaga aagtatgtta aaaatttatt tagtagaaaa tttaaaaata 2640  
gtttttttt gtttaagagg tttttaaatt tttatttata tagttaaatt ttgaaatfff 2700  
agttaggttt gttttattat tataattatt gtataaatat ttttaaggat tttgttttta 2760  
gttttaagta tgatttattt ttataagttt gattagttat tatattagtt ttgttatgga 2820  
aatgatgatg tttttatttt ttgttgtaga gttgttaaat tttgatttat atttatggtg 2880  
ttttttttg tgaaggtttg tagcgaaga aatttttaatt tttttgtttt gtaattatag 2940  
ttggtagttt tatttaatgg gtattttgtt tttttaaaga atttagttgt tttgtttaga 3000  
agtcgatttt ttgatgtttt taacgtttgg ttttaattgat ttgttttaatt ggagttttcg 3060  
tcggtgagga gcgagatggt atcgattaga atgttgggat ttgttgttta attgtagga 3120  
gtgagagata ttgagattta gaaatttttg gaggtgggag gggagaggga tagtttcgga 3180  
cggaggcga gatgtaagat aaagggatgg attttatata ggaaaaaaaa aaagatttcg 3240  
ttgaggtatt gaggtgttg acgattatat tttttaaagg agaagttaaa aagtaaggaa 3300  
gtgggaggag gttggagggt aaagtattta aaagattat tcgggtataa tttgtttttt 3360  
tgttggtgtt tgtaaaggat agatagtttc gtttttaaag tatatgaatg ttttttttaa 3420  
gtgattggga atggatatta attgtttgtt aaatgttatt aaatgttttt ttaaatttag 3480  
gggatataga aagaggggta taaaaggaga atttaaatag aaaaagggag gattcggagg 3540  
ttttgaaag cggggggaga agaaggagga gggataatag agaggaatag agaaggagag 3600  
cggagagaag ataaataaaa ataaaaatag gaattattga ataattatat attaaaaaga 3660  
aagttttttt ttatggggta tttaaaatat tgagattgta atagtgattt cggttatgga 3720  
agaagatgt tttttttat ttttgtttc gaaagttttt ggtttcgta ttggcgatta 3780  
aaattttatt aggttaaaga gtgtgtttaa ttgtttgaag aatgtagtag acggaaggcg 3840  
ggtttcgta tgctgtttt tttttcgtt ggagagaatg aaagaaacgc gtagagttag 3900  
agatttttgt caggttagat ttttttcgt cgttttaggt tatcggttat tcggtaaaga 3960  
ttcgagtaag gaacgtaggg ttattgttg ggttaataaa tggagttcgt ttttttttt 4020  
tcgacgtcgt ttgttcggtc gatgttttcg gtaatttatt cgcggcgtat gtagaggagt 4080  
ttttttttt ttttagatt attgtttcg attaatttga ttttttaaat atatttgatc 4140  
gtatttttta ggtggatata ttaatagggt acgggttga gaggagcggg tgatgaggag 4200  
agggatttaa atttcgaac gtttgggtt ggtcggagt cgggggggtt tgggaggaga 4260  
gaggggagaa gagagaagga aggagagcgt ttgtcgggat ggttgagttg tttcggcag 4320  
tagttttggg gttgtacgtt ttttgggag atgttgtgtt tgtttttagg tcggtaaagag 4380  
cggttttaat attatcgttt tttatttttt tttttgtaa tttttagaga aacgtttttg 4440  
gttttttcgt cgcgatattt ttagtttga tttttttata gtttaggcgg cgcgttttcg 4500

---

-continued

---

tacgttgag cgtcggcgt tagtaggacg tttttttcgc cgtcgattcg ttttttttg 4560  
ttttgttgt gttgtttttt tgatattttc gtttttatta tttttagttc ggagagacgt 4620  
tatttagtcg cggttcgtat tcgcggttcg gggttacgcg cgggaagagg gcgtagttc 4680  
ggatttcggt ttcggtaggg ggcgttttgg agcggagagt gaggcgaatg gtatatgagt 4740  
gtcgggtag tttattttga agttcaggtt tttattttga gttatgtttc gtttagtttt 4800  
attcgggtta gcggtttggc agcgagtta tttgtggttt tcgcggtcgt ttttttttg 4860  
tatttttga tttattcgtc gatttttttt tttcgggatt tgtattttgt tttattaatt 4920  
agagttcgtat tgtttttttt tacgtgattt cgggcgggtt gaggatttgt tgttttttaa 4980  
acgttagagg gatcggggcg gtagagtccg agaggcgggt gtcgggttcg ggggcgtttt 5040  
gatttttttt tttattttgt ttttcgggtt ttattcgttt gtttttgat tttcgttttt 5100  
ttttgttttt cggttttttt gagttttttt tttatggtag tagtttttcg cgttttcggc 5160  
gtagtttttt agcggacgat tttttcgttt cggggttgag ttttagtttt ggatgttgtt 5220  
gaaattttcg agattatgcg cgggtttggg tgttgttttt tcgtcgggtg ttattgttat 5280  
cgtcgtcgtt tttgttgcg tcgttcggcg gatgtttagt agttcgttgt tcggttttcg 5340  
cgattttgtg tttttcggaa gtcgtttgtt gttgtagagt tgaacgaatt agttatggtg 5400  
ttgtgggagt tttcgcggta gtgtagtagt tggatatttt gcgaggggtt ttgttgggtg 5460  
ttgttgtgt tcgttatgtt atttatcgtg gttcgttcgg tgaagttcgt tgtttttttt 5520  
atttttttaa gtgattgtta aacgtttatc ggttgaatt gttttggtta gtttagaatt 5580  
ttcgttttcg atttttttaa ttcgtagaag aatacgcgta tttagtagat attagtttat 5640  
tttagcgcgt ttttttagtt tttttttttt tattgtttta gatttttaatt attatttatt 5700  
tttatttaga gaaataaggg gaattgttgt aggttcgggg gtgaggggtg gttttgggat 5760  
gggtagaaag ttaggtgtga gtaggaaatt tttgtatgtt tgcgtttata ttggagtgc 5820  
gaggattttg agaaatatta aacgggatgg ttttttgggt ttattgtttt gaaagagtat 5880  
taattttagg gaaatattg aatagaagt tttgttatta ttaaagaaaa aagttttatt 5940  
aggatgagga agaaataatt ttatgagaaa gaatgagcga gaaagtaata aattaaatgg 6000  
tgattgtagg ggaatcgttg atttttggtt aagggtttat gaggtcgtat tggtttttcg 6060  
ttgaagatta ggttatatag attttagagg agttgggttt taatagaatt tttttttttt 6120  
tttttttttt tttttttttt tttttttttt tttttttatt tatttttttt tttttttttt 6180  
ttattttttt ttttttttag cggtaaaaga tattgttttt gtagtttaga tatgtttttt 6240  
tttttgtttt ttttaagtttt aaggtagtat aggggagttg agaaaaagaa tattttgcgg 6300  
gttttttagg tcggagtggg tatgattgag gttggtagg ttttatgtag gcgagtcgag 6360  
ggcggaaatcg attttagtgg gcggtgattt tttttttttt ggataggttt ttgtggagtg 6420  
ggttaggat tttttttgtt cgttcgggtt ttttttagatt ttgacggcga acgtttggtt 6480  
ggtttcgttt tgttgaagt ttttaattaa atagggttag aggatgggag ttgttgtatt 6540  
tttagttggt atagtattcg gtttgatagt ttgtagtata gggtagtatg aattttttat 6600  
tttttgtgaa tataattttg ttgtagttaa atttgtttt gaataaagtg ttttttaaag 6660  
atgtatataa gttgaagtgt atgtaatttt agagaggagg gaatgattaa ttgtaattta 6720  
gggtgaaagt ttgtatagtt tttagttatt attgatgtaa atgttaaaag gaaaattatt 6780

-continued

---

```

atgtattatt ttaatttatt ttttataaag ataagttgag atatgtaatt ttattagatt 6840
tgggttaata gattgttttt ttttttggta gtttttaaat ttggtatttt aataaaattt 6900
aatatgtttt tataattttt tgatttatgc gtatatgtgt gttgtttttg aaagaataag 6960
ttttattttg ttattgttta attatttttt agatgtttta ttatggtaat aattatgagt 7020
ttgtaaaaat aatttttggg aatgttgatg gttttgtagt ttaatataga ttggtttgtt 7080
ttattttttag tttttgtatt gttttaggaa ataattaatt taaatgtgaa gttgatattt 7140
gtaattaaga aattatata ttattagata ttttaaaggg gattgtataa attaaagaga 7200
ataaattggt tttgtagata ggtgtgtaag aatttggtat ttcgttttta tttttgtaa 7260
tttagaggtg attaattttt atttgagtta aatagattat tatagaaaat attgtgtttg 7320
tttattttta ttattgaggt tttgtttttt tttgttttgg atatatttta aataaggggt 7380
tgtttttagtc gttgaagtaa aagaataatt aaagatgggg aaatggtaaa agggatatta 7440
gagattatta ttagtttttt tttaaaatgt ggagttttgt ggttataaat attgtttatt 7500
taatgagtaa aaaataaaaa taaaaaaaaa ataggaagta aatgtaagt ttttatttat 7560
tattgttagt attaacgtaa gttttaaaaa atagtattat tagaaaagga tattaagga 7620
gaattgatta gaaaagaatt gtggaaaatg gaaacgaata ttgattattt aattagattt 7680
tgaggttatt agtagatagt gattttgtag tatagttata gttgttggat ttaaaattta 7740
ggataagtat tttaaagttt taaagtagtg tttttttttg ttaaaaattt gtaagatgtt 7800
ttaatgattg gagtgttttt tttgaatttg agg 7833

```

&lt;210&gt; SEQ ID NO 15

&lt;211&gt; LENGTH: 7833

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 15

```

ttttaaattt aaagagaata ttttagttat taaaatattt tatagatttt taataaaaaa 60
aagtattatt ttgaagtttt aaaatatttg ttttaaattt taaatttaat aattatagtt 120
gtattgtaag gttattgttt attgataatt ttaaaattta gttaagtgat taatattcgt 180
ttttattttt tataattttt ttttagttta ttttttttta gtattttttt ttgatagtgt 240
tattttttta agtttcggtt aatattgata gtgggtaatg aaagttaaatt attgtttttt 300
tgtttttttt ttatttttat tttttgttta ttagggtgat aatatttatg attataaaat 360
tttatatttt ggaaaagagt tagtgatgat ttttgaatat tttttatta tttttttatt 420
tttaattggt tttttgtttt aacgattgaa ataatttttt atttgaatg tatttagata 480
aagaggaaat aaagttttaa taataaagat aaataggat agtggttttt gtgatggttt 540
gtttggttta aatgaagatt gattattttt aagttaatag ggggtggaagc ggggtgtaaa 600
gtttttgata atttatttgt aaaattagtt tatttttttt agtttatgta gtttttttta 660
aaatatttgg taaatatgta attttttgat tgtaaatggt aattttatat ttaagttagt 720
tattttttta aataatgtaa gggttaggaa tgaagtaaat tagtttgtgt tggattataa 780
agttattaat atttttaaaa attgtttttg taggtttata attattatta taataaagta 840
tttaaaaagt gattaggtaa tagtaaagtg aaatttattt ttttaaaaat aatatatatg 900

```

-continued

---

tacgtatgaa ttaagaagtt atagaaatat gttgagtttt attaaaatgt taaatttaga	960
aattgttaaa aaagagaata atttattgat ttaaatttaa tagggttgta tattttaatt	1020
tgtttttgta aaggataaat tagaatgatg tataataatt tttttttgg tatttatatt	1080
agtaataatt aggaattata taggttttta ttttgagtta tagttggta ttttttttt	1140
tttaaagtta tatatatttt agtttatata ttttttgaa agatatttta tttagagtta	1200
gatttaatta tagtaaaatt atatttatag aagatgaaaa attatatata ttttatatta	1260
taggttgta aatcgaatgt tatgttagtt aggagtgtag taatttttat tttttggtt	1320
tatttaatta ggaagtttta gtagagcgaa gtttgtaag cgttcgctgt tagaattga	1380
aggaattcga gcgagtaaga agagtgttg atttatttta tagaagttg tttagaaatg	1440
gaggagttag cgtttattga agtcggtttc gttttcggtt cgttttatag gagtttgatt	1500
agttttagtt atgtttattt cggtttggga gattcgtaaa gtgttttttt ttttaatttt	1560
tttgatttat tttgaagttt agggaaagtaa agagaggggt atatttggat tgtaaaatta	1620
atgttttttg tctgttagga gagaagggaa tgagagagag agagagatag atagatagag	1680
agagagagag agagagagag agagagagag agagagagag agaaatttta ttgaaattta	1740
gttttttag aatttgtgtg atttggtttt taacgggaga ttagtgcgat tttatggtat	1800
ttttgttag aattagcgtat tttttgtag ttattatttg atttattggt ttttcgttta	1860
tttttttta taaagtattt tttttttat tttagtaaga ttttttttt taatgatgat	1920
aaagtttttg ttttagtgtt ttttttagga ttggtgtttt tttaaaatag tgaatttaga	1980
aaattatttc gtttaaatat ttttaaaatt ttcgtagttt taatgtaagc gtaagtatgt	2040
aaaggttttt tgttatattt gtattttttg tttatttttag aattattttt tattttcggg	2100
tttgtaatag tttttttgt ttttttgat agaggtgggt ggtattaggg gtttaggta	2160
gtaggaggtg aggggttgag gaggcgcgtt agggtaggtt ggtttgtgtt ggatacgcgt	2220
gttttttgc ggagttaaag ggtcggggac ggggttttg gatttattag agtaatttta	2280
gtcgggggc gtttgtagt tatttaagga ggtagggaaa gtagcaggtt ttatcgggcg	2340
ggttacgatg agtagtatga cgggtagtag tagtagttag taaaagtttt cgtaaagtgt	2400
ttagttgttg tattgtcgcg gggattttta tagtattatg attagttcgt gtaattttgt	2460
agtagtaaac ggttttcgag gaatatagga tgcgggggt cgggtagcgg gttattgagt	2520
atttcgcgga cggcggtagt agagggcgcg gcggtgtag tggattcgg cggggaagta	2580
gtagttaaat tccgctatga tttcagagat tttagtaata tttaggatt ggttttagtt	2640
tccggagcag agggcgttc gttgagaagt tccgctcggag acgcgggaag ttgttttat	2700
aaggaggag ttttgggaag tccgaggata ggaggagac ggagtttagg ggtagacgag	2760
tggagttcga ggaggtaggg tggagggaga gttaaaggcgt ttcgtagttc gtagtcgtt	2820
tttcgagttt tgtcgttcgt atttttttg cgtttggaa gtagtaggtt tttagttcgt	2880
tccgggttac gtgggaagag gtagtcgggt tttgattggt ggagtaggat gtaggtttcg	2940
ggagggagg gtcgacgagt aggtgtaagg atgtaaggag gaggcggtcg cggaaagtta	3000
agatgggttc gttcgttagg cgttggttcg agtggggtta ggcgggtat ggtttaaag	3060
agaagtccg gtttagggg gggttattcg tatatttata tattattcgt tttatttttc	3120
gttttaggac gttttttatc gaaggcggg ttcggattag cgttttttt tccgcgctga	3180

-continued

---

tttcgggtcgcgagtcggg	tcgcgggttg	gtggcgttt	ttcagattgg	agatgggg	3240	
ggcggagggt	ttagaggagt	agtagtagta	ggtagagag	ggcgagtcg	gcgcgggaga	3300
ggcggttttg	ttggcgatcg	gcgttttagc	gtgcgggagc	gcgtcgttta	ggtttaggg	3360
ggatgtaggt	tgggaatgtc	gcgcgggaga	ggttagggac	gttttttag	ggattatag	3420
gaaagaggg	gagagcgat	ggtgttagaa	tcgttttgt	cgatttgaa	gtaatagtag	3480
tatttttat	aagagcgtg	aatttaagg	ttgttcg	aggtagtta	gttattcgg	3540
taggcgttt	tttttttt	tttttttt	tttttttt	ttaggtttt	cgtagtttcg	3600
atttagtta	agcgttcgta	ggttgaatt	ttttttta	ttattcgtt	tttttagt	3660
cgtagttat	tagtgtgtt	atctggagg	tcggttaga	tgtgttga	aggttagatt	3720
ggtcgggata	agtgtttga	gagaaagaga	aaggttttt	tgtatacgtc	gcgggtggg	3780
tgtcgggagt	atcggtcggg	tagcggcgtt	cgggaagggg	agagcgggtt	ttattgttg	3840
gtttaggtag	tgattttcg	ttttttatc	gggttttgt	cggatggtc	gtgattggg	3900
gcgacgagag	aaggttta	tcggtaggag	tttttggtt	tgcgcgttt	tttttttt	3960
tttagcggga	agggtaaacg	gtatagcggg	atcgttttt	cgtttgtgt	attttttag	4020
tagttagata	tatttttag	tttaatgaa	tttagtcgt	tagtaacggg	attaagagtt	4080
ttcggggata	agggtggaga	ggaatattt	ttttttatga	tcgggttat	tattgtagtt	4140
ttagtgttt	ggatgtttta	tagggaagag	tttttttt	ggtgtgtgat	tatttagtga	4200
ttttgtttt	tgttttgtt	tattttttt	tcgtttttt	ttttatttt	ttttgttat	4260
ttttttttt	ttttttttt	tcgtttttaa	aagttttcgg	atttttttt	ttttattta	4320
aattttttt	ttgtgtttt	tttttgtgt	ttttgaatt	taggagagta	ttgataata	4380
tttaataggt	aattaggtt	tattttta	tatttaaaag	aggtattat	atatttgaa	4440
aacgggatta	tttttttt	gtagatatta	gtagaaaaat	aaattgtatt	cgagtaattt	4500
tttaagtat	tttaatttt	aattttttt	tattttttt	tttttaatt	tttttttga	4560
gagatgtgat	cgtgtagat	tttagtgtt	taacgaaatt	ttttttttt	ttttgtgta	4620
aattatttt	ttattttat	atcttcgtt	tcgttcgaga	ttgtttttt	tttttttat	4680
tttaagat	ttttgaatt	tagtgtttt	tatttttgg	aattaagtag	tagattttag	4740
tatttttagtc	ggtgttatt	cgtttttat	cgacgaagat	tttattaaa	tagattaatt	4800
agattagacg	tggagggtat	tagaaaatc	gttttagat	agagtagtta	aattttttaa	4860
ggaaatagaa	tatttattag	atagagttg	taattaatat	tgtaaaata	ggaattagaa	4920
attttttcg	ttataggtt	ttagtagaga	agtaatata	aatatagatt	aagattta	4980
aattttatag	tagagaatga	gaatatgta	ttttttatag	taaggttgg	gtggaatta	5040
attaggtta	tgaataaag	ttatgtttga	aattaaaggt	aaagtttta	aaagtgtta	5100
tgtagtaatt	atgataatga	aataggatt	gttaggatt	tagagtttg	ttatgtaagt	5160
agaattttag	agaattttt	agtagaggaa	aattgtttt	gaatttttg	ttagtaaat	5220
ttttgtata	tttttaata	atatatgtt	tttttaagac	gtttgttaa	aagtaagtta	5280
aaattttaaa	ggagtaatt	attggtgta	attggtta	aaatgcggt	gtttttatag	5340
aggtttttta	aattttaaa	tagtttgaag	taaagtttt	ttaatggga	tgttgaatt	5400
ttgtgtatt	tattttgat	ttagtgtat	agtgtatta	agaaataat	tttgaattg	5460

---

-continued

---

gtaagtatta ttaagtggtga gaagaatatt atttattgag tagagaattg tattattgaa 5520  
tatgtaaata aaaatatata tattatttag atttgttatt aggtattaaa gaagtagata 5580  
agattgtatt agtaattgga ttagtgtttt aatTTTTTTT tagtaaggta aaattagttt 5640  
atTTattaga attaaattta agtttatgaa ttgtattttg tattgCGtat tatatgattg 5700  
ttagtaaatat gatataatta tattatgtat ttgtaaaatt tttattttta aatattatat 5760  
tatatttatt tttaaTTTTT ttgagttaga atattttatt tgggtatat atattttaga 5820  
attgatgtag aggagtagag tttagttggt agatttttta gtgaaatag tgtagatata 5880  
TTTTTTtag aaaatttaag aatatttttt ttttttatgg aaagaatatt attataaagt 5940  
gtgagattat ttagtTTta agtaggggggt ttgggagttt ttttttaata agaatagttt 6000  
aagataaata aatgaatttg ggaataaag atatattggt aattagaatt tttatttttt 6060  
ttatgatttt atattttttg attgttttaa taaaggtaag atgtattttt tgttttttag 6120  
gtgtaggta ttgtgtatg taggatagaa tatgttattt ttatttaatt tttaaaatat 6180  
ttttatgaga taaagaatat tttttattt tatataaaag gaatatggtt ttgaaagtat 6240  
taggtaattt gtttaagaa ataaatttcg ttagtgatat tgttgggatt ttgtgaaatt 6300  
ttgtttgatt ttagagtata agatataagt tattaatttt tgttgattg ttgtttgta 6360  
gtttttgaga ggggaaatta attgggaacg tattagtttt gtttatgata tttattttgt 6420  
tttttttggt gagtcgtagt aaggtttaaa ttttaatttt taaatttcgg taataagatt 6480  
tagtgatttt tgaatttgggt tgattatatg aattttttga gaaattttga aataatagat 6540  
aaattttaag ttttattatt agggatttag aaaatttgga gttgggtttt gggatttata 6600  
ttttaatatt tatttttggt ggagaagtaa ggtattattt atatttatat gataaatgaa 6660  
aggatatttc gatttttggt tagtttttat agagagggtt ttgagatagg ttaaaagttt 6720  
ttatttagtg taaagatgag ttgttgata tgtttgttt tttgtttta tgtatatgtt 6780  
tattaatat gtattaagtt tttaatatgt ttgatatagt aattttgtga ttgtagataa 6840  
tttttttatt tgtaaaatgt gagtaaat aatatttgtt ttttgggttt gttttaaaga 6900  
ttaaaataaa aatgtatggt ttaatggtag tggatttggg gggatttttt atataataa 6960  
gtgaatggag gtatgaaata ataaatatat aaagatgtgt gtatttatat ttatatatat 7020  
aattaaaaat agttaaagat gtataaatta aagttaaatg tatgtgatat tgaagtatat 7080  
gttgattatt gataatgaag tatagtataa tttttaattt tatattttta tttttatat 7140  
ttaataattt atattttaat ttttagttta ttttaagatat atagatatag atagaatgtt 7200  
ttaaatgtat tattgtttta gttttaatgt ataatttta tatgaaaag tatttattag 7260  
tgtgtagtaa atgtattaga atattattta taggttaaat gatattgata atgtttttaga 7320  
gtcgttattt tttttttggg ataatttttt aaattgagag taattaaatt tgttattttt 7380  
ttattttttt ttataatggt aaataatata taaataatga gttgtgattt aaagaaataa 7440  
ttttataatg taaaagtttt tttatgtttt attgattttt atgtttttta ttaattttgg 7500  
tagtttaagt ttgtgatttt agtgttgagg aaagtttatt ttattttaga gtagtgggg 7560  
ttagttatga ttttatata taattatttg gataattaaa gaattattgat gtagagtccg 7620  
tatttaaaga tttggaattt ttagaagtga tatagaagta ttggtatata tatattttta 7680  
aagttttttt ggagaaatat atagttatga ttgagaatta ttgttttggg gaaagtgatt 7740

-continued

---

```

atTTTTtTat tattTaaata gTtaagTttt agGagGtaaa atattTatat tttttttTat 7800
taaaattTgt aaaatatata ttttattTaa gat 7833

```

```

<210> SEQ ID NO 16
<211> LENGTH: 965
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

<400> SEQUENCE: 16

```

```

gattaattaa gggTatttta gaagTtaggt gttttTgtg tttttttTga gTatGgaggt 60
tattaatttt tagggggaag agatGtagtg tgagGtaggg gTgtTgtgtt aagaaatttt 120
gatgtttttg gggattgagg ataaaggtgt ggatattgatt ttggggtatt tggagttttg 180
tgatttTgtg tatggatggt atatttaggg gTtaattttt gttttTgttt aaagaatttt 240
aagTtagagt ttttTgtttt gtttatagtt ttgggatggt gttgtTgtgt ttattgtata 300
ggtagTgttt ggattggTtg tagtagattg tgtgtTgtgt gttttattgg gagatggTgg 360
agatgtTgaa aagTtttttt tttgttattt tggatgtTgt gggTgGtaag tgttttagtt 420
tttatttttg ttgagttgaa tgtttaggta tagtGgaatt gaaattTggt tttttatttt 480
tgttgagTtg aatgtttagg tatagtGgaa ttgaaatttg gttttgtggg atgtgagagt 540
tgttgaggtt atgtgtaatt gggTgtgatg gagggTgttt gtttTgtatg tgtgtaggtt 600
tgatgtaagt aggttattgt tgtgtgagtg tgtggatgtg attgtttgag agatttggag 660
gtaggtttgg gatattTttg agtgaatatt ttaggatatt tttttGgtta gTattTgttt 720
tttagTgttt gtgattTtag gtgggtatat gttgggagat agTaatgggt ttgggtgtgt 780
gTaaatgagt gtgattGgaa gtgagTgtga gtttgattta ggtagggatt atatagtatt 840
gttatatttg tttgtttttt agtagaggat tgaagtgtgg gggTgggggt atggggTttg 900
aatagaatgt ttttgggata ttttGgTaaa tagtagttgg aagTaaaggg gtagttgtgt 960
aatg 965

```

```

<210> SEQ ID NO 17
<211> LENGTH: 965
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

<400> SEQUENCE: 17

```

```

tgttTgtata gttgtttttt tgtttttTgt tgttTgtttg taagatgttt tagagatt 60
ttattttTgt tttgtatttt tattttTgta ttttagtttt ttattaaaga gtaggtaggt 120
gtgatagtgt tgtgtggTtt ttgttttagat taagtttata tttgtttttg gttatatTta 180
tttatatata tttaaattta ttattgtttt ttaatatgtg tttattttga attatagata 240
tTaaaaaata gatattgggt agaagagtat tttgaggtgt ttattTaaat gTgtttTaaG 300
tttgtttttg agttttttgg gtggtTgtat ttatatattt gTatgatgat gattTgtttg 360
tattaaattt gTatatatta tgaataggTg ttttttatta tattTaaata tgtgtgattt 420
taatagtttt tatattttTgt agaattgggt tttagtttta ttgtgtttga gTgttttagtt 480
tagtagaggt agggaaTttg gtttttagttt tattgtgttt gagtgttttag ttttagtagag 540

```

-continued

---

```

gtaggatta aggtgtttgt tgtttatagt gtttagagt gtaagaaaga agtttttag 600
tgtttttatt attttttggg ggaatgtgta gtgtgtgatt tgtttagatt ggtttgggtg 660
ttgtttgtgt ggtgagtgtg gtagtggtat tttgggggtg tgggtggagg taaggatttt 720
agtttgaggt tttttgaggt agagtagaaa ttagttttta ggtgtgtgtg ttgtggtgtg 780
agttatggaa ttttaggtat tttgggggtg tgtttgtatt tttgtttta gtttttagaa 840
gtgttgaat ttttagtat gatattttt ttttgtgta ttttttttt ttttaggggt 900
tgggtggttt tgtgtttaa gagggtagt ggaatgtta attttaaga tgttttaaat 960
tgatt 965

```

&lt;210&gt; SEQ ID NO 18

&lt;211&gt; LENGTH: 16579

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 18

```

tgttttttaa aatgatgttt agttgaatta gttaatggag ataaaaatat ttatatataa 60
tttttttta agtttattag ttattttata gattagtatt aattttggtt attttaatga 120
gttaattttt ttttagagat attataaatt ttttaaaagt ggaaaataat tgttattttt 180
ttgaataatg tttattattt aaaataatgt atataaaagt tgttaataa atttttgttg 240
aattgagaga aaatataaaa aattgaaagt agtttgttg taataagtgt ttgttaagtt 300
ttatagattt tttttttttt ttggaattat tgtaataaat atttgatttt tgattttag 360
gattaaaatg tgatttttta tattttaata ggagtaagta ttagttttgt aatttttttt 420
gttaatagaa taaaaatgga atgtatgttt gttgttttta tatgttggtt gatatttatg 480
ttgaatattt tttttataat gtagaggttt tttgtttttt tgttttgtgt tttatttttg 540
tttattatag agatattaat gatgttatgt gtttttttgt attattagtg tagtagtttt 600
tattatgttt tttgttttgg ttttattgta gtttaattaa ataaatggaa atgggtagt 660
attgaaaata aatagtaatt gtaaaagttt tgttttttaa gaaatttttt gtattgtatt 720
ttgttttgtt ttattgattt taagattttt attgtttttt tgtgtttaga aattagattt 780
tttagtttgt tttgttttgg ttagataaat gttttataaa aatgatttat atttttttaga 840
ttgtggggag agaagaaaaa gattttgaga aagtgattaa ggagggtgta agagtgatat 900
aatttttgat ttaatttagt tggattgtgt tttttatggt tttatttttt tttttataaa 960
aaagatagtg tttttgatgt tttgttttag gtagtttttt aggggtgagt tagatgtgtt 1020
gagttatttg gttgttttga ggattggttt agagagtttt tttttttatt aatgaagtta 1080
tttttttttg ttagtttggg tttttttatt attttatggt ggaatttggg tttggtttgt 1140
atagttttgt tgtgatgtag atggtttttt agttaagttt tagagaaatt tgtttaatta 1200
tgaataggat atttatagta taggagaaaa aaaaatttat taaaaataat attttattgt 1260
aatattttga tttgggggtt ggagagttaa aatttaatat ttgatgtttg ggaaaaattt 1320
aatgttttat ggttttttaa ggaatttgt ttgtgtttta tttattagag taatttatat 1380
attttgggtt aaagtataat gagattgtta ttatttttt gttttttgta agattttttt 1440
tttaagaata aaaagaaggt aataaagatt ttatattttt gtttttttta gaaagttaaa 1500

```

-continued

---

gagaatttaa agtggtgat gaaatgttg tttttaatt gtagtgttt atttaaaaa	1560
aaaattatga tataatttta gaggataggt ttttttttt tagggtttt gaaagttaa	1620
agtatatttt taattttag tattattgta ttatttagtg aggataattt gtttagttgt	1680
gttaggatt agttagatta aggagttgaa attaagtttt ttttgaattt gaatgggatt	1740
ggaattttgt tattttttag gtagtaatag atgatgttat gtttaagaat attttatttg	1800
attaataaaa aaggagaagt ttgaaatgta attttaaaga attttagaag gaggtttaga	1860
tgaagataat aaaatttttt ggagatatat ttgttattta ttttttttt ttttaagttta	1920
ttggtttgtt gtttttgagt ttagttaaat ttttaagat aagattaaag taaggaagta	1980
aggatatag tttttggttt atggtttttt ttagatataa gtattgtgaa tgaagatttt	2040
gttttattta tatttgggat tttggagttt ttataaagaa tatttttagt ataaggatgt	2100
gatttaaagg gggatttata gaataatatt ttaattattg tttgattat atttgagtgt	2160
aatggataaa taaatataat tttagtata aaattgaaa aatataatatt tttttattta	2220
aaaatttttg atagttttgt taggtattgg agaattagaa ataaatgttt ttttattttt	2280
tttttttga tgaaggaaa ttattaaaa tagttatata gaaataataa tataattgta	2340
gattttattt taaatatagt tggatgttt ttttttttt tttttattta tttttattt	2400
tagatagttta gaggggaaa tttgttttag agatagaaa aagatattta tgggtataa	2460
ttttatttta gttatttgat agtttgaaga agtatttaatt ttttgaata ataattgtat	2520
aagttaaaga agagaaaaa taggtagttt aaagtttttt aatattgtgt atattttaat	2580
tttttaagtt agttttgaga aagtgttttt gttttgtaga agtatttggg taagtgttgt	2640
gaagtttatt ggtattaaaa gaaataaaa attttggagg attagagaat taaatttggg	2700
aatataaatg aaaatatttt agaattgaaa aagtatttaa tttattattt ttaattatgt	2760
tttattatat atagatggta ttatttatag tgaatgttaa ttattaatat ttttttaaaa	2820
atgggtatta atttgttatt agttgtttgt ttttgattta gttttttttt gttgaaagg	2880
tttatttaatt tgtattataa agggataaaa gggtaaaaga taataagttt tttagattgg	2940
tagatttgtt atttataaaa tggttttggg taggttattt gaattttttg aaatttaatt	3000
agattttattt aaatttaaga atgagaatat aaatttatat tttgaagtgt tttatagaaa	3060
ggtttatttt aatgtttgga gtatatattt taatgaatat ttattttatt ttattttttt	3120
ttatttttga attaagtaat tttgaattta aagttgttat gattagtatt gaaaagatta	3180
ttggattatt aattgtgtga ttttgggata gtaatttttt gtatttttagt ttgtttatat	3240
gttatatag aaggttgaag tttgattttg ttttgtgatt attattttaa atatttgatg	3300
aaattaaatt ttagtgtttg gaatggtagt ataataaatt tattaagaat aaataattta	3360
ttgtaaaaat atattgattt ttaaatgatg taattgatag ttatattatt gtagagggtt	3420
gataaataat aaaagaaatg aaagatgtat atgggtgagaa ttgaaattat tttgataagt	3480
tttttatttg tttattattt aaaattaatg attatgttga atgtttataa attataaaat	3540
ataaaagaaa ttttataaat gtgtatgtat aggagttaa gttattaaaa gttttaaagt	3600
ataagtttaa attaaattaa ttaagaagt tgagaggaaa aattggtttt tatttttaat	3660
tattattgtt ttgaggtttt atgtttaata taatttttta agtagagggtt ttagagagaa	3720
gagttgtgag gatattttta tatttgtgta gaaggaaaag tttgttattt atttttagt	3780

-continued

---

ttttagtggt atattgatgt gtattttgga tttattttgt ttttattgta taaatttata	3840
tttgatttta aagaaaagga aaatttaaag tttttttttt ttaaggggat agaaattttt	3900
tgtgttaatt gtttgatttt tttttttgta aggttttatt ggaaattttt tgtaataata	3960
tgtaggggat ttttttatgt gttgatgttg tttatatagt ggggtgggtt tgattgaaga	4020
aaaaaattg tatatatgta tgaaagatta tggttttatt tttggaaagt atgaaagggtg	4080
attgatattt ttaagaagtt tttgttattt aggaaaatta ttaaatattt tatttagaga	4140
tatttgaaa gattgaagga aaggaagaat gaagaaagta gaatttagat ttagtggggg	4200
agagatttgt ggtagaggaa aagtattttt tttgaatttg ataagggtt tgtttggggg	4260
aatttttgt ttagtttttt attattaggg ttttttgaag ttgggttttt tattgggtag	4320
ttttttgga gtgtagtggg gaatttttat attttttttt taggtttttg aaggattttg	4380
tttttttagt gtttttttta ggttggtagg agttttgagt ttgatatttt tttttgatgg	4440
gataggaag ttttgggtt gtgtaaatat gttgtaatta agtttttgt tgattttata	4500
gttttgtgtg tttttgagaa gaagtgattg tatttaattg tttattgttg gtttgttttt	4560
taagagtttg ggggtttttt ttttttaatt tagaattagt tggatggggg gtggggaagt	4620
gggggtgggg aaggagtggg agggtagtgg tttttgtgag tagagtgatg ttattgagtg	4680
agttttttaa tggggagtgt gtgtgttttt aagttgattg gtattttttg tttaggaagaa	4740
atgtaagag gtgggagtgt ttggggaggg aggtaggtgg tttttattgt aggtgtgggg	4800
agttgttttt ttgttttttt gtttgttttt taagtttga tttttaggag tggttgaagt	4860
tgtggagtgt ttttggagt ttgtgaatgaa tttttttttt tttttttttt tttttttttg	4920
ttgagttttt tttttgggtt tgatgggata gtgatataat gatgatgggt gttataattt	4980
gtatttgaat ttgtagggtg gttgttttga gtttttttgg ggaagaattt taggtgtgtg	5040
gatgtaatag ttgagaatat taggtgttgt ggataggagt tgggattaag atttttggtt	5100
agttttgat ttttttgat ttttttagt tgttttgat tttttgtatt tttttttggg	5160
ttattatggt ttttatgtga tttgtttggg taatgttga ttttagttgt tagtgttga	5220
gtgaattttt tttttaaatt gtaataagtt gttttttaa gtaattatgt ttttttgtt	5280
ttttttttaa aaaaaaaat taaaaaatt atagaaaaa atttgtgagt tttagaaaaa	5340
gaagtaattg gtagaaggtt ttaattaagg taaagagttg taaggtaag ttaagaaaat	5400
gtaggatatt aaaaaatgta ggtaattttt ataaagggtt ttggggagag gtatatagag	5460
ggattttggt gttgaaaaag atttagataa aagaaattt ggggtgggtg ggggtaaaat	5520
gattaatgga attgggggaa gggaggggat aaattgtaa gaaattatag aaaagtgggtg	5580
ggtttttgag ttggagagaa gagagggatt tttggtattt tgattttttt gttgttgggtg	5640
tttttaatat gtttaggta aaagtttga tggggattat taagatttgt tatagataag	5700
tttttgaagt tgttttgggt taggttattt ggtttttttag tttttgggtg gtggtttagt	5760
tttgggtgtt ttggaaggt atttttgggt agtttttgat agtgtgattt ggtgtttaa	5820
tagtttggga ttttgtgtgg atttgatttt ttttagattg aggtagtttg ggaggaggtt	5880
tggttggggg aggtgtagga tttttgtgtt gtttttttga tgatttgggg attgttatgg	5940
tttttttttg gtgttttttg gtttttttgt tttgtatgtg gtgtgaaggg gtttagtagg	6000
aaggagtaga ggatgggggg tggggttggt ggagttttgt ggaggttttg gaggtttttg	6060

---

-continued

---

ggtgggaaaa gtttgttttg aattggtagg gatgtgtaat aatTTTTTTT attttgaaga 6120  
gtgaaatagg gtttgttgtt tttattttta taagtaaatt ggtattttaga gtgtatttga 6180  
gataaagggtg gtttggggat ttgaatagg ggtttttggg ttagtttttt tgttttagatt 6240  
ataggagttt tttgtttttt tataatattt ttattttttt ttagtttttt gtttttagttg 6300  
agtaatttta ttggtaggtt tagtggtagt ggtgtgtgtg tgtgttggga gttttggggg 6360  
atgtttttgg ttggagtgtt ttattgtttt gtagagggtg ggggtttga ggggtatatt 6420  
tttgggtgtg gtagatttag ggtatttggg ttgttttggg ttgtggtttg tgttattggg 6480  
tgtggagtgg tttgggtgtt aggtagagga gagtgggga gaaaaatagt gttatagttt 6540  
aaattatttg ttttttaatt taattgtgtt ttggttaagt ttgtatgtt ttgggagatg 6600  
ttgtgggaag ggggagaaat ttatatggg ttggagagt tgttttgtt tgtgtatatt 6660  
tgtggtagag ttttttggg ttgtgttttt ttttttggg tttttttttg tttttatttt 6720  
ttttgttttt tttttttttt ttttagtttt gttttttttt ttttagtttt gtgttttttt 6780  
ttttatagtt gataaatgaa tggtttaggt gaaatttttg tttttttgt gttttaaggt 6840  
agtagggagg gaggagttag ggaggtggg tgtttttttg gatttgtttt tagtttagtt 6900  
tataagattt gtagaattta gatgtttagg aattgggagt tttgtggtgg gtgtgggtgt 6960  
tttttgatgg agaagttttg tatagggtga gaaaaataag ttttttagat taagtgtgta 7020  
ttttttataa ttttgtttt aaggatggaa ggttttagtt ttttttaag ttatttattt 7080  
tgtattttat aatttgggt gtatatttag gagtttggg gattttgaaa aaaggttttg 7140  
gtttgtgtaa agttagtaga tgtttttttg tgagtgggtt ggaatgtat tgggtttttg 7200  
ggtttagttt gttgttttag ggatgaatgt atatggttgg ggtggagtta ggttttggag 7260  
gtttagatgt gttatttttg gtgtttttta gaaagataat atatataatt tgtttttaaa 7320  
atttatgatt atttgaattt aatggtaggg ttttgtttta ggatattgta aaagaagggt 7380  
ttttaattta aaatttaaat ttttttaatt ttaggtggg tgttggatgg gaaagtggag 7440  
agaaggtggg tagtgggagg aaaaaagaaa gggaaaggaag ggaagtggaa ggaagagga 7500  
ttgtggaggg gaggaaaagg agaaaggaat ggagatggg tagatggtat ttaaattagt 7560  
aggaatgat tggaaaggggt tgttttggg tttgggtttg aattttgaga tttgttttag 7620  
gatatttttt tttttttttt agtggtaata aatgtgattg tagaggtgtg ttgggatgga 7680  
agggtgggaa agaattagat aagtggaggt ggttttagtt tgtttaaagg ggttgggttt 7740  
gggtttgtt tttatttgtt tattttgtgt taaggagga tttgtaatgg gtgttttttg 7800  
tagtggttgg gttagttagt tgttatgtgg gttggtgttt tggtttttag ttgagatggg 7860  
tgtgattttt tgggttttag gattgtagta tagttgtata ggtggttggg gtagttgagg 7920  
ttttatttta ttgaggttaa tgggggtggg gtggggaagg tagggtttag ttttaattgag 7980  
agagttgagt ggttttgttt ttgggatttg gggttggtt tttttggagt tgatttgagt 8040  
tttattttaga tagtgtttaa tttttttatt aaaaggttat ttaagtttta ttttatttta 8100  
ttttattttt gtttgttttt tgggagtggg aggaggggta ttgtgttttt gttttgaggt 8160  
aggggagagg taaaaattta ttttaagtaa attttttgtt agaaaaatga gattttatat 8220  
tgtattttat ttgttgattt tttaaatata tttaggttat tttaaagttt attttttagtt 8280  
aattaataga ttaattttgt tttagattgt tataaagagt ttgtttgga tggtaaatat 8340

-continued

---

tgttatttga taagattaat agaattttat gggggtaaat atggaagggt tgtgagtga	8400
attgagggtt ttgtgtttta atgaggaatt ttagtgggga tttggtgggg attggaagtt	8460
tagtatttat ttagggagtg attgagttgg aaatttgag aagtttggg tagatgagat	8520
ggtgtttttt agagttgtat ttttaagttt tttattttgt tgtatattgg aagaattgga	8580
ttttaagat tttttttatt attttagatt ttgtagttaa tagttatggt tttttttttt	8640
tgtttgttga aagttgtgat ttaattgttt taagtatttt ttgataattt tgttaaattt	8700
attttgaaaa tatttatata aagttttaaa ataggatggt gtgagtgtaa tataaaaaata	8760
ggaaggtatt aatggaagta ataattagtg ttttgagata gaagtggttt ttagtttga	8820
atttgaaat gtaagagta gtatattagg gagaagggtt ttaagagttt gattttttgt	8880
tttgagtttt gagtaagaa agaaatatgt tagaattgta ttgtttgttt gtagtaatat	8940
tgtaaagggt ttaaatggt ttaaatgtga ttgagtaatt atttttttta gtattaagtt	9000
tatagttttt aatttttttg tgagaagatt aggtaggaag tgggatgaga ttttaagataa	9060
ttatttagga tttatgttag aatttaaagt tttgattttt ttggtagata aaaatagagt	9120
tggttttaaa ttgtattttt tgtttttaaa aaaatgaggt ttgaaaggaa tttttttaga	9180
tttatatagg tttatagatg taagtttttt tttttttttt gtgtttgtaa attggaggga	9240
gatgtttttg atatttttga atatttgtat ttatatattt gtaaattttg ttaggaaagg	9300
atttaaatag ttaaattaa ggtgattatg ataatgtta aagatattat ttattgtttt	9360
ttttttgtt gtattattag gtttaaataat tttggagtta atgattgggt atttaattga	9420
gtaatagagt ttttgagagt gtaaagatta attaattttg ttttttaaag tttggttaatt	9480
aagttttttt tttttggtga gatttaata tttttgttga tgttattat ggataaataat	9540
tggtaaatat tgtattttaa atgggatatt gttttgtgat ggagtagaat tttttaagat	9600
ataggtttt tttgaaagaa agaattaatg attgtagtaa ggtatttttt tattatttat	9660
gtagagaaat tagttaggat tgataaataat ttattgttgg gagttatttg ttttggtttg	9720
agaaaatag ttgattgtta tataggtttg tattttatta gaaaaataa agtatgtatt	9780
ttttttgatt tggaaattaga aaagggggat tagatatgga tgttttggtt aagttttatt	9840
ttaattagtg aatttttgtt agtttttttt tgtagtttag atttggtaa ttatatatat	9900
ataaagtttg tgtgtgtgtg tgggggtgtg gtgtgtgtgt atgtatatag ggtttttaaa	9960
tattgttaag tttttattat tgaatgtgtt agagggtggt gtattttaga ggttagtata	10020
aagaattttt gtttaagtgt ttagttttgt tttttgaaat atatttttgg gtagatttat	10080
taattataaa aggatataaa tgtatatatt ttttatattt gtatatattt tattttattt	10140
ttttaaggat ttaagaggt gtttatgttt ataaaatgag attatatatt ttataataaa	10200
tatatattgat aaattataat ttttttatta ttaaattatt gatttgatat tatttgaggt	10260
atagtgaat gagtattaga tttggaatta gaaagtttga gttggattta ttagttgaga	10320
aaattaagta aagatattat ttttgattt ttagtgttat aatttattaa atgaatatgt	10380
agttataaag attaaaagag taagtatttt gaaaaattat ataaaaagg gatgataata	10440
tattatgttt tgggatttat attgatttaa ttttataatt ttaataataa ttgtattatt	10500
taagttatgt agtttttatt gaatttataa aataaattag agaggggtat attgttttaa	10560
aaattgttta gttttattaa ttagaagata gattatagaa attgatttta atgaagttgg	10620

-continued

---

atatttagtg gaaaggaaga attatttaaat aaatggtttt aggataattg atgggtttat 10680  
gtaaaataag aattagtttt tatttgattt tttattttaa aagaaatata agatgaatta 10740  
aatatgaaat gtagggaata attttttata aaatgataaa agaaaattat gggattaaaa 10800  
aaaataagtt aggaatagag aaatttgttt taagtaaaat ataaaattta gtgttatttt 10860  
aggaaatatt gataggtttg attaagtaaa aattatataa aagtaataa tatatttgga 10920  
aaatatttat gatatttgag agataaaaag ttaataattt ttgatatttt atggattttt 10980  
atgaataatt ttgaaaagt tataatttatt aaaatagag taaaaattat gtataggtag 11040  
tggtagaag agatatagag ttttaaaatt atagaaaaa ttgtttagt ttattttaa 11100  
ttaataaat agtaattaa atattaatga tatgttatat tttttttta ttaattagag 11160  
gagtaataaa aaaaaataat tttttttttt tttttttttt ttgagatgga gtttgttatt 11220  
taggttgag tgtagtggtg tgattttagt ttattgtaat ttttggtttt taggtttaag 11280  
taattttttt gtgttagttt ttttagtagt tgggattatg ggtatttatt attatgttg 11340  
gttaattttt gtgttttag tagagtggg ttttattatg ttggttatgg tagttgatt 11400  
tgaatttttg attttaagt atttatttat ttttagtttt taaagtgttg ggattatagg 11460  
tatgagttat tgtgtttggt aaaaatgaaa taaaaagtt tgatgtaagt ttgtgttgg 11520  
taggttggtg agaaattggg ttttttatat attgtgtat gaaagtaaat aggttttttt 11580  
attttaaaaa ataatttggg aattattttt ttttgatttt ttattaaaag ttttaaaaag 11640  
atgtatataat tttttgattt agtaatttta ttttttgata ttattatag gatggatatt 11700  
ggtaataata ttatttaaga taaatggata aggatgtttt ttgtagtatt gtttttaaga 11760  
gtaaatattt gaaatagtg taataggttt attataggga attatttaa taaaagtgt 11820  
tgtatttatt taatggaaa ttatgtagtt gttaaaaata ataggaata attgtatttg 11880  
ttgatttga atatttttta aaatatattg ttttttaagt gaaaaattg atagataaaa 11940  
ttgagtatat attatataat taattatgta aaatgggtgt atatttttt ttttaataaa 12000  
tatttattga gtgttttagg tattgtggat ataggagtga ataaaatatt aaagattgtt 12060  
ttgttgaat ttttattatg gtaggtagag aaagtaatt ttaattata tatgtaaaat 12120  
ggtagatagt gtaggatga aaatagagta tgggaaagg aatagagtat agtatgtttt 12180  
aggtagagta atgagaaagg agtttattag gtagagaata gaggagtgtt aaggggtgat 12240  
aatgtgaaaa agtttagtta agtttaggt taggaattta tatgatgta ttatttgat 12300  
gaagtaaat tatgaaattt tgtattttaa ttatgtgatt tttttttgt gtataatgat 12360  
ttttttgtg agaaagaatt agtgagtaat atgatatgt atgtttatta ttttatttt 12420  
gtagttgaag ttttttagaa agttgggtgt ttagggtatt taattagtag aatgggtgat 12480  
aaaattattg tgtttttgt ttaagttat tttttgttag ttgagtttt ttaaaattaa 12540  
gggagaatat taaatattta ttaatagaa ttttggtttt tttttggagg ttgtttattt 12600  
ttggtgtgtt ttttattttt tttgtagata ttttttatgg ataggagttt tattgatttt 12660  
gtttttatg ttttttaggt aattaaatag ttttagtatt ttgtatttg attgtttaat 12720  
gatttttttt ttttttatt gattttattg gagtagaagg ttaaatatag gtatttattt 12780  
ttatattttt taggtttttt tatgtagata gagttgttat tattagtaga tgggttttat 12840  
aagagagata tggaaaagt gaagaatgga aatagggttt tggagtgttg ttaaaaaata 12900

-continued

---

agaaagaggt ttttgaaagt tttttattta atatgattaa atatagaaga taaaagatgt 12960  
aatttaatga ttaaggatag ggatatagtt ttgagaaaat tatagaagtg gaatgaatgt 13020  
tttgattttt aaagaagagg gaatggaaat gtagttttga ggtagttaaa gttataaag 13080  
ttttttgata tatttattgt agaaataaat gattatTTTT taatttaagg tttgtaaatt 13140  
taggtgatta taggggtgag gtatggagt gaaatgggta agagattaag tgggggtagt 13200  
aatggttttg agtatataga ttttaagggga gagattggtt tttggttaat gtttatgatt 13260  
agttagttag aatatagatt tagtggtggt agattttttg attttttgaa agaagttaga 13320  
gattttatat ttagtgtaaa tagtttaatt tttaaattga gtaatttatt taaatatatt 13380  
tatgtaattt atttaaatgt atttagagta tgtttgattt tttttttgta tttttattta 13440  
gttttagggg taattttaag gaaagtgatg tgatatattg ttgtattat gtagatggag 13500  
gattatTTTT ttttttttaa taaaagatat ttttataagg attgagtgtg gtataaagag 13560  
tttgattttt agtattttaa ttgagttaga aatttaata tagaagattt ttttttgggt 13620  
tatatttttg gtttttttaa atgagatttg attttttaga tattgataaa tattatatta 13680  
ttgggttatg gtaattttt aagtggattg atttaaaagt gtttaaaagt attttaggtg 13740  
ttaaggaatt gttggtatgg agttttaaag gtgtttttgt ataagaaata gaggtgttga 13800  
ttatgtattt gattataaaa tattttttaa ggatatattt tgaatttggg aatgaagaaa 13860  
tattgtggga attattaagt tgaataaagt attagttttg ttttttagaa gtttattggt 13920  
ttatagagga gattagatat gtttatagat aatttgaagt atatagtaat aaattatagg 13980  
atgatttatt ttaagtataa agtgtatatt gttattatag tggtaattta ggttataatt 14040  
tttaaagtga agagttttaa atattatttt tttatttttt agtgatatga aattatggtta 14100  
gaatatttgt agaaagtttg aattagtaat ttttaataa aatattgagt attgatatat 14160  
tgatatattt tgagattttt tttatatata ttattttatt aatttttatg attattttgg 14220  
taagtattaa tttttataat taagttagta tagaaaatga ggtttagaga gatatagtat 14280  
tttattttaa gttatagaat ataagtagag aatgtagaat ttgtattttg ggttatttga 14340  
tagtggagtt tagattttta aaattaggtt tataaagatt ttttgaaaa aaagaaaaat 14400  
aaaaataga gaataaattt ttttgaagaa aagtaggaaa ttattattta taagtataat 14460  
taaagagaga tttttaattt tttttttttt tttttttttg agatggagtt ttgttttgtt 14520  
gtttaggtgt aattttagtt tattgtaatt tttatttttt gggtttatgt tataaagaaa 14580  
gtttttattt gattggtttg atttgattgt ttattagaat ataggttttt tagtaggggt 14640  
aataagttat aatagtttta gtgatataag ataaataaa attaataata aggttaattt 14700  
agaattatta atatatgttt tttattgtga attattagtt ggtttataga tggaaaatat 14760  
aagatatata atgtattagt ttaaaaaatg ttttttagag agtgttttaga attatagaat 14820  
ttaaaaaatt tgttttgaaa aggtttaatt tttattttta aatatataaa gaaattaaat 14880  
gtaaaaaaat taagagaaaa ggttttaatt aattaaatgt tgtataattt taagttttat 14940  
ttatgtttaa gattagtttt gtaatttatg attaatttta gtattagttt gtttttaatt 15000  
taggtttgtt aaataatata gaattttttt tttaatgaag ttttttaggg ttttagttga 15060  
aaaatatata ttagttagga tgtttttta tgttaaggaat aaaaagttat gattttattg 15120  
ggtttaaaata ataagtatat ttataatttt atataaaatg aagtttaag gtatggttta 15180

-continued

---

```

ttttaggttt aattttaag gatttatggt ttgttttgt tgattttttt ttagttttat 15240
ttttagaggt attttattta tttatggatg taaaattttt ataatagttt tagtagaaaag 15300
attttagtat gataatgttt aggggaagaaa attgattatt tttttttttg tttttttttt 15360
aggatgaatt tatttttttt agggattttt tattggagga agagattttt ttatatttta 15420
ttggttaaaa ttatattata tattgtttta aatttaatta ttgaaaagg gaatggtatt 15480
tttataattg atttaggata ttaggattat tttttgggtt ggggtttgat ttaggatttt 15540
ttaaagtatg tgattgaggg taggtattag aatggaatta gagttgtggt agaaaggaga 15600
aatgtatggt tattatagaa tagttttaa tgttattgag gagggataa ttgtaaaaat 15660
taaataattt ttgttattt ttttgaaggt ggtattttga ttttagatg gttttttaat 15720
ataggttttt ttttttttta ttattatagt tgttttgaat tttgagttgg aagggaatat 15780
tttgagattt agattgttaa atgttttttt taaagttagt taataaatta aatggtaaag 15840
ttagggtagt tttttgattt agtatagggg attttttttt attttttatt tttgagattt 15900
tagaattggt ggtattgttt taaaatttat ggtaagaatt ggttattttt gtaattaata 15960
tttttttata atatatgtt ttgttttgtt tagtttagta gaaattatat ggagtttgtg 16020
ttttaaaaag ttgttgaag tttttatttt ttgttttggg attatgtgta tgaattatta 16080
attggttttt tttttattta tatttgatga agatgttttt tttttaatat ttttttttat 16140
tgttttttat ttttttttgt tttatttatg attagtgttt tgtttttaa tagattttgt 16200
ggttatttat ttttttttgt gttagttttt atttattagt tatttgaatt gtggttttta 16260
ttgtttttta tatagtttat ttatttggg ttgttatata tatttatatt gttatatggt 16320
tttaaattgt attttggat gattttggta atgggtgtat tggatgagat ttaaattaat 16380
aattaaagta ttgagatagt tttgttatt ataagtatt tttgttttt atagtttaag 16440
aggagttatt tttttattt ttattattta atgtttaata ttatttttat tataataat 16500
gtataaaaag tatgtgattt atgatttatt ttaaatttga atgtttgtga tttattttgt 16560
gttttttttt atttttagg 16579

```

&lt;210&gt; SEQ ID NO 19

&lt;211&gt; LENGTH: 16579

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 19

```

tttagaaatg aagaaaaata taaagtaaat tataagtatt taggtttaaa gtgggttata 60
aattatatgt tttttgtgta ttgtatgtaa taagggttagt attaagtatt aagtaataag 120
agtagaagaa gtagtttttt ttggattata aaaataggag taaatttga atagtagaag 180
ttattttaat gttttaatta ttagtttgaa ttttatttaa tgtagttatt attaaaatta 240
ttttagaata tagtttggga atatataata atgtggatgt gtgtgataat tataagtgag 300
taggttgtat gaagaataat agagattata gtttaagtaa ttagtggata ggagttggtta 360
taaaaagaaa tgggtgatta taaagtttat ttagaatag atagttgatt ataaatagag 420
taagagaaga tggggagtaa tagagaaaga tgttgaaga aaggattttt tattaagtgt 480
aaggtagaag ggaattaatt ggtggtttat gtatataata ttaaaataga gaataaggat 540

```

-continued

---

tttgtaggt ttttaaaagt atagatttta tgtagttttt agttgggttaa gtaataaaaa	600
taaatgtatt ataaggaggt gttaattata aaagtgatta gtttttgta tgaattttta	660
agtagtatta atagttttaa agttttaaga gtaaagaatg aaagaaaata ttttgattg	720
agttaagaaa ttgttttggg tttgttattt aatttattgt atgattttgg aaaagatatt	780
taataatttg ggtttttagaa tgtttttttt tagttttaa ttaaggtag ttataatgat	840
gaggaggaga agaatttttg ttggggaatt atttagaaat taaagtgtta attttaaaag	900
aatggtaaaa gattattttg tttttatagt tgtgtttttt ttggtagtat ttagaattat	960
tttataatag ttatgtgttt ttttttttta atataatttt ggttttattt tagtgtttat	1020
ttttagtat atattttgga gaattttggg ttaaatttta gtttaaggga tgattttgat	1080
attttaagtt aattatagag gtattatttt tttttttaat gattggattt ggaatagat	1140
ataatatagt tttggttagt gagatgtgag gggatttttt tttttagtga agggtttttg	1200
ggaagatag atttattttg agaagagaat aaaggaagag gtggttagtt ttttttttg	1260
gatattatta tgttggaaatt tttttgttg aattgttga gggattttgt atttatgagt	1320
aagtgagata tttttgaaa tgaagtgtgag aaagaattag taaaggtaag gtgtaagttt	1380
ttaaggatta gatttagagt ggattatggt tttgaatttt attttatag agattataaa	1440
tgttattatt gtttaaaatt aataaaattg tgattttttg ttttttatag ttgaaagtat	1500
tttaattggt atataatttt tagttgggat tttgaaaaat tttattgga aagaaatttt	1560
atattattta gtaaatttag attgaaaata agttaatggt gaaattagtt ataaattata	1620
aaattgattt tgaataataa tagggtttg agttgtatag tatttaatta attaagattt	1680
tttttttaa ttttttgta tttgattttt ttatatattt tgaataaga attagatttt	1740
tttaaaataa aatttttaa ttttataatt ttagatattt tttgaagagt attttttaag	1800
ttaatatatt atataattta tattttttat ttatagatta attgataatt tataatgaaa	1860
aatatagtt gataatttta aattaatttt attattagtt tgtatttatt ttatattatt	1920
agaattgtta tgatttgta attttattgg aaagtttatg ttttagtaga taattaggtt	1980
ggattaatta aatagaaatt tttttatgg tatgaatttg ggaggtggag gttgtagtga	2040
gttgagattg tgtttgggtg atagagtgag attttgtttt aaaaaaaaa aaaaaaaga	2100
aattaaagat ttttttttaa ttatatattt aaataataat tttttatttt tttttaggaa	2160
aatttgtttt ttatttttta tttttttttt tttttaggaa atttttataa gtttaatttt	2220
aaaagtttg gttttattgt taaatagttt agaatgtaag ttttgatttt tttatttata	2280
ttttgtgatt ttgggtaaaa tattatattt ttttaattt tattttttgt tattatttag	2340
ttgtgaaaat tagtatttat tgaagtgatt ataggaattg ataaaataat gtatataaaa	2400
taaattttag tgatatataa tatattagta tttaatattt attattaaaa ttattaattt	2460
agattttttg tagatgtttt aatatagttt tatattatta agagtataaa gaatagattt	2520
tggggttttt atatttgaat attataattt aggttattat tatagtagta atatatattt	2580
tatgtttgga ataaattatt ttataattta ttgttatgta ttttaatta tttgaaata	2640
tgttaattt tttttgtaga atagtaggtt tttggaaggt agagtttagta tttattttaa	2700
tttaataatt tttataatatt ttttttattg ttaggtttag ggtatatttt tgaggaatgt	2760
ttgtagatta aatatatgat taatgttttt attttttatg tagaaatatt tttgaaattt	2820

-continued

---

tgtgtaata attttttaat atttgaaata tttttgaata tttttgaatt agtttatttg 2880  
aagatttatt atgatttaat aatatgatgt ttgttaatat ttagagagtt aaattttatt 2940  
taaaaaaatt aaaggtgtga tttaggaaag gattttttgt atttgagttt tttatttaat 3000  
ttaaataatta aaaattgaat tttttgtggt atatttaatt tttataaaag tgttttttgt 3060  
tggggagagg agaatggttt tttatttata taggtataat aagtatatgt attatttttt 3120  
ttaaagattat ttttaagatt aggtaagaat ataaggagga gattagatat gttttaaata 3180  
tatttgaata agttgtataa atagtgttgata ataagttatt taatttaaaa attagattgt 3240  
ttatattaaa tataaaattt ttaatttttt ttaaaaaatt agaaggttta gtaatattaa 3300  
gtttatattt tagttgggtg attatgagta ttaattaagt agtagttttt ttttttaagt 3360  
ttgtgtgttt aaggttatta ttgtttttat ttagtttttt atttattttt attttatgtt 3420  
ttgtttttgt agttatttaa gtttgtaaat tttgaattaa aagatgatta tttattttta 3480  
tagtaaatat attagagggt tttattgatt ttagttgttt tagaattgta tttttatttt 3540  
tttttttttg gagtttaaga tgtttatttt atttttgtaa ttttttaag gttatgtttt 3600  
tatttttggg tattaagtta tattttttgt tttttatggt tgattatatt aaatggaaag 3660  
tttttagaag tttttttttt tgttttttagt attattttta atttttattt ttattttttt 3720  
atttttttat gtttttttta taagtattgt ttattagtga taataatttt gtttgtatgg 3780  
aaagatttga gagatatgag agtgaatggt tatatttagt tttttatttt agtgaagtta 3840  
atgagaaggg aagaaagtta ttgaataatt taggtatagg gtattgagat tgtttaattg 3900  
tttgggagat ataaaaata aggttaatgg aattttttatt tatgagaaat gttttagggg 3960  
aggataaggg atatatagg gatgggtagt ttttgagaga aagttaagat tttattaaat 4020  
agatatttag tatttttttt tgattttggg gaagtttagt tgataagggg tgatttaagg 4080  
atagagatat ggtgggtttg tttattattt tgttaattag gtattttaga taattaattt 4140  
tttaggatat tttagtata aagtggatgt aatgagtatt atatatata ttgtttattg 4200  
gttttttttt atagaagaa tttattgata gtagagagaa attatgtgat ttaggtgtga 4260  
aattttatga atttatttta tataagtaat agtattatgt gaatttttga atttggattt 4320  
ggttaagttt ttttatattg ttgtttttta gtattttttt gttttttggt tgggtgaattt 4380  
tttttttatt atttatttta agatatattg tttttattt ttttttttat gttttatttt 4440  
tatttttagta tatattatta ttttgatat gtgattggga attgtttttt tttatttatt 4500  
atagtaaaag ttttaataga gtagtttttg atgttttggt tttttttgta tttataatgt 4560  
ttagaatatt tagtaaatat ttgttaaaag aatgaatata tttttatttt atataattag 4620  
atgtataata tatatttagt tttatttatt aattttttta tttagaatat aatataattt 4680  
ggagaatggt ttaaatttagt aaatataatt tttttatggt atttttgata gttgtatagt 4740  
tttttattag gtggatatag taatttttat ttaaattgggt ttttatgatg agtttgttat 4800  
attgttttta gatatttgtt tttagaagta atgtttgtaa gaatattttt gtttatttgt 4860  
tttgggtgat attggtatta atatttattt atataataat gttgagaagt ggaattgttg 4920  
aattaaagag tatatatatt ttttaaaaa ttttgataaa gaattaaaa aaagtgattg 4980  
ttaaattatt ttttgaagta gaggaattta tttattttta tataatagta tatgagagat 5040  
ttagtttttt tataatttgg ttaatatata tttgtattaa attttttaatt tttatttttg 5100

-continued

---

ttaggtatag tggtttatgt ttgtaatfff agtattttgg gaggttgagg tgggtgaatt	5160
atftgagggt aggagtttaa gattaattgt tatggttaat atggtgaaat tttattttat	5220
taaaaatag aaaattagtt aagtatggtg gtgggtgttt ataattttag ttattagga	5280
ggttgatata ggagaattgt ttgaatttgg aagggtggagg ttgtagttag ttgagattgt	5340
gttattgtat ttttaattga gtaatagatt ttattttaaa aaaaaaaaaa aagaaaaaaaa	5400
ttattttttt ttgtttttt tttgatttgg gaaaaaatgg tgtagtatat tatttgggtt	5460
ttaattatta tttatttaat tatagatgag gttgagttag tttttttata attttaaaat	5520
tttatgtttt tttttgttat tatttgtata taatttttgt ttatattttg atagatgtgg	5580
atftttttaaa gttatttata aaagtttatg gaatattaag agttattagt tttttgtttt	5640
ttaaatgttg tagatgtttt ttaaatatat tttttgtttt tgtgtaattt ttatttagtt	5700
agatttatta gtatttttta gaatggattt gggttttatg ttttgttttag aataggtttt	5760
tttattttta atttattttt ttttaattta tgattttttt ttattatttt gtaaaaagt	5820
atftttttata ttttatattt gatttatttt atattttttt taggatgaga aattaggtag	5880
agattgattt ttatttttga tggatttatt aattgttttg aaattattta ttgaataatt	5940
tttttttttt atgaaatgtt taattttatt ggagtttaatt tttgtaattt gttttttgat	6000
taatgaaatt aataatfttt taaaataata tgtttttttt tgatttattt tatagattta	6060
gtagaaatta tatggttttg gtaatgtagt tattattaaa attatagaat tagattaata	6120
tgaattttta gatataatat attattattt tttttttgta tagtttttta aagtatttat	6180
ttttttgatt ttataaattg tataattatt taatagatta tgatattaaa ggtttaaaag	6240
taatattttt gtttggtttt tttagttagt gagtttaatt taaatttttt gatttttaagt	6300
ttagtgttta ttttattata ttttaagtaa tattaaatta ataatttaat aataaaaaga	6360
ttataattta ttaaatatgt ttattgtaaa atatgtaatt ttattttata gatataaata	6420
ttttttaaaa tttttagaag aatgaggtag gatgtatata gatgtggaaa gtatatatat	6480
ttgtattttt ttatagtttg taaatttatt tagaggtgtg ttttaggaaa tagaattgat	6540
atattaaata aagatttttt gtattgattt ttgaaatata attattttta atattttta	6600
tggtaagaat tttagtagtg ttaaaaattt tgtatatata tatatatata tgtattttat	6660
atatatatat agattttata tatatgtaat ttattaagtt taggttataa gaggaaattg	6720
gtaaaaattt attaatgaa atggagtttg gtttaagatat ttatattta tttttttttt	6780
ttgattttta attagggaga gtatagtgtt tgtgtttttt ggtaaaatat agatttatgt	6840
aataattagt atattttttt agattagagt aagtaatttt taataatgag tgtttgttag	6900
ttttaatttg tttttttatg tgggtgtgta aaaaatattt tattatagtt attaatfttt	6960
ttttttaaaa gaattatgtg ttttgagaag ttttgtttta ttataggta atattttatt	7020
ttaaatatag tgtttattga tgtttgttta tgaatgatgt taataaaaat gtttaaaatt	7080
tattaggaaa ggagggttta gttgttagat tttgggagat agagttagtt gatttttgtg	7140
tttttaaagg ttttgttatt taattaagtg attaattatt aattttaaag tatttaagtt	7200
taataatata ataaaagaga gaataatggg tagtattttt aatattaatt atggttattt	7260
ttggtttagt tgtttaggtt tttttttagt agaatttata agtgtataga tataggattt	7320
ttggaatatt aaaggatttt ttttttagtt tgtaaatata gggaaaaaga agaaatttat	7380

-continued

---

atgtgaaat ttatgtaaat ttaggggat tttttttaa tttgttttt ttgaaagtag	7440
gaaatataat ttaaaattaa tttgttttt atttattaaa agaattaaaa tttgaattt	7500
tgatataaat ttaaaatgat tgttttgaat tttattttgt tttttattta atttttttat	7560
aagaagatta aggattgtgg atttggtgtt ggaggaagta attgtttaat tatattggaa	7620
tttattggaa ttttttataa tattattgtg ggtagataat ataatttaa tatatttttt	7680
ttttattta gagtttagaa taagaagtta gatttttga gatttttttt ttgatgtatt	7740
gtttttaata ttttaagtta ataaattggg aattgttttt attttagaat attaattgtt	7800
gtttttattg gtattttttt atttttatat tatatttata atattttatt ttaggatttt	7860
atgtgaatat ttttaggatg ggtttggtga aattgttaag aagtgttaa aataattagg	7920
ttatagtttt tagtagataa aaagaaaaag atgtggttgt tagttataga atttgaggtg	7980
atggaaaaga tttttaaaa ttagtttttt taatatgtag taaaaaggga ggatttaaaa	8040
tgtagttta agaagtattg tttgtttgt attaagtttt tttaggtttt tagtttagtt	8100
atttttaaa taggtgttga gtttttagtt tttgttaggt tttgttaag gttttttatt	8160
gaggatagg aattttagtt atattttagt tttttttatg tttgtttta tgaattttta	8220
ttagttttgt tgaatgtag tatttgttat tttaggtaa tttttatga tagtttgaa	8280
tgaattaat ttgttaattg attaagaatg ggttttagag tgatttgaat gtatttaaaa	8340
agttaataaa tggaaatggt tataaagttt ttttttttg ataaaaagt tgtttggagt	8400
aaatttaaat ttttttttg ttttaaaaa aaagtatgat gttttttta ttattttta	8460
gaagtaagta aaaataaaa aaaataaaa aaaatttaa tagttttta ataagagtgt	8520
ttagtgttgt ttggatggaa tttgggttag ttttagaagt agttggtttt gaattttggg	8580
ggtgaggta tttggttttt ttggttaagt tgggtttgt tttttttgt ttgtttttat	8640
tggttttggg ggtgggggt tttggtgtg ttggtgttt gtgtggtgt attgtagttt	8700
ttgagtttag gaagtatat ttgttttggg tgggagtga aggttaatt tatgtggtgg	8760
ttggttaatt tagttgttgt ggggagtgt tgttgtgggt tttttttgg tgtgaggta	8820
gtgagtgagg gtgggtttta gatttagttt ttttgggtgg ggtggatta ttttgtttg	8880
tttagttttt ttttattttt ttattttggt atgtttttgt agttatattt gttgtgttg	8940
ggaaaggaa aagagtattt taaaatgagt tttgaaatt agatttaaat tttgaaatga	9000
atttttttgg tttgtttttg gttgtttgaa tgttatttgt tttatttttg ttttttttt	9060
ttttttttt tttttatggt ttttttttt tttatttttt ttttttttt ttttttttt	9120
tttttattgt ttgttttttt tttatttttt tgtttgatgt ttgtttttaa gttgaggaag	9180
tttaagtttt ggattaggag ttttttttt gtaaatattt ggggtgggt tttattgttg	9240
ggtttaagtg gttatgggtt ttgagagtat ggtatgtgta ttatttttt taaagatgtt	9300
taggtggtg tgtttagatt tttgaaattt ggttttattt tggttatgtg tgtttgtta	9360
tttaagtgtg aattggattt aggtgtttgt gtgtgtttta gttgattgt gagaaggat	9420
ttttgtatt tgtatagatt aagatttttt tttagggttt tttaaatttt tagatgtatg	9480
ttgtagattg tggagttaa aatgaatgat ttaagggaaa attagggttt tttatttttg	9540
aatataaggt tgtgaaagat gtatgtttgg tttgggggt ttgttttttt ttgtttgtgt	9600
gggatttttt tattaggggt tgtttatatt tgttgtaaaa ttttaatttt ttagatattt	9660

---

-continued

---

aggttttgta agttttgtga gttgaattgg gggtagattt gaaggagtgt attgtttttt 9720  
ttgttttttt tttttgttg ttttgagta tgggaaaggt agggatttta tattagtatt 9780  
ttatttgta gttgtggggg aggggatgta gaagttagag agagagaata gaggttggga 9840  
gaggggaagg gaggtagaga agatggagat agaggagat ataaagatga ggaatgtgat 9900  
ttagagagat tttgttggg gtgtgtgtgt ggtaggataa ttttttaggt gttgtgtaga 9960  
tttttttttt ttttttagt attttttagg gatatgtagg atttgttgg atgtagttgg 10020  
attggagagt aaatagttta ggttatgata ttattttttt atttgtttt tttttgttta 10080  
atatttagat tgttttatgt ttagtgatat agattgtggg ttagaataat ttggatgttt 10140  
taggtttgta tgtattagg atgttgtttt atagtgttta tgtttttgta ggatggggg 10200  
gtgttttagt tgaggatgtt ttttggggtt tttggtatat gtgtgtgttg ttgtttag 10260  
atltgttagt aaagtgttt aattagagtg gagaattagg agagggtggg gaatatgtaa 10320  
ggaaatggag gatttttga ttttgggtga gaggattgat ttagaggttt ttgatttggg 10380  
tttttagatt atttttgttt gtagtgtgtt ttgggtgttg gtttgtttgt tgagatggga 10440  
gtgatagatt ttattttatt ttttaggggt agaaaaattg ttgtatattt ttgttgattt 10500  
agaatagggt tttttgttt aggggttttt tagatttttg tggggtttta ataattttat 10560  
ttttattttt ttgttttttt tttgttgggt tttttgtatt gtgtgttagga taaaagaatt 10620  
taggggtgtt gggagagaaat tgtgatagtt ttaagttgt taaagagata tgggtgggat 10680  
tttgattttt ttttggttgg attttttttt agattgtttg tggtttgaag gggttagggt 10740  
tgtgtaaggt tttaggttgt ttgggtgttg ggttgtattg ttaggagttg tttgggagtg 10800  
gttttttagg aatattagg attgattata tattaggggt tgggaaatta ggtggtttgt 10860  
attaaggtgg ttttggggat ttgtttgtgg taagttttgg tagtttttat ttaaattttt 10920  
gttttgagtg tgttaagaat aataataata aaaaaattaa agtgttaaag gttttttttt 10980  
tttttttagt ttaagaattt attatttttt tatgattttt ttataattta tttttttttt 11040  
ttttttaatt ttgttagtta ttttattttt attttatttt gggttttttt tgtttgaatt 11100  
tttttaata ttaaggtttt tttgtatggt tttttttaa agtttttatg aaagttattt 11160  
gtatttttta agtgtttata tttttttaa tttgttttat agttttttgt ttaattaaa 11220  
gttttttatt aattgttttt tttttttaa tttgtgggtt ttttttata agttttttgt 11280  
ttttgttttt taagggggga ataaaagaaa tgtgattatt ttggaagggt gtttattgta 11340  
gtttgggggg aaaatttatt gtagtgttgt gtgattgggt ttggtgttgt ttaggtgggt 11400  
tatataggaa gtgtgggtgt ttggggaagg atgtggaggg tgtgggatgg tgttgaaga 11460  
tgtgggagga tgtgggggtt gttgaagatt ttggtttttag tttttgttta taatatttaa 11520  
tgtttttggt tgttgtttt gtatgttttg agtttttttt tagaaagggt tgggtagtt 11580  
tgtttgtaag ttaaatgtg ggttgtgata ttattatta ttatattatt gtattgttag 11640  
agttgaggag gagatttagt gagaagaagg aggagggaga ggaggagggt ttatttatag 11700  
gttttaaaag tgtttttag ttttagttat ttttaagagt ttaggttttg aaagtagtg 11760  
gaggggtgga aagtagttt tttgtgtttg tggtagggat tgtttgtttt ttttttagg 11820  
tatttttatt ttttgggtt tttttttgat aagaagtatt aattggtttg gggatagtag 11880  
tgttttttat ttagggattt atttagtaat attgttttgt ttgtggaat tattgttttt 11940

-continued

---

ttatTTTTTT tttatTTTTa tttttttggt ttttTgttag ttagttttgg gttaggggaa 12000  
aggagtTTTT aggtTTTTag ggggttaggt agtaatagat aattgagtat gattatTTTT 12060  
ttttgggagt atataaaatt gtaaaattag taaagaattt ggttatagtg tgtttatgtg 12120  
tttatagagt ttgtttgttt tattaagggt aagtgttagg ttaaggttt ttgttaattt 12180  
gaaagagata ttgagaaaat gagatTTTTt ggggatttag agggaaaagt taagaatttt 12240  
ttattgtatt tttagggaaat tgtttaatgg ggagtttggt tttaaaagat tttggtaata 12300  
aaaggttgga taggaaattt ttttaggtaa attttttgtt ggatttaaag agaataattt 12360  
ttttttgta taaatTTTTt tttatataag tttagatttt gttttttttg tttttttttt 12420  
tttttagttt ttttaagtat ttttgagttag aatatttgat aatttttttg agtaaataggg 12480  
attttttggg agtattaatt attttttatg ttttttggaa ataagattat aattttttat 12540  
gtgtatatgt gatttttttt ttttagtttag gtttatttta ttgtgtaaat agtattaata 12600  
tatggaagag ttttttTgat tgtgttataa aagattttta ataggatttt atagagaaaa 12660  
gggttaaata gttgatataa aggatttttg tttttttaga aaagagggat tttggatttt 12720  
ttttttttt gaagttaagt atgagtttat ataataaggaa taaaataaat ttaagggtga 12780  
tattagtata atattaggga tattagaatg gatggtaaat tttttttttt atataaatat 12840  
gaaagtattt ttataatttt tttttttgaa gtttttattt agaaaattat attaaatata 12900  
ggatttttaa atagtagtga ttaaagatga aagttaattt ttttttttaa tttttttgat 12960  
tagtttggtt taaatTTatg ttttaaaatt ttttagtaatt tagatttttg tatatgtgta 13020  
tttataagat tttttttata ttttgtaatt tgtaggtatt tagtatgggt attgatttta 13080  
agtgataaat aggtagaaga tttgttagga taatttttagt ttttattatg tgtatttttt 13140  
attttttttg ttatttatta gttttttgta gtaataaat tgttagttat attatttggg 13200  
aattaatggt tttttgtagt gaattattta ttttttagtaa atttattgta ttattatttt 13260  
aaatattgaa atttgatttt attagatggt tagaatgata gttatagagt agaattagat 13320  
tttaattttt atgtataata tgtaaaataa ttaagggtga gaaagttatt gttttaaagt 13380  
tatataatta atagtttagt ggttttttta gtgtaatta tagtaatttt agatttaaga 13440  
ttgtttgatt taggaatgga gagaataaa ataaaatgaa tgtttattga aatataatatt 13500  
ttagatatta agatagattt ttttTgtgaa ttttttaaga tgtggatttg tatttttatt 13560  
tttaaatTTg agtaaaattt gttgagtttt aaagagttta agtaatttgt ttaaggttat 13620  
tttgtaagtg atagattttg tagtttgaaa agtttattat tttttatttt tttgtttttt 13680  
tatgatatag ttaagtaaaa ttttttaata aaggagaatt agattagaaa taagtaatta 13740  
atagtagatt aatgtttatt tttagggaaa tgttaatagt taatatttat tataaataat 13800  
gttatttTga tataatggaa tatgattgaa agtgatgggt tagatatttt ttatattttg 13860  
agatgttttt atttatggtt ttaaaattga ttttttaatt ttttaaaatt ttttagttttt 13920  
tttaatgTta atagatttta taatatttga ttaaattttt ttataggata aaggattttt 13980  
tttagggTtg atttaaggaa ttagaatatg tataatgTtg aaaagtttta aattgTttgt 14040  
tttttttttt ttttgatttg taatattatt atttagaaaa ttaaattttt ttttaggtta 14100  
ttagataaatt aaaatgaaat tatattttat ggatattttt tttttatttt tagggtagat 14160  
tttttttttt agttgtttta aatggagaat ggataaaaaa gaaaaaggaa aatatattag 14220

-continued

---

ttatatttag aatagagttt atagttgtat tattatTTTT atgtaattgt ttttaatgat 14280  
TTTTTTTTat taagagaaaag aaaatgagag agtatttatt tttgattttt taatgtttga 14340  
taaaattatt agaagttttt gaatgggaaa atatatattt tttaggtttt ataattaa 14400  
ttgtatttat ttgtttatta tatttaagta tgaattaaat agtaattaa atattatTTT 14460  
gtgagttttt ttttaaaata tatttttgta ttggggatat tttttgtgaa aattttaaga 14520  
ttttaggat aaataaagta agatttttat ttataatggt tgggttttaa agaaattata 14580  
agttagaagt tatgtgtttt gtttttttgt tttgattttg tttttgggaa tttggttgag 14640  
tttaggggta gtaagttaat gatttgaga ggaaagggtg agatggtaga tgtgttttta 14700  
gggaatttta ttatttttat ttaaattttt ttttgagggt ttttaggatt atattttaga 14760  
ttttttttt ttatttgatt aggtgaagtg tttttagata tgggtattatt tattgttggt 14820  
ttagggatga taagatttta gttttattta aatttagaaa gagtttagtt ttaatttttt 14880  
gatttgattg gttttaggta taattaagta aattatTTT attggatgat atagtaatgt 14940  
tatagattaa aaatgtattt ttaattttta aagattttgg aaaaaggaag tttatTTT 15000  
aagattgtgt tataattttt tttttaaata aaatattgta aattagggat agatatttta 15060  
tttagtattt tgaattttt ttggttttt aggaaaggta gaaatataag gtttttattg 15120  
ttttttttt attttaaaa aagaaatttt gtaaaaggta aagtgatggt ggtaatttta 15180  
ttgtgtttta atttgaagta tataaattgt ttgataaat taaatataga taaattttt 15240  
tagaaagtta taaatatata gattttttt agatattaga tattgagttt ttgttttta 15300  
atTTTaaat taaatgtta tagtaaaata tttgtttta taaattttt tttttttgt 15360  
gttgtaaag tttgtttat ggttgagtag atttttttag aatttaattg aaagattatt 15420  
tatattatga tagaattatg taagttaaag ttaagtttta atagggagtg atgggagagg 15480  
ttgggtggt gagaggagat gattttatta gtgagaagag aggtttttt agttggttt 15540  
taaggtagtt aaatggttta gtatattgt attatTTT ggaggttgt tggagtagag 15600  
tattaaaaat gttattttt ttataaagaa aaagatgaag gtataaagaa tatagttaa 15660  
ttaaattgaa ttaagattg tattatTTT atttttttt tagttattt tttaaagtt 15720  
ttttttttt tttttatagt ttgaaaaata taggttattt ttataaagta tttgtttgat 15780  
tgaagtaaga taaattgaag agtttggtt ttgaatataa ggaagtggta gagattttag 15840  
aattaatgag gtgggtgga atataatgta gaaggtttt taagaggtaa aatttttga 15900  
gttggtggtt atttttagtg attatTTT tttatttatt taattagatt atagtgggat 15960  
tgagataagg agtataatga aagttgtgt attgatggtg taaaaagata tatggtatta 16020  
ttaatgtttt tgtggtgag aaaggtagaa tgtaagatag aaaataagga aatttttgtg 16080  
ttgtaagaaa agtgtttagt atgggtatta attggtatgt agaagtagta agtatatatt 16140  
ttatTTTat tttattgata aaaagaattg taggattaat gttgttttt attaaaaat 16200  
gagaaattat atTTTaaatt tataaattaa aggttaata tttgttgtaa taattttaaa 16260  
aaaaagagaa gatttgtaag atttagtaaa tatttattat aaatagattg ttttaattt 16320  
tttatattt ttttaattt aataaaagt tattgagtag ttttatgtg tattgtttta 16380  
gatagtaggt attattttaaa aaagtaataa ttatttttta ttttaagga gtttataatg 16440  
tttttagaaa gaaattaatt tattggagta gttaaagtta taattgattt ataaaaatgat 16500

-continued

---

```

tagtgaattt agaggaaaat tatatataaa tatttttggtt tttattgggt ggtttagtta 16560
gatattattt taaaaaata                                     16579

```

```

<210> SEQ ID NO 20
<211> LENGTH: 3000
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

<400> SEQUENCE: 20

```

```

atatttaaat tgtatatttg gtttgaata tattatata ttaaagaat tattattttt    60
ttgaatatta ttaatattaa taatagtatt atagtgatta agagtggatt taaataaagg    120
ttggtttgaa atttaggttt gttatttatt agtaagtttt taaaaatttt gagtttttag    180
tttttttata tgtgaaatgg agaaaataaa tttgtaaaat taattaaatt tgtttattta    240
tatttttttt tttagtgtat attaattggt attttttggtt aggttagaat gtgtttttta    300
ttatgtttta aatttgtttt gtgttaattt tattgttaga atttttttta tttgagaat    360
tagaaaagga aatattatgt ttggtaatgt ttatattttt taaaataaat ttgtaggaaa    420
gaatatttag taagtgatga gagtagtaat gattgttttt attattttta atttataata    480
ttattttttt tagagttttt taagtattgt agataatttt ttatttatta aaaaataaat    540
tgaattata agtatttagg gttgatattg tttttgaatt agatagtgtt tatattagtt    600
gtataagatt aatttaaagt agaggatgaa attttttttt tgaatttttt ttagaatgta    660
atttagttaa tatattaaa ttaaattttt tttgaatggg agtaattttt atggattaat    720
ttgtaatttt ttagattata ttaaggtaa tgtagagggt gttgtattat aggtttttgtg    780
attagagatt attggatttg tttttggaaa agttatttat gattttttgg gttttggttt    840
ttttattttt tatattgggt taataatgat tattgtagag ttgttatgga agttaaataga    900
atztatgtat tataagtgat taatataatg attgatattt agtgatttat taataaatta    960
aaatatttat tattaatatg atagagaagg tgttgtaaa atagataata ggtttttgga    1020
agaggtgatt aaatggatgt aaaaattatg gattgtttat tttgtttatt tttgttgtgt    1080
tttttggttg tggatataat atgtgtgggt ataaaattgt aaattttatg tagttgtgta    1140
gtgtatgtgt agaaggttta gatatgaaat gttatttttag taatgtgttt agagaagttt    1200
tgatgtgtgt ttggaagtaa gttgtgtgtt tttgattttt ggggtgtttg gatggatgtt    1260
tatatttga tttagtagta ttggaagggg tttaggtttt ttgtagtata gtttattttt    1320
agattgttta gttttttata atatataatt ttatggaaaa gggatttttt tttgttaga    1380
aaaagtgttt tagtttggtt tgggttggtt tttattttat gttgtgttaa gtaggtgaag    1440
ttttttttgt tttttttttt ggggtaagtg gaaaggagtt tggtaggggg tttgtagtgg    1500
tttगतatagg ggaattgggt agtgagagag ttttaggtaa ttttgggggt tgttttatag    1560
aagttagtgg ggattgatag tggttttttg gtttagggag gagagtgtgg tttggtgttt    1620
tttttttttag tttggaggtt gtagttgttt gagttggttt ggggtgggggt ggggtggggg    1680
tgggtgggag ggtatggaga ttatgggtgt gttatttggg atatttaggg ttttagggtt    1740
ttgggtggtt tttatgtgag attgtaaatt atgataatag gtagttattt gaggttaaat    1800
aaaaatggag tgggtttttt gtgtgtgtgt gttttttgtg tttttggtgg ttttttttga    1860

```

-continued

---

```

ggtttttggg ggttttatga gttttagta gttggtggg atgttgtttt tgttttattt 1920
ttttgtgtaa ggtgaggggt gttggtagtg tgggtgatgt tttggttgtt tttggttttt 1980
gtgtaaaatt tttattttgt ttatgtgaag ttgttgttgt tttagagagg gggaaagagt 2040
tgtgggaaaa gttggggagt gatgattgtg gtggttgggt gtgttttttt attttttttt 2100
tttttttttt tttttttgtt gtagtttggg gttttggttt tttttttttt tttttttttt 2160
ttggagttgg tttttttttt tgttttgttt tttttttgtt tgtgtatggt atttgttgtg 2220
gggtggttga aggggatggt ttgtttttat tagaggtata gtgtgaaggg gaaattttga 2280
tattggaagg aatgagaata aatatttaat tatggatgta ttgaattgtg gttgggatag 2340
atattttggg aatttgaggt ggattgggtg atgaggtgag tgattttttt ttttaatttt 2400
tgttttaggg tttttggggg agtttgagtt gagagaattt ttaaattttt tgggaaagtg 2460
tgtgaggttt tgttggggat gttgaggttt gggatttgag gatgtgtagt tggatgggtg 2520
gtgggtgttt gtgttttttg ggggtgtttg gaggttgggt gttttatggt tgagggtttg 2580
ggttgtttgg attgtagtgg tgtttttgtt tttagaagat gtttttaagt ttaagggtt 2640
ttttttgagt ttgtttgttt tttttgggtt tgggttggag tttgtgtgta atggagtta 2700
tttagtagtt tagtgtgtgg tttttatttg tattttgttt ttatttggtg gaggtgtgag 2760
tattgggggt ttttttatat tttttttatg atgtgtatta ttttttgatg atttttttaga 2820
tggtttaggt gtgaggatgt tgatttagag ttttttgag ggttataggt gtttgggttt 2880
ttttggtgtt ggggtgtgtg gtattttaaa ggtttgtgtt ttaattttta ggtattgatt 2940
gggtttttta attgtggtga ttttatttta atagttttta tgtggtgtgg attgaaatgt 3000

```

&lt;210&gt; SEQ ID NO 21

&lt;211&gt; LENGTH: 3000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 21

```

gatatttagt ttatgttata taaaaattat taaagtggga ttgtttagt tgaagagttt 60
gattagtggt tggagattag aatgtgagtt tttaaagtat atatgtattt ggtattggga 120
aagtttaggt gtttgtgatt ttttgaagga ttttgggtta gtatttttgt gtttggatta 180
tttagggggt tattagaaag taatatatgt tataagaaag atgtggggga gattttgatg 240
tttgtgtttt tgttaggtgg aggtgggggt taggtagaag ttgtgtgttg gattgttggg 300
tgaattttgt tatgttagg ttttgtgttg attttggag ggataggtag gtttggagg 360
gatttttggg gtttggggat gttttttagg gtagagagta ttgttgtggt ttgagtggtt 420
tgggttttta ggtgtggggg atttgggttt taagtgtttt ttggggatgt aggtgtttat 480
gtattgttta gttgtgtgtt tttagtattt agtgtttggt gtttttgggt gagttttgtg 540
tatttttttg gaaagtttgg ggggtttttt aatttaggtt tttttgggag ttttgggggtg 600
ggggttggaa gaaggggtta tttattttgt tgtttggttt gttttgggtt tttgaagtgt 660
ttgttttagt tgtggttag tgtgtttgta attaagtatt tatttttgtt ttttttagtg 720
ttgaagtttt ttttttgtgt tgtgtttttg gtgaaaaatg gatatttttt ttggttattt 780
tataataaat agtgtatata agtggggggg aagtgggggt gagggagaag ttggttttga 840

```

---

-continued

---

gggggaggag gaaaggagag gagttaaataa tttggattgt gatagggggg aaaggagaag 900  
aaaagaaaat gagagagtgt gtttagttgt tgtagttgtt attttttggg tttttttgta 960  
gttttttttt tttttttaag gtagtgataa ttttatgtgg ataggatgga agttttgtgt 1020  
ggaagtggg aatggttga gtgtgtgtt tgtgtgtgtt agttttgtat ttgtgtaggg 1080  
agggtgggtg ggggtgatgt tgttattggt tattgtgggt ttgtgagggt gttgggggtt 1140  
ttggggagg ttgttaggga tgtgggggt ggtggtgtgt gggggattta tttgttttt 1200  
atttgatttt ggggtgattt ttattgttat ggtttgtgat tttgtgtggg gattgtttag 1260  
ggttttggg ttttgatgt tttgggtggt gttgtgttaa tttttgtgtt tttgtgtgtg 1320  
tttttatttt gttttatttt ggggtgattt gagtggtgtt agtttttagg ttgaggggag 1380  
ggatttgtga ttgtgttttt ttttttgggt tggagagtta ttgttgattt ttatttgttt 1440  
ttgtgggta gtttttggaa ttgtttggaa tttttttgtt atttagtttt tttgtgtagg 1500  
ttattgtggg ttttttgttg gatttttttt tatttttttt aagggaggag atagaaggga 1560  
ttttgtttat ttgaataat gtgaaataaa aattaattta gattagattg ggggtttttt 1620  
tttgatggga gaaatatttt ttttttggg agatatatgt tatgaaggat taagtgggtt 1680  
gggataggt tgtgtgtga agggtttggg ttttttttag tgtgttggg tgtaggata 1740  
ggtatttgtt ttgatgttt aaaggttagg tagtaatgat ttatttttaa ggtggtgta 1800  
gagttttttt aggtatattt ttgaaatgat attttgtgtt taagtttttt gtgtatgtat 1860  
tatgtgatta tataggattt atgattttat atttatatgt gtgtatgta taattagggg 1920  
atatagtaaa ggtagatgga ataaataatt tataaatttt gtatttattt aattattttt 1980  
tttaaaaatt tattgtttat tttagtggta ttttttttgt tatattgata ataaatgttt 2040  
taatttattg ataaattatt gagtgtaaat tattgtattg gttatttatg gtatatgagt 2100  
ttatttaatt tttatgataa ttttatgggt gttattatta agttggtata aaagatgagg 2160  
aaattaaagt ttaagaatt atgagtgatt ttttaaaaga taaatttagt ggtttttaat 2220  
tataaagttt atgatataat aatttttata ttattttagg tgtggtttaa gagattatag 2280  
attaatttgt agaaattatt tttattttaa gaaagtttag ttttaataata tttattaagt 2340  
tatgttttga aaaaggttta gaaaaaagat tttatttttt attttaggtt ggttttatgt 2400  
aattgatatg agtattgttt aatttaaaag tagtattaat tttgaatatt tatggttata 2460  
atttattttt taatgagtgg ggaattattt ataattttta agaggtttt gaaaagggtg 2520  
tgttgtaaat ttaaaatgat aaaggtagtt gttgtgtttt ttattattta ttgggtgttt 2580  
tttttttag atttattttg gaggggttag gtattgttag gtataatgtt ttttttttg 2640  
gtttttaagg tagaaagggt tttggtagt ggggtggtat aaagtggatt tggagtatgg 2700  
ttgagagtat attttgggtt aatgaggaat gttagttaat atatattgga gagaaaaata 2760  
tgaatggata gatttaatta atttttaga tttatttttt ttattttata tgtgagaaaa 2820  
ttaagggttt agaattttta gaagtttgtt aataaatggt agatttgagt ttttaagttag 2880  
ttttatttta agttattttt tagttattgt gatgttatta ttagtattaa tggatttttag 2940  
aaaaatagta atttttttaa atatgtaata tattataagt taaatagta atttaaatgt 3000

<210> SEQ ID NO 22

<211> LENGTH: 1984

-continued

---

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (9, 17, 18, 33, 41, 49, 50, 52, 54, 62, 80, 84, 127,
157, 159)
<223> OTHER INFORMATION: unknown base
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (175, 207..209, 214, 1181)
<223> OTHER INFORMATION: unknown base

<400> SEQUENCE: 22

agttttagnt attatnnta tttttattaa tgnntttggg nttataggnn gngntttttt    60
gnttaaatag tagttaatgn aggnatntta aaaggatat tttttatat ttatttagtg    120
gttatgntat ttttaattt ttattataat tttgtgngnt tagtataatt tgtgntttta    180
ttttatttgt gaataaatgg aggtagnnnt ttgngtaatt tataagtaag tggtagattt    240
gggggttggtg ttagataat gtggttttag attttttggt tttttttggt ttttgttga    300
ttttgattga gtttagttat ttattttttt aatataatatt aattggtggt tttatgtttt    360
aggttttggtg ttagatgtta gaaatatatt tttatttatt ggattttggt atggtttatt    420
gtttagtgtt ttgttatgtt atgtgttata taatgtatnt attttggtat ttagtttatg    480
gaaaataatg ttgaaagata ttgtatgtat gtttttataa taaaattatg taatttatgt    540
tttttttatt gttttggtta ttaggaaggt tttatttaaa ttagagtaaa tatattagaa    600
tttgggtttg ttatttttagt ttgttgaatt tttttttttg gtttagattt tttattttgg    660
tttataaatt ttattgtata aatgtttttt attgtaaaat attttaaatt ttttttaagg    720
gaagggttga tggaaatgat atagtaaggt tttttttgta tttttttaga tttttattag    780
ggaaggtaga ttagatggtt ttttttattt ttagttttaa agttttttat ttataaaatt    840
ttttaagtag tgattattgt tgttttgagt atttgagggt agtttagagt tttgaaagg    900
taaaatttat atataaaaag aagttttttt tgatttagat tttagttttg gtgtagatg    960
tatatttgag gagtagggtg ttagttagat ttttatttga gtagttaata aaatttattg    1020
tttttaaatg gaattttttt tgtaagtttg aattgatttg ttatttttgt gattgtgag    1080
atggtagggg agtattaaat attattatga tttgggttat agtgggggga aaaaaggaaa    1140
aaagaaaaaa aaaatttatt gtttaagttt gttaggtgta naaaggggtg gaattgttg    1200
ggttatttta tttgatttat tggaaataga gtggatttta ttaatatatt aataaagaga    1260
atttttttgt attagggttg aagtgggtgt tagttttttg tgtaatttta ttttttgga    1320
aagtggaaatt agttggtatt gtttagtgtg atttgtgagg ttgagtttta atagtttaaa    1380
gaagtaaatg gtagttatt tttgtggggt ttgttttttg tgagggtttt atttgtatt    1440
tgttatgtaa aatgagggag tgttaggaag gaatttgttt tgtaaagtta ttggttttg    1500
ttattagttt ttatttaagt tttttgtgat gttgtgttg atttatttg gaagttggtt    1560
ggttggtgag gtagagtttt tttttaaagt ttggttttta tggaaaatat gtttagtgta    1620
gttgtgtgta tgaatgaaaa tgttgttggg tgttttttag tggataaaat gtagttgaga    1680
atttgtttg ttttgtgtgt tttttgttt taggtagggg agaggggttg ttgggtgtgt    1740
tttgtgtttt gtttttgat ttggattgtt tggatgggt aggggtgagg ggtttttgg    1800

```

-continued

---

```

gggttggggg ttttggtgt ggtggtgaag atagattggg gtttgtagg gaggttattt 1860
tgagtttaga gattttaggt atttttata tataggtttt tattttggtg tgtgtgtgtg 1920
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtatgtt tgttaatggg 1980
agga 1984

```

```

<210> SEQ ID NO 23
<211> LENGTH: 1984
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (804, 1771, 1776..1778, 1810, 1826, 1828, 1858, 1901,
1905)
<223> OTHER INFORMATION: unknown base
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (1923, 1931, 1933, 1935, 1936, 1944, 1952, 1967, 1968,
1976)
<223> OTHER INFORMATION: unknown base

<400> SEQUENCE: 23

```

```

ttttttgtt aatgaatgta tatatatata tatatatata tatatatata tatatatata 60
tatatatgta tgtatgtag agtgggagtt tgtgtgtggg gggtgttag gatttttggg 120
tttgaatga ttttttatt gagttttgat ttgtttttgt tgttgtgatt ggaggttttg 180
atttttgaa agttttttta ttttgttgt gttgggtgat ttgaatgtag aaatggggtg 240
tagagtgtgt ttggtagttt tttttttta tttgggatag gagaatgtat agaatgagtg 300
gagtttttgg ttgattttt tttgattaga agtgtttggg ggtgtttta tttatgtatg 360
tggttgtatt gagtatatt tttgtgggag ttaggttttg aggagagggt ttgttttgtt 420
agttagttaa tttttaaat agattagtag tagtattatg aaagtattgg gtagagggtg 480
atgattagga ttaatggttt tataagatgg attttttttt aatgtttttt tgttttgtat 540
ggtagatatg gggtagat tttgtgagga gtgagttttg tggagggtgg attttatttg 600
tttttttga ttgttggggg ttagttttat aaattatggt gggtaatggt agttgatttt 660
atttttttag agaatggaat tgtatggggg attggtggtt atttttagtt tagtgtaaaa 720
agattttttt tattaanaatg ttaataagat ttattttatt ttaataaat tagataaaat 780
ggtttttagta gtttttagtt tttntatggt tggttaagggt ttgtagtgga ttttttttt 840
ttttttttt tttttttat tattgtggtt taagttagta tgggtgttgg tgtttttttg 900
ttattttata aattataaag atgatagatt aatttgagtt tgtagaaaaa attttattaa 960
gggtaatag attttattaa ttgttttagt gagagtttga ttagttattt attttttaa 1020
tgtgtatttg gtattaaagt tgggatttgg attggaagga atttttttt gtatatggat 1080
tttattttt agaagttttg aattattttt aggtatttag aatagtagta gttattgttt 1140
aagggtttt ataaatggg ggttttggga ttgagggtaa gagggatatt taggtttgtt 1200
tttttaata ggaatttaag aaaatgtagg gaagatttta ttgtattatt tttatgtagt 1260
tttttttag aaagaattta aggtgtttta taataaagga tatttgtgta atagaattta 1320
tgaattaaaa tagaaaattt gggtagaag gaaaagtta gtaaatgaa atgataagtt 1380
taaatttaa tgtatttgtt ttgatttggg taagagtttt ttgatagta aagtaataa 1440

```

-continued

---

```

agaatataa gttatatagt tttgtttag aatatgtat ataatgtttt ttagtattat 1500
ttttatggg ttaaatattg gaataaatgt attatgtaat atgtaataa gtaggatatt 1560
gagtaatgga ttatgttaag atttagtgga tggggatgta tttttagat ttggatagg 1620
gtttgaata tggagatagt aattagtggt tgtaaagaa atagatgggt agatttaatt 1680
aagatgtagt agaaggtagg gaggggtaag gagttggag ttatattggt taggtatagt 1740
ttttgattt gttattttt tgtaaaatt ntaannntt gtttttttt gtttatagat 1800
aaaatgaan tataaattgt attgantnta taaggttgtg gtgaggatta agtgatgnta 1860
taattattga ataaatataa agtgatgtgt tttttaagat nttntattg gttgttgttt 1920
ggntagaagg ntnnttttg tggntttaag gntattgata aagatgnnag tgatgnttga 1980
agtt 1984

```

&lt;210&gt; SEQ ID NO 24

&lt;211&gt; LENGTH: 7833

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 24

```

gtttttgggt agatatgtgt tttataagtt ttaatggaga aaaatgtaag tattttattt 60
tttgaaattt ggttatttga gtaatgagaa aatagttatt ttttttagga tagtggtttt 120
taattatggt tatgtgtttt tttaggaaaa ttttaaaaat atatataat taatgttttt 180
gtgttatttt tagggatttt aagtttttga atatgaattt tgtattagta ttttttaatt 240
atttagtgga ttgtgatgtg aaattatgat tgagttttat tgttttaaga tgaataaaat 300
tttttttagt attgaaatta taaattttaa ttattaaaat taattaaggg tatgggaatt 360
aataaggtat agggaagttt ttatattata aaattatttt tttaaattat agtttattgt 420
ttatatgtta tttgttattg tagaaaaggg tgaaaaaata gtaaattdaa ttatttttag 480
tttgaaaaat tatttagaaa tgaagatgat gattttgaaa tattgttaat attatttgat 540
ttataaataa tgttttaata tattttattat atattgatag atattttttt atatgaatat 600
tatatattaa aattaaggta ataatgtatt tagaatattt tatttatatt tatgtatttt 660
aagtaggtta gaaattaaga tatgagttat taagtatgag atgttaaggt gtggggtag 720
aaattatatt gtattttatt attaataatt aatataatatt ttaattattat atatatttaa 780
ttttaatttg tatattttta attattttta attatgtgta taaatataag tatataatatt 840
tttatgtatt tattttatta tatttttatt tattttatta tataggggat ttttttaaat 900
ttattattat taaattatat atttttattt taatttttag aataagttda ggaggtaggt 960
attgttatta tttatatttt ataaatgagg aaattgttda tagttataaa gttattgtgt 1020
tagatatatt agaagtttaa tatatatttg gtgaatataat gtataaaaat agagagatag 1080
atatgtataa tagtttattt ttatattgag taaaagtttt taatttggtt tagaaatttt 1140
tttgtgaaaa ttgagtaaaa attgaggtat ttttttattt gttatatagg tataggtggt 1200
attttatttt ttttaataagg atgaatattg aaatgtggat ttttaaggttt aatttttagat 1260
tttttgaatt tttgatagtg ggatttggaa tttgtttatt gttttaaagt tttttaagga 1320
atttatatga ttaattaggt ttagaaatta ttggatttda ttgttgaagt ttgagaatta 1380

```

-continued

---

aagtttgggt tttattgtgg ttttatagaa agggtaaatg aagtattatg gatagaattg	1440
atatgttttt agttagtttt tttttttaga agttaatagg tagtaatata gtagaaatta	1500
gtgatttatg ttttgtgttt tgaagttagg tagaatttta tagagtttta gtagtgttat	1560
tgatgagatt tgttttttgg ggtaagttgt ttgatgtttt taaagttata ttttttttat	1620
ataaatgag ataatttttt ttgttttata ggggtgtttt aaagattaaa taaaaataat	1680
atgttttatt ttatatggta taatgtttga tatttaagaa gtaaaggata ttttttattt	1740
ttattgaagt aattagaaaag tatgaaatta tgaaggagat aagagttttg attggtagtg	1800
tattttattt ttttaggttt atttatttat tttaaattat ttttgttga gaataatttt	1860
taagtttttt atttaagttg tgagtaattt tatattttat aatgatgttt tttttatgag	1920
aaaaaaaaat gtttttaagt tttttggaga aaatatattt gtattatttt tattgaaaaa	1980
tttaataatt ggattttgtt tttttgtatt aatttttagag tgotatgtt ataaataaag	2040
tgttttagtt taagaagatt gaaagtaaat atggtatagt attttaaat aagaattttg	2100
taaatatatg gtatgattgt gttatattat tagtaattat atgatatgta atgtaaagta	2160
tagtttatag atttaaattt aattttaata agtaaattga ttttgttttg ttggggaaaa	2220
gttaaagtat taatttaatt gttaatgtag ttttgtttat ttttttgga ttttagtgata	2280
agtttaata atgtatata ttttatttat atatttagta atataatttt ttgtttaatg	2340
agtgatgttt ttttgttatt tgggtgtgtt tgttagtttt agaatttgtt ttttgggtgt	2400
attataatat taagatagaa gtaagtgtaa taaaattgta gtatttttat tgaagagggt	2460
ttgttttaaa ttgtttaata atttaaagga tttttgtgga agtaattgta tttgttaatt	2520
agttataatt agtaattaat ttttttgag ttttaattta tttttggtaa aatgttttag	2580
gaagagtata tattattaga aagtatgtta aaaatttatt tagtagaaaa tttaaaaata	2640
gttttttttt gtttaagagg tttttaaaat tttatttata tagttaaatt ttgaaatttt	2700
agtaggtttt gttttattat tataaattat gtataaatat ttttaaggat tttgttttta	2760
gttttaagta tgatttattt ttataagttt gattagtatt tatattagtt ttgttatgga	2820
aatgatatg tttttatttt ttgtttaga gttgttaaat tttgatttat atttatgttg	2880
ttttttttgt tgaaggtttg tagtgaaga aatttttaat tttttgtttt gtaataatag	2940
ttggtagttt tatttaatgg gtattttgtt tttttaaaga atttagttgt tttgtttaga	3000
agttgatttt ttgatgtttt taatgtttgg ttttaattgat ttgttttaat ggagtttttg	3060
ttggtgagga gtgagatgtt attgattaga atgttgggat ttgttgttta attgtagga	3120
gtgagagata ttgagattta gaaatttttg gaggtgggag gggagaggga tagttttgga	3180
tggagggtga gatgtaagat aaagggatgg attttatata ggaaaaaaaa aaagattttg	3240
ttgaggattt gaggtgttgt atgattatat tttttaaagg agaagttaaa aagtaaggaa	3300
gtgggaggag gttggagggt aaagtattta aaaggattat ttgggtataa tttgtttttt	3360
tgttggtgtt tgtaaaggat agatagtttt gtttttaag tatatgaatg ttttttttaa	3420
gtgattggga atggatatta attgtttgtt aaatgttatt aaatgttttt ttaatttag	3480
gggatataga aagaggggtg taaaaggaga atttaaatag aaaaaggag gattttggagg	3540
tttttgaaag tggggggaga agaaggagga gggataatag agaggaatag agaaggagag	3600
tggagagaag ataaataaaa ataaaaatag gaattattga ataattatat attaaaaaga	3660

-continued

---

aagtTTTTTT	ttatggggta	tttaaaatat	tgagattgta	atagtgattt	tggttatgga	3720
agaaagatgt	TTTTTTTTat	TTTTGTTTT	gaaagTTTT	ggTTTTgtta	ttggtgatta	3780
aaatTTtatt	aggTtaaaga	gtgtgtTTaa	ttgtttgaag	aatgtagtag	atggaaggTg	3840
ggTTTTgtta	tgttgtttgt	TTTTTTgtt	ggagagaatg	aaagaaatgt	gtagagttag	3900
agatTTTTgt	tgagttagat	TTTTTTTTgt	tgttttaggt	tattggttat	ttggtaaaga	3960
tttgagtaag	gaatgtaggg	tattgtttg	ggTtaataaa	tggagtttgt	TTTTTTTTTT	4020
ttggatgttg	ttgtttggtt	gatgtTTTTg	gtaatttatt	tgtgggtgat	gtagaggagt	4080
TTTTTTTTTT	TTTTtagatt	attgtTTTTg	attaatttga	TTTTTaaat	atatttgatt	4140
gtatTTTTta	ggTggatata	ttaataggtt	atgggtTgga	gaggagtggg	tgatgaggag	4200
agggattTaa	atTTgtgaat	gtttgggttg	ggtTggagtt	gtgggggggt	tgggaggaga	4260
gaggggagaa	gagagaagga	aggagagtgt	ttgttgggat	ggtTgagttg	ttttggtgag	4320
tagTTTTggg	gttGtatgtt	tttGtgggag	atgttGttgt	tgtTTTTagg	ttggtaaagag	4380
tggTTTTaat	attattgttt	tttattTTTT	TTTTtGaaa	TTTTtagaga	aatgtTTTTg	4440
gtTTTTTTgt	tgtgatattt	ttagtttGta	TTTTTTtata	gtttaggTgg	tgtgtTTTTg	4500
tatgtTggag	tgtTggtTgt	tagtaggatg	TTTTTTTTtg	tgtTgatttg	TTTTTTTTtg	4560
TTTTgttGt	gttGTTTTt	tgatattTTTT	gtTTTTatta	TTTTtagttt	ggagagatgt	4620
tatttagttg	tggTTTTgat	ttgtggTTTTg	gggttatgtg	tggaaagagg	gtgttagttt	4680
ggatTTTTgt	tttGttaggg	ggTgtTTTTg	agtggagagt	gaggtgaatg	gtatatgagt	4740
gtgtgggtag	tttattTTTTg	agtttGagtt	tttattTtga	gttatgtttt	gttttagttt	4800
atTTgggtta	gtgtttggTg	agtGagttta	tttGtggttt	ttgtggTtgt	TTTTTTTTtg	4860
tattTTTTgta	tttattTgtt	gattTTTTtt	ttttgggatt	tgtattTtgt	tttattaatt	4920
agagtttgat	tgtTTTTttt	tatgtgattt	tgggtgggtt	gaggatttgt	tgtTTTTtaa	4980
atgttagagg	gatgtgggtg	gtagagtttg	agaggtggtt	gttgggttgt	gggggtttt	5040
gattTTTTtt	ttattTtGtt	TTTTtgggtt	ttattTgttt	gtTTTTggat	TTTTgttttt	5100
TTTTgttttt	tggTTTTtta	gagTTTTttt	tttatggtag	tagTTTTttg	tgtTTTTggt	5160
gtagTTTTtt	agtGgatgat	TTTTtGttt	tggggtTgag	tttagTTTTt	ggatgtTgtt	5220
gaaatTTTTg	agattatgtg	tgggtTtggT	tgtTgttttt	ttgttgggtg	ttattgttat	5280
tgtTgtTgt	tttGttgttg	ttgtttgtgg	gatgttttagt	agtTtGttgt	ttggTTTTtg	5340
tgattTtGtg	TTTTttgGaa	gttGtttGtt	gttGtagagt	tgtatGaaat	agttatggTg	5400
ttgtgggagt	TTTTgtggta	gtgtagtagt	tggatatttt	gtgagggttt	ttgttgggtg	5460
ttgtTgtTgt	ttgttatgtt	atTTattGta	gtttGttttg	tGaaGttTgt	tgtTTTTttt	5520
atTTTTtTaa	gtgattgtta	aatgtTtatt	ggtTggaatt	gttttggTaa	gtttagaatt	5580
tttGTTTTtg	atTTTTtTaat	tttGtagaag	aatatgtGta	tttagtatag	attagTttat	5640
tttagtGtgt	TTTTttagtt	tttattTTTT	tattGttTta	gattTTTTaat	attatttatt	5700
tttattTtaga	gaaataaagg	gaattgtTgt	aggTTTTggg	gtgaggggtg	gttttgggat	5760
gggtagaaag	tgtaggTgta	gtagGaaatt	tttGtatgtt	tgtGttTata	ttggagTtgt	5820
gaggatTTtg	agaaatatta	aatgggatgg	TTTTttgggt	ttattGtttt	Gaaagagtat	5880
taattTtagg	gGaaatattg	aaatagaagt	tttGttatta	tTaaGaaaa	aagtTttatt	5940

-continued

```

aggatgagga agaaataatt ttatgagaaa gaatgagtga gaaagtaata aattaaatgg 6000
tgattgtagg ggaattgttg atttttggta aagggtttat gaggttgat tggttttttg 6060
ttgaagatta ggttatatag attttagagg agttgggttt taatagaatt tttttttttt 6120
tttttttttt tttttttttt tttttttttt tttttttatt tatttttttt tttttttttt 6180
ttattttttt ttttttttag tggtaaaaga tattggtttt gtagtttaga tatgtttttt 6240
tttttgtttt ttttaagtttt aaggtagtat aggggagttg agaaaaagaa ttttttgagg 6300
gttttttagg ttggagtggg tatgattgag gttggttagg ttttatgtag gtgagttgag 6360
ggtggaattg attttagtgg gtgttgattt tttttttttt ggataggttt ttgtggagtg 6420
ggttaggtat tttttttgtt tgtttgggtt ttttttagatt ttgatgggta atgtttggta 6480
ggttttgttt tgttgaagt ttttaattaa atagggttag aggatgggag ttgttgatt 6540
tttagttggt atagtatttt gtttgatagt ttgtagtata ggggtgatgt aattttttat 6600
tttttgtaa tataattttt ttgtagtaa atttggtttt gaataaagt ttttttaaag 6660
atgtatataa gttgaagtgt atgtaatttt agagaggagg gaatgattaa ttgtaattta 6720
gggtgaaagt ttgtatagtt tttagtattt attgatgtaa atgttaaaag gaaaattatt 6780
atgtattatt ttaatttttt ttttataaag ataagttgag atatgtaatt ttattagatt 6840
tgggttaata gattgttttt ttttttggtt gtttttaaat ttggtatttt aataaaattt 6900
aatatgtttt tataattttt tgatttatgt gtatatgtgt gttgtttttg aaagaataag 6960
ttttattttg ttattgttta attatttttt agatgtttta ttatggtaat aattatgagt 7020
ttgtaaaat aatttttgga aatgttgatg gttttgtagt ttaatataga ttggtttggt 7080
ttatttttag tttttgtatt gttttaggaa ataattaatt taaatgtgaa gttgatattt 7140
gtaattaaga aattatata ttattagata ttttaaaagg gattgtataa attaaagaga 7200
ataaattggt tttgtagata ggttgtaag aatttggtat tttgttttta tttttgtaa 7260
tttagagggt attaattttt atttgagtta aatagattat tatagaaat attgtgtttg 7320
tttattttta ttattgaggt tttgtttttt tttgttttg atatattta aataaggggt 7380
tgttttagtt gttgaagtaa aagaataatt aaagatgggg aaatggtaaa agggatttta 7440
gagattatta ttagtttttt tttaaaatgt ggagttttgt ggttataaat attgtttatt 7500
taatgagtaa aaaataaaaa taaaaaaaaa ataggaagta aatgtaagt tttattttat 7560
tattgttagt attaatgtaa gtttaaaaa atagtattat tagaaaagga tattaagga 7620
gaattgatta gaaaagaatt gtggaaaatg gaaatgaata ttgattattt aattagattt 7680
tgaggttatt agtagatagt gattttgtag tatagttata gttgttggtt ttaaaattta 7740
ggataagtat tttaaagttt taaagtagtg tttttttttg ttaaaaattt gtaagatgtt 7800
ttaatgattg gagtgttttt tttgaatttg agg 7833

```

&lt;210&gt; SEQ ID NO 25

&lt;211&gt; LENGTH: 7833

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

&lt;400&gt; SEQUENCE: 25

```

ttttaaattt aaagagaata ttttagttat taaaatattt tatagatttt taataaaaaa 60

```

-continued

---

aagtattatt ttgaagtttt aaaatatttg ttttaaatTT taaatttaaat aattatagtt	120
gtattgtaag gttattgttt attgataaatt ttaaaattta gttaaagtgat taatatttgt	180
ttttattttt tataattttt ttttagttaa ttttttttta gtattttttt ttgatagtggt	240
tattttttta agtttgtggt aatattgata gtgggtgaatg aaagtttaaat atttgttttt	300
tgttttttt ttatttttat tttttgttta ttaggtggat aatatttatg attataaaat	360
tttatatttt ggaaaagagt tagtgatgat ttttgaatat ttttttatta tttttttatt	420
tttaattggt tttttgtttt aatgattgaa ataatttttt atttgaaatg tatttagata	480
aagaggaaat aaagttttta taataaagat aaataggat agtggttttt gtgatggttt	540
gtttggttta aatgaagatt gattattttt aagttaatag ggggtgaagt ggggtgtaa	600
gtttttgata atttatttgt aaaattagtt tttttttttt agtttatgta gtttttttta	660
aaatatttgg taaatatgta attttttgat tgtaaatggt aattttatat ttaagttagt	720
tattttttta aataatgtaa gggttaggaa tgaagtaaat tagtttgtgt tggattataa	780
agttattaat atttttaaaa attgtttttg taggtttata attattatta taataaagta	840
tttaaaaagt gattaggtta tagtaaagtg aaatttattt ttttaaaaat aatatatatg	900
tatgatgaa ttaagaagtt atagaaatat gttgagtttt attaaaatgt taaatttaga	960
aattgttaaa aaagagaata atttattgat ttaaatttaa tagggttgta tatttttaatt	1020
tgtttttga aaggataaat tagaatgatg tataataaatt ttttttttgg tatttatatt	1080
agtaataaatt aggaattata taggttttta ttttgagtta tagttggta ttttttttt	1140
tttaagttta tatatatatt agtttatata tttttttgaa agatatttta ttttagagttta	1200
gatttaatta tagtaaaatt atatttatag aagatgaaaa atttatata ttttatatta	1260
taggttgta aattgaatgt tatgttagtt aggagtgtag taatttttat tttttggtt	1320
tatttaatta ggaagtttta gtagagttaa gttgttaag tgtttgtgt tagaattga	1380
aggaattga gtagtaaga agagtgttg atttatttta tagaagttg ttagaaatg	1440
gaggagttag tgtttattga agttggtttt gtttttggtt tgtttatatg gagtttgatt	1500
agttttagtt atgtttattt tggtttggga gatttgtaaa gtgttttttt ttttaatttt	1560
tttgattat tttgaagttt agggaaagtaa agagaggggt atatttggat tgtaaaatta	1620
atgtttttt tggttagga gagaagggaa tgagagagag agagagatag atagatagag	1680
agagagagag agagagagag agagagagag agagagagag agaaatttta ttgaaattta	1740
gttttttag aatttgtgtg atttggtttt taatgggaga ttagtgatgat tttatggat	1800
ttttgttagg aattagtgat tttttgtag ttattatttg atttattggt tttttgttta	1860
ttttttttta taaagtattt ttttttttat ttttagtaaga tttttttttt taatgatgat	1920
aaagttttt ttttagtggt ttttttagga ttgggtgttt tttaaaatag tgaatttaga	1980
aaattatttt gtttaaatat ttttaaaatt tttgtagttt taatgtaagt gtaagtatgt	2040
aaaggttttt tgttatattt gtattttttg tttatttttag aattattttt tatttttggg	2100
tttgtaatag tttttttgt ttttttggat agaggtgggt ggtattaggg gtttaggta	2160
gtaggaggtg aggggttagg gaggtgtgtt agggtaggtt ggtttgtgtt ggatagtggt	2220
gttttttgt ggagttaaag ggttggggat ggggttttg gattatttag agtaatttta	2280
gttgggggt gtttggtag tatttaagga ggtagggaaa gtagtgagtt ttattgggtg	2340

-continued

---

ggttatgatg agtagtatga tgggtagtag tagtagttag taaaagtttt tgtaaagtgt	2400
ttagttgttg tattgtgtg gggattttta tagtattatg attagtttgt gtaattttgt	2460
agtagtaaat ggtttttgag gaatatagga ttgtgggggt tgggtagtgg gttattgagt	2520
atthttgtga tgggtgtagt agaggtgggt gtggtggtag tggattttgg tggggaagta	2580
gtagttaaat ttgtgtatga ttttgagagt tttagtaata tttagggatt gggtttagtt	2640
ttggagttag agggttgttt gttgagaagt tgtgttgag atgtgggaag ttgttttat	2700
aaggaggag ttttgggaag ttggaggata ggaggagatg ggagtttagg ggtagatgag	2760
tggagtttga ggaggtaggg tggagggaga gttaagggtt ttttagttt ggtagttgtt	2820
ttttgagttt tgtttttt atttttttgg tgtttggaa gtagtaggtt ttttagttt	2880
ttggggttat gtgggaagag gtagttgggt tttgattggt ggagtaggat gtaggtttt	2940
ggaggagggt gttgatgagt aggtgtaagg atgtaaggag gaggtggtt tggaaagttat	3000
agatgggttt gtttgttagg tgttgggtt agtggggtta ggtgggtat ggtttaaag	3060
agaagtttgg gttttaggtt gggttattt tatatttata tattatttgt tttattttt	3120
gttttaggat gttttttt gaagggtggg tttggattag tgttttttt ttgtgtgtga	3180
ttttgggtt gtagtgggtt ttgtggttgg gtggtgtttt tttgagttgg agatggtggg	3240
ggtggagggt ttagaggagt agtagtagta gggtagagag gggtaggtt gtgtgggaga	3300
gggtgtttt ttggtgattg gtgttttagt gtgtgggagt gtgtgttta ggtttaggg	3360
ggatgtagg tgggaatggt ttggtggaga ggttagggat gtttttttag gatttatag	3420
gaaagagggt gagaggtag ggtgttagaa ttgtttttt tgatttggaa gtaatagtag	3480
tattttttat aagagtgtt aattttaagg ttgtttgtt aggtagttta gttattttg	3540
taggtgtttt ttttttttt ttttttttt ttttttttt ttaggtttt ttagttttg	3600
atthagtta agtgtttgta ggttgaatt tttttttta ttatttttt ttttttagt	3660
tgtagttat tagtgtttt atttgggagg tgtggttaga tgtgtttga aggttagatt	3720
ggttgggata agtggtttga gagaagaga aaggttttt tgtatatgt gtgggtgggt	3780
tgttgggagt attggttgg tagtgggtt tgggaaggg agagtgggt ttatttgtt	3840
gttaggttag tgattttgt tttttttt gggttttt tggatggtt gtgatttgg	3900
gtgatgagag aaggtttaat ttggtaggag tttttggtt tgtgtttt tttattttt	3960
tttagtggga agggtaaat gtatagtggg atttgtttt tgtttgtt atttttagg	4020
tagttagata tattttttg ttaaatggaa ttttagttt tagtaatggg attaagagt	4080
tttgggata aggtgggaga ggaatattt tttttatga ttgggttat tatttagtt	4140
ttagtgttt gtagtttta tagggaagag ttttttttt ggtgtgtgat tatttagtga	4200
ttttgtttt tgtttttgt tattttttt ttgtttttt tttttattt tttttgttat	4260
ttttttttt ttttttttt ttgtttttaa aagtttttgg atttttttt tttttatth	4320
aattttttt ttgtgtttt ttttttgtt tttttgaatt taggagagta tttgataata	4380
tttaaatagg aattagtgt tatttttaaat tttttaaag aggtatttat atattttgaa	4440
aatgggatta tttattttt gtagatatta gtagaaaaat aaattgtatt tgagtaattt	4500
tttaagtat ttaattttt aattttttt tattttttt ttttttaatt tttttttga	4560
gagatgtgat tgttagtat ttttagttt taatgaaatt ttttttttt ttttgtgtga	4620

-continued

---

aatttatttt tttattttat atttttgttt ttgtttgaga ttgttttttt ttttttttat	4680
ttttaaagat ttttgaattt tagtggtttt ttttttggg aattaagtag tagattttag	4740
tatttttagt ggtgggtatt tgttttttat tgatgaagat tttattaaaa tagattaatt	4800
agattagatg ttggaggtat tagaaaattg gtttttagat agagtagtta aattttttta	4860
ggaaatagaa tttttattag atagagttgt taattaatat tgtaaaataa ggaattagaa	4920
attttttttg ttataggttt ttagtagaga aggtaatata aatatagatt aagattta	4980
aattttatag tagagaatga gaatatgta tttttatag taaggttggg gtggaatta	5040
attaggttta tgaataaag ttatgtttga aattaaggt aaagtttta aaagtgttta	5100
tgtagtaatt atgataatga aataggattt gttaggattt tagagtttgg ttatgtaagt	5160
agaattttag agaatttttt agtagaggaa aattgttttt gaattttttg ttaagtaaat	5220
ttttggtata ttttttaata atatatgttt tttttaagat gttttgttaa aagtaagtt	5280
aaattttaaa ggagtttaatt attggttgta attggttaat aaatgtgggt gtttttatag	5340
aggtttttta aattattaaa tagtttgaag taaagttttt ttaatgggaa tgttgtaatt	5400
ttgttgattt tttttgtat ttagtgttat agtgttatta agaaataaat tttgaaattg	5460
gtaagtatta ttaagtggtga gaagaatatt atttattgag tagagaattg tattattgaa	5520
tatgtaataa aaaatatata tattatttag atttgttatt aggtattaaa gaagtagata	5580
agattgtatt agtaattgga ttagtgtttt aatttttttt tagtaaggta aaattagttt	5640
atttattaga attaaattta agtttatgaa ttgtattttg tattgtgtat tatatgattg	5700
ttagtaatat gatataatta tattatgtat ttgtaaaatt tttattttta aatattatat	5760
tatatttatt ttttaatttt ttgagttaga atattttatt tgggtatat atattttaga	5820
attgatgtag aggagtagag tttagttggt agatttttta gtgaaatag ttagatata	5880
tttttttag aaaatttaag aatatttttt ttttttatgg aaagaatatt attataaagt	5940
gtgagattat ttatagttta agtagggggg ttgggagttt ttttttaata agaagttt	6000
aagataaata aatgaatttg ggaaaaaag atatatgttt aattagaatt tttatttttt	6060
ttatgatttt atattttttg attgttttaa taaaggaag atgtattttt tgttttttag	6120
gtgttaggta ttgtgttag taggatagaa tatgttattt ttatttaatt tttaaaatat	6180
ttttatgaga taaagaatat tttttattt tatataaag gaatatgggt ttgaaagat	6240
taggtaattt gttttaagaa ataaattttg ttagtgatat tgttgggatt ttgtgaaatt	6300
ttgtttgatt ttagagtata agatataagt tattaatttt tgttgattg ttgtttgta	6360
gtttttgaga ggggaaatta attgggaatg tattagtttt gtttatgata tttattttgt	6420
ttttttgtg gagtttagt aaggtttaaa ttttaatttt taaattttgg taataagatt	6480
tagtgatttt tgaatttggg tgattatag aattttttga gaaattttga aataatagat	6540
aaattttaag ttttattatt agggatttag aaaatttgga gttgggtttt gggatttata	6600
ttttaatatt ttttttgtt ggagaagtaa ggtattattt atatttatat gataaatgaa	6660
aggataattt gatttttgtt tagtttttat agagagggtt ttgagatagg ttaaaagttt	6720
ttatttagtg taaagatgag ttgttgata tgtttgttt tttgtttta tgtatagtt	6780
tattaaatat gtattaagtt tttaatatgt ttgatagat aattttgtga ttgtagataa	6840
tttttttatt tgtaaaatg gagtaaat aatattgtt ttttgggtt gttttaaaga	6900

-continued

---

```

ttaaataaaa aatgtatggt ttaatggtag tggatttggg gggatttttt atataaataa 6960
gtgaatggag gtatgaataa ataaatatat aaagatgtgt gtatttatat ttatatatat 7020
aattaaaaat agttaaagat gtataaatta aagttaaatg tatgtgatat tgaagtatat 7080
gttgattatt gataatgaag tatagtataa tttttaattt tatattttta tattttatat 7140
ttaataattt atattttaat ttttagttta ttaagatat atagatatag atagaatggt 7200
ttaatgtat tattgtttta gttttaatgt ataatttta tatgaaaaag tatttattag 7260
tgtgtagtaa atgtattaga atattattta taggttaaat gatattgata atgttttaga 7320
gttgttattt ttatttttgg ataatttttt aaattgagag taattaaatt tgttattttt 7380
ttattttttt ttataatggt aaataatata taaataatga gttgtgattt aaagaaataa 7440
ttttataatg taaaagtttt tttatgtttt attgattttt atgtttttta ttaattttgg 7500
tagtttaagt ttgtgatttt agtgttgagg aaagtttatt ttattttaga gtagtggggt 7560
ttagttatga ttttatatta taattatttg gataattaa gaatattgat gtagagttag 7620
tatttaaaga tttggaattt ttagaagtga tatagaagta ttggtatata tatattttta 7680
aagttttttt ggagaaatat atagttatga ttgagaatta ttgttttggg gaaagtgatt 7740
atttttttat tatttaaata gttaagtttt aggaggtaaa atatttatat ttttttttat 7800
taaaatttgt aaaatatata ttttattaaa gat 7833

```

```

<210> SEQ ID NO 26
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 0002044:2092P22

```

```

<400> SEQUENCE: 26

```

```

ggataggagt tgggattaag at 22

```

```

<210> SEQ ID NO 27
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 0002044:2462O22

```

```

<400> SEQUENCE: 27

```

```

aaatcttttt caacacaaa at 22

```

```

<210> SEQ ID NO 28
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 0002064:1038P21

```

```

<400> SEQUENCE: 28

```

```

ggaagagggt attaaatgga t 21

```

```

<210> SEQ ID NO 29
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 0002064:1224O19

```

---

-continued

---

<400> SEQUENCE: 29  
ccccaaaatc aaacaacaa 19

<210> SEQ ID NO 30  
<211> LENGTH: 22  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1287P22

<400> SEQUENCE: 30  
tttgattag gttggaagtg gt 22

<210> SEQ ID NO 31  
<211> LENGTH: 22  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1529022

<400> SEQUENCE: 31  
cccaataaa tcaacaacaa ca 22

<210> SEQ ID NO 32  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1127P21

<400> SEQUENCE: 32  
ttgtttgggt taataaatgg a 21

<210> SEQ ID NO 33  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1381020

<400> SEQUENCE: 33  
cttctctctt ctccctctc 20

<210> SEQ ID NO 34  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER CALCITONIN

<400> SEQUENCE: 34  
aggttatcgt cgtgcgagtg t 21

<210> SEQ ID NO 35  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER CALCITONIN

<400> SEQUENCE: 35  
tcaactcaaac gtatccaaa ccta 24

---

-continued

---

<210> SEQ ID NO 36  
<211> LENGTH: 26  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER VERSICAN  
  
<400> SEQUENCE: 36  
  
tgggattaag attttcggtt agtttc 26  
  
<210> SEQ ID NO 37  
<211> LENGTH: 23  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER VERSICAN  
  
<400> SEQUENCE: 37  
  
cactacaacg ctacgcgact aaa 23  
  
<210> SEQ ID NO 38  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER TPEF  
  
<400> SEQUENCE: 38  
  
tttttttttc ggacgctcgtt g 21  
  
<210> SEQ ID NO 39  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER TPEF  
  
<400> SEQUENCE: 39  
  
cctctacata cgccgcgaat 20  
  
<210> SEQ ID NO 40  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER EYA4  
  
<400> SEQUENCE: 40  
  
cggagggtac ggagattacg 20  
  
<210> SEQ ID NO 41  
<211> LENGTH: 15  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER EYA4  
  
<400> SEQUENCE: 41  
  
cgacgacgcg cgaaa 15  
  
<210> SEQ ID NO 42  
<211> LENGTH: 26  
<212> TYPE: DNA

---

-continued

---

<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER HCADHERIN  
  
<400> SEQUENCE: 42  
gacggatttt tttttaacgt tttttc 26  
  
<210> SEQ ID NO 43  
<211> LENGTH: 22  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: MSP PRIMER HCADHERIN  
  
<400> SEQUENCE: 43  
aaataaaata ccacctccgc ga 22  
  
<210> SEQ ID NO 44  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: HM PRIMER EYA4  
  
<400> SEQUENCE: 44  
ggtgattggt tattgttatg gtttg 25  
  
<210> SEQ ID NO 45  
<211> LENGTH: 23  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: HM PRIMER EYA4  
  
<400> SEQUENCE: 45  
cccctcaacc taaaaactac aac 23  
  
<210> SEQ ID NO 46  
<211> LENGTH: 23  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: HM PRIMER CALCITONIN  
  
<400> SEQUENCE: 46  
ggatgtgaga gttgttgagg tta 23  
  
<210> SEQ ID NO 47  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: HM PRIMER CALCITONIN  
  
<400> SEQUENCE: 47  
acacacccaa acccattact atct 24  
  
<210> SEQ ID NO 48  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: TPEF FORWARD PRIMER

---

-continued

---

<400> SEQUENCE: 48  
ggacgttttt tatcgaaggc g 21

<210> SEQ ID NO 49  
<211> LENGTH: 15  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: TPEF BACKWARD PRIMER

<400> SEQUENCE: 49  
gccaccaac cgca 15

<210> SEQ ID NO 50  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2183A188

<400> SEQUENCE: 50  
atgtgattcg ttgggta 18

<210> SEQ ID NO 51  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2183B188

<400> SEQUENCE: 51  
atgtgatttg ttgggta 18

<210> SEQ ID NO 52  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2194A186

<400> SEQUENCE: 52  
gggtaacgtc gaatttag 18

<210> SEQ ID NO 53  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2194B186

<400> SEQUENCE: 53  
gggtaatggt gaatttag 18

<210> SEQ ID NO 54  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2324A187

<400> SEQUENCE: 54  
aaaaattcgc gagtttag 18

---

-continued

---

<210> SEQ ID NO 55  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2324B187

<400> SEQUENCE: 55

aaaaatttgt gagtttag 18

<210> SEQ ID NO 56  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1101A188

<400> SEQUENCE: 56

tatatatacg tgtgggta 18

<210> SEQ ID NO 57  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1101B188

<400> SEQUENCE: 57

tatatatatg tgtgggta 18

<210> SEQ ID NO 58  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1147A188

<400> SEQUENCE: 58

agtgatgtagc tagaaggt 18

<210> SEQ ID NO 59  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1147B188

<400> SEQUENCE: 59

agtgatgtg tagaaggt 18

<210> SEQ ID NO 60  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1156A188

<400> SEQUENCE: 60

tttagatacg aaatgtta 18

<210> SEQ ID NO 61  
<211> LENGTH: 18  
<212> TYPE: DNA

-continued

---

<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1156B188  
  
<400> SEQUENCE: 61  
  
ttagatag aaatgta 18  
  
<210> SEQ ID NO 62  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1222A188  
  
<400> SEQUENCE: 62  
  
aagtaagtcg ttggtgtt 18  
  
<210> SEQ ID NO 63  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002064:1222B188  
  
<400> SEQUENCE: 63  
  
aagtaagttg ttggtgtt 18  
  
<210> SEQ ID NO 64  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1279A188  
  
<400> SEQUENCE: 64  
  
gaagtggtcg ttagtttt 18  
  
<210> SEQ ID NO 65  
<211> LENGTH: 19  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1279B188  
  
<400> SEQUENCE: 65  
  
gaagtggttg ttagttttt 19  
  
<210> SEQ ID NO 66  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1346A188  
  
<400> SEQUENCE: 66  
  
ttgttagcg tgattgt 18  
  
<210> SEQ ID NO 67  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002383:1346B188

---

-continued

---

<400> SEQUENCE: 67

ttgttttagtg tgatttgt 18

<210> SEQ ID NO 68

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002383:1475A188

<400> SEQUENCE: 68

aaggaattcg ttttgtaa 18

<210> SEQ ID NO 69

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002383:1475B188

<400> SEQUENCE: 69

aaggaatttg ttttgtaa 18

<210> SEQ ID NO 70

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002383:1524A188

<400> SEQUENCE: 70

aatgttttcg tgatgttg 18

<210> SEQ ID NO 71

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002383:1524B188

<400> SEQUENCE: 71

aatgtttttg tgatgttg 18

<210> SEQ ID NO 72

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002393:1223A188

<400> SEQUENCE: 72

attgttttcg attaattt 18

<210> SEQ ID NO 73

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: 0002393:1223B188

<400> SEQUENCE: 73

attgtttttg attaattt 18

---

-continued

---

<210> SEQ ID NO 74  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1294A188

<400> SEQUENCE: 74

ataggttacg ggttgag 18

<210> SEQ ID NO 75  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1294B188

<400> SEQUENCE: 75

ataggttatg ggttgag 18

<210> SEQ ID NO 76  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1332A186

<400> SEQUENCE: 76

aattgcgaa cgttggg 18

<210> SEQ ID NO 77  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002393:1332B186

<400> SEQUENCE: 77

aattgtgaa tgttggg 18

<210> SEQ ID NO 78  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Calcitonin ML probe

<400> SEQUENCE: 78

cgaatctctc gaacgategc atcca 25

<210> SEQ ID NO 79  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Versican ML probe

<400> SEQUENCE: 79

tcgacgttac ccaaacgaat cacataaaaa ac 32

<210> SEQ ID NO 80  
<211> LENGTH: 22  
<212> TYPE: DNA

-continued

---

<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: TPEF ML probe  
  
<400> SEQUENCE: 80  
  
aattaccgaa aacatcgacc ga 22

<210> SEQ ID NO 81  
<211> LENGTH: 29  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: EYA4 ML probe  
  
<400> SEQUENCE: 81  
  
cgaaacccta aatatcccgataacgccc 29

<210> SEQ ID NO 82  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: HCadherin ML probe  
  
<400> SEQUENCE: 82  
  
gctcctcgcg aaatactcac cccg 24

<210> SEQ ID NO 83  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Calcitonin HM probe  
  
<400> SEQUENCE: 83  
  
acctccgaat ctctcgaacg atcgc 25

<210> SEQ ID NO 84  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: EYA4 HM probe  
  
<400> SEQUENCE: 84  
  
aaaattacga cgacgccacc cgaaa 25

<210> SEQ ID NO 85  
<211> LENGTH: 28  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Calcitonin HM blocker  
  
<400> SEQUENCE: 85  
  
tgttgaggtt atgtgtaatt ggggtgtga 28

<210> SEQ ID NO 86  
<211> LENGTH: 26  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: EYA4 HM blocker

-continued

---

<400> SEQUENCE: 86  
aaactacaac cactcaaadc aaccga 26

<210> SEQ ID NO 87  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: EYA4 HM blocker

<400> SEQUENCE: 87  
ggtatgggtt gtgattttgt gtggg 25

<210> SEQ ID NO 88  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2096A188

<400> SEQUENCE: 88  
aagattttcg gttagttt 18

<210> SEQ ID NO 89  
<211> LENGTH: 18  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: 0002044:2096B188

<400> SEQUENCE: 89  
aagatttttg gttagttt 18

<210> SEQ ID NO 90  
<211> LENGTH: 22  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: TPEF ML PROBE

<400> SEQUENCE: 90  
accgaaatc acgcgcaaaa aa 22

<210> SEQ ID NO 91  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: control primer

<400> SEQUENCE: 91  
tggtgatgga ggaggttag taagt 25

<210> SEQ ID NO 92  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: control primer

<400> SEQUENCE: 92  
aaccaataaa acctactcct cccttaa 27

---

-continued

---

<210> SEQ ID NO 93  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: control probe

<400> SEQUENCE: 93

accaccaccc aacacacaat acaaacaca 30

<210> SEQ ID NO 94  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Homo Sapiens

<400> SEQUENCE: 94

ccttagtccc tacctctgct 20

<210> SEQ ID NO 95  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Homo Sapiens

<400> SEQUENCE: 95

ctcatttaca cacacccaaa c 21

<210> SEQ ID NO 96  
<211> LENGTH: 26  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 96

tggataggag ttgggattaa gatttt 26

<210> SEQ ID NO 97  
<211> LENGTH: 33  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 97

cttattacaa tttaaaaaaa aaattcacta caa 33

<210> SEQ ID NO 98  
<211> LENGTH: 39  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: blocker

<400> SEQUENCE: 98

aaattcacta caacactaca caactaaatt caacattac 39

<210> SEQ ID NO 99  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Probe

---

-continued

---

<400> SEQUENCE: 99  
ttttcgtatt ttttttcggg ttattacggt tt 32

<210> SEQ ID NO 100  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Probe

<400> SEQUENCE: 100  
atgtgattcg tttgggtaac gtcga 25

<210> SEQ ID NO 101  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 101  
gtagggttat tgtttgggtt aataaat 27

<210> SEQ ID NO 102  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 102  
taaaaaaaaa aaaaaaactc ctctacatac 30

<210> SEQ ID NO 103  
<211> LENGTH: 29  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: blocker

<400> SEQUENCE: 103  
aactcctcta catacaccac aaataaatt 29

<210> SEQ ID NO 104  
<211> LENGTH: 22  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Probe

<400> SEQUENCE: 104  
cgaaaacatc gaccgaacaa cg 22

<210> SEQ ID NO 105  
<211> LENGTH: 25  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Probe

<400> SEQUENCE: 105  
gtccgaaaaa aaaaaaacga actcc 25

-continued

---

```

<210> SEQ ID NO 106
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Forward Primer

<400> SEQUENCE: 106

gttagttagt taatttttta aatagattag tag                33

<210> SEQ ID NO 107
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Reverse Primer

<400> SEQUENCE: 107

caaaaaaaca aataaaatac cacctcc                27

<210> SEQ ID NO 108
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Blocker

<400> SEQUENCE: 108

cctccacaaa actcactcct cacaaaatac                30

<210> SEQ ID NO 109
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: LC Probe

<400> SEQUENCE: 109

tttcgttttg tatggtagat acggggtga                29

<210> SEQ ID NO 110
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: LC Probe

<400> SEQUENCE: 110

attaatggtt ttataagacg gatttttttt taacgt                36

```

---

1. A nucleic acid comprising a sequence at least 18 bases in length of a segment of the chemically pretreated genomic DNA according to one of the sequences taken from the group comprising SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto.

2. An oligomer, in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer, said oligomer comprising at least one base sequence having a length of at least 9 nucleotides which is complementary to, or hybridises under moderately stringent or stringent conditions to a pretreated genomic DNA according to one of the SEQ ID NOS:6 to SEQ ID NO:25 or sequences complementary thereto.

3. The oligomer as recited in claim 2, wherein the base sequence includes at least one CpG, TpG or CpA dinucleotide.

4. The oligomer as recited in claim 3, characterised in that the cytosine of the CpG, or the thymine of the TpG or the adenosine of the CpA dinucleotide is located approximately in the middle third of the oligomer.

5. A set of oligomers, comprising at least two oligomers according to any of claims 2 to 4.

6. A set of at least two oligonucleotides as recited in claims 2 to 5, which can be used as primer oligonucleotides for the

amplification of DNA sequences of one of SEQ ID NOS:6 to SEQ ID NO:25 and sequences complementary thereto.

7. A set of oligonucleotides as recited in claims 2 to 5, characterised in that at least one oligonucleotide is bound to a solid phase.

8. Use of a set of oligomer probes comprising at least four of the oligomers according to any of claims 5 through 7 for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) within one of the sequences according to SEQ ID NOS:1 to SEQ ID NO:5, or sequences complementary thereto.

9. A method for manufacturing an arrangement of different oligomers (array) fixed to a carrier material for analysing colon cell proliferative disorders associated with the methylation state of the CpG dinucleotides of one of SEQ ID NOS:1 to SEQ ID NO:5, and sequences complementary thereto wherein at least one oligomer according to any of the claims 2 through 5 and 7 is coupled to a solid phase.

10. An arrangement of different oligomers (array) obtainable according to claim 7.

11. The array of different oligonucleotide- and/or PNA-oligomer sequences as recited in claim 10, characterised in that these are arranged on a solid phase in the form of a rectangular or hexagonal lattice.

12. The array as recited in any of the claim 10 or 11, characterised in that the solid phase surface is composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold.

13. A method for detecting, differentiating or distinguishing between colon cell proliferative disorders associated with at least one gene and/or their regulatory regions from the group comprising Versican, TPEF, H-Cadherin, Calcitonin and EYA4 in a subject, said method comprising contacting a target nucleic acid in a biological sample obtained from said subject with at least one reagent or a series of reagents, wherein said reagent or series of reagents, distinguishes between methylated and non methylated CpG dinucleotides within the target nucleic acid.

14. The method according to claim 13, comprising:

- a) obtaining, from a subject, a biological sample having subject genomic DNA;
- b) treating the genomic DNA, or a fragment thereof, with one or more reagents to convert 5-position unmethylated cytosine bases to uracil or to another base that is detectably dissimilar to cytosine in terms of hybridisation properties;
- c) contacting the treated genomic DNA, or the treated fragment thereof, with an amplification enzyme and at least two primers comprising, in each case a contiguous sequence at least 16 nucleotides in length that is complementary to, or hybridises under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NO 6 to SEQ ID NO 25, or complements thereof, wherein the treated DNA or a fragment thereof is either amplified to produce one or more amplicates, or is not amplified; and
- d) determining, based on the presence or absence of, or on a property of said amplicate, the methylation state of at least one CpG dinucleotide sequence of SEQ ID NO 1 to SEQ ID NO 5, or an average, or a value reflecting an average methylation state of a plurality of CpG dinucleotide sequences of SEQ ID NO 1 to SEQ ID NO 5,

whereby at least one of detecting or distinguishing between colon cell proliferative disorders is, at least in part, enabled.

15. A method according to claim 13, comprising the following steps of:

- a) obtaining, from a subject, a biological sample having subject genomic DNA;
- b) treating the genomic DNA, or a fragment thereof, with one or more reagents to convert 5-position unmethylated cytosine bases to uracil or to another base that is detectably dissimilar to cytosine in terms of hybridization properties;
- c) amplifying one or more fragments of the treated DNA such that only or preferentially DNA originating from colon or colon cell proliferative disorder cells are amplified; and
- d) detecting the amplicates or characteristics thereof and thereby deducing on the presence or absence of a colon cell proliferative disorder.

16. A method according to claims 13 to 15, wherein said colon cell proliferative disorders are taken from the group comprising adenocarcinomas, polyps, squamous cell cancers, carcinoid tumours, sarcomas and lymphomas.

17. The method of claims 14 to 16, wherein in step a) the biological sample obtained from the subject is selected from the group consisting of histological slides, biopsies, paraffin-embedded tissue, bodily fluid, stool, blood, serum, plasma, urine, sputum and combinations thereof.

18. The method of one of claims 14 to 17, wherein step b) treating the genomic DNA, or the fragment thereof, comprises use of a bisulfite solution.

19. The method of one of claims 14 to 16, wherein treating in b) is subsequent to embedding the DNA in agarose.

20. The method of one of claims 14 to 19, wherein step b) treating the genomic DNA, comprises treating in the presence of at least one of a DNA denaturing agent or a radical scavenger.

21. The method of one of claims 14 to 20, wherein contacting or amplifying in step c) comprises use of at least one method selected from the group consisting of: use of a heat-resistant DNA polymerase as the amplification enzyme; use of a polymerase lacking 5'-3' exonuclease activity; use of a polymerase chain reaction (PCR); generation of an amplicate nucleic acid molecule carrying a detectable labels; and combinations thereof.

22. The method of claim 21, wherein the detectable amplicate label is selected from the label group consisting of: fluorescent labels; radionuclides or radiolabels; amplicate mass labels detectable in a mass spectrometer; detachable amplicate fragment mass labels detectable in a mass spectrometer; amplicate, and detachable amplicate fragment mass labels having a single-positive or single-negative net charge detectable in a mass spectrometer; and combinations thereof.

23. The method of claim 22, comprising in step d) use of mass spectrometry for detecting the amplicate, or detachable amplicate fragment mass labels.

24. The method according to one of claims 14 to 23, wherein in step c) of the method 2 or more different fragments are amplified.

25. The method according to one of claims 14 to 24, wherein one or more of said primers comprise sequences

taken from the group according to SEQ ID NOS:34 to SEQ ID NO:49, SEQ ID NOS:96, 97, 101, 102, 106 and SEQ ID NO:107.

26. The method according to one of claims 14 to 24, wherein one or more of said primers comprise one or more CpG, TpG or CpA dinucleotides.

27. The method of claim 26, wherein said primers comprise between 2 to 5 CpG, TpG or CpA dinucleotides.

28. The method according to one of claim 26 or 27, wherein said one or more CpG, TpG or CpA dinucleotides are located within the 3' half of the primer.

29. The method according to one of claims 26 to 28, wherein said primers comprise one or more bases which hybridise to positions that were converted in the treatment of step b) in claim 14, or of step c) in claim 15.

30. The method of claim 29, wherein at least one of said bases are located within the 3' half of the primer.

31. The method according to one of claims 14 to 30, wherein said amplicates obtained in step d) comprise at least one 20 base pair sequence that comprises 3 or more CpG, TpG or CpA dinucleotides.

32. The method according to one of claims 14 to 31, further comprising in step c) the use of at least one nucleic acid molecule or peptide nucleic acid molecule at least 18 base pairs in length comprising one or more CpG, TpG or CpA dinucleotides and wherein the sequence of said molecule is complementary or identical to a sequence selected from the group consisting of SEQ ID NOS:6 to SEQ ID NO:25, and complements thereof, and wherein said nucleic acid molecule or peptide nucleic acid molecule suppresses amplification of the nucleic acid to which it is hybridised.

33. The method according to claim 32, wherein the sequence of said nucleic acid(s) or peptide nucleic acid(s) is selected from the group consisting SEQ ID NOS:85 to SEQ ID NO:87, SEQ ID NOS:98, 103 and SEQ ID NO:108, and sequences complementary thereto.

34. The method of claim 32, wherein amplification of DNA that was unmethylated prior to treatment of step b) is suppressed.

35. The method of one of claims 32 to 34, wherein said nucleic acid molecule or peptide nucleic acid molecule is in each case modified at the 5'-end thereof to preclude degradation by an enzyme having a 5'-3' exonuclease activity.

36. The method of one of claims 32 to 35, wherein said nucleic acid molecule or peptide nucleic acid molecule in each case lack a 3' hydroxyl group.

37. The method of one of claims 32 to 36, wherein the amplification enzyme is a polymerase lacking 5'-3' exonuclease activity.

38. The method of one of claims 32 to 37, wherein the binding site of the oligonucleotide or PNA oligomer is identical to, or overlaps with that of the primer and thereby hinders hybridisation of the primer to its binding site.

39. The method of one of claims 32 to 38, wherein the binding sites of at least two of the oligonucleotides or PNA oligomers are identical to, or overlap with those of at least two of the primers, and thereby hinder hybridisation of the primers to their binding site.

40. The method of claim 39, wherein hybridisation of at least one of the oligonucleotides or peptide nucleic acid oligomers hinders hybridisation of a forward primer, and the hybridisation of at least one of the oligonucleotides or peptide

nucleic acid oligomers hinders the hybridisation of a reverse primer that binds to the elongation product of said forward primer.

41. The method of one of claims 32 to 38, wherein said oligonucleotide or peptide nucleic acid oligomer hybridises between the binding sites of the forward and reverse primers.

42. The method of one of claim 14 or 15, wherein determining in step d), comprises hybridisation of at least one nucleic acid molecule or peptide nucleic acid molecule in each case comprising a contiguous sequence at least 9 nucleotides in length comprising one or more CpG, TpG or CpA dinucleotides and wherein the sequence of said molecule that is complementary or identical to a sequence selected from the group consisting of SEQ ID NOS:6 to SEQ ID NO:25.

43. The method of claim 42, wherein at least one such hybridising nucleic acid molecule or peptide nucleic acid molecule is bound to a solid phase.

44. The method of claim 43, wherein a plurality of such hybridising nucleic acid molecules or peptide nucleic acid molecules are bound to a solid phase in the form of a nucleic acid or peptide nucleic acid array selected from the array group consisting of linear, hexagonal, rectangular, and combinations thereof.

45. The method of one of claim 14 or 15, wherein determining in step d), comprises sequencing of the amplicate.

46. The method of one of claim 14 or 15, wherein determining in step d), comprises: hybridising at least one nucleic acid molecule comprising a contiguous sequence at least 9 nucleotides in length that is complementary to, or hybridises under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NOS:6 to SEQ ID NO:25, and complements thereof; and extending at least one such hybridised nucleic acid molecule by at least one nucleotide base.

47. The method according to claim 42, wherein the sequence of said nucleic acid(s) or peptide nucleic acid(s) is selected from the group consisting SEQ ID NOS:50 to SEQ ID NO:84, SEQ ID NOS:88 to SEQ ID NO:90, SEQ ID NOS:99, 100, 104, 105, 109, SEQ ID NO:110, and sequences complementary thereto.

48. The method according to claim 42, wherein said oligonucleotides or PNA oligomers are fluorescently labelled, and wherein detection thereof is by either an increase or a decrease in fluorescence or fluorescence polarisation.

49. The method according to claim 42, wherein the hybridisation of the oligonucleotides or PNA oligomers is detectable by fluorescence resonance energy transfer, and wherein the detection is by either an increase or a decrease in fluorescence.

50. The method of one of claim 14 or 15, wherein the background DNA concentration is at between 100 to 1000 fold excess of the concentration of the DNA to be investigated.

51. A method according to claim 13, comprising:

- a) obtaining, from a subject, a biological sample having subject genomic DNA;
- b) extracting the genomic DNA;
- c) contacting the genomic DNA, or a fragment thereof, comprising SEQ ID NOS:1 to SEQ ID NO:5 or a sequence that hybridises under stringent conditions to SEQ ID NOS:1 to SEQ ID NO:5, with one or more methylation-sensitive restriction enzymes, wherein the genomic DNA is either digested thereby to produce digestion fragments, or is not digested thereby; and

- d) determining, based on a presence or absence of, or on property of at least one such fragment, the methylation state of at least one CpG dinucleotide sequence of SEQ ID NOS:1 to SEQ ID NO:5, or an average, or a value reflecting an average methylation state of a plurality of CpG dinucleotide sequences of SEQ ID NOS:1 to SEQ ID NO:5, whereby at least one of detecting the prostate cell proliferative disorder, or distinguishing between a transitional and a peripheral zone of origin of the prostate cell proliferative disorder is, at least in part, afforded.
- 52.** The method of claim **51**, further comprising, prior to determining in step d), amplifying of the digested or undigested genomic DNA.
- 53.** The method of claim **52**, wherein amplifying comprises use of at least one method selected from the group consisting of: use of a heat resistant DNA polymerase as an amplification enzyme; generation of an amplificate nucleic acid carrying a detectable label; and combinations thereof.
- 54.** The method of claim **53**, wherein the detectable amplificate label is selected from the label group consisting of: fluorescent labels; radionuclides or radiolabels; amplificate mass labels detectable in a mass spectrometer; detachable amplificate fragment mass labels detectable in a mass spectrometer; amplificate, and detachable amplificate fragment mass labels having a single-positive or single-negative net charge detectable in a mass spectrometer; and combinations thereof.
- 55.** The method of claim **54**, comprising use of mass spectrometry for detecting amplificate, or detachable amplificate fragment mass labels.
- 56.** The method of claim **55**, wherein the mass spectrometry is selected from the group consisting of matrix assisted laser desorption/ionisation mass spectrometry (MALDI), electron spray mass spectrometry (ESI), and combinations thereof.
- 57.** The method of claim **51**, wherein the biological sample obtained from the subject is selected from the group consisting of histological slides, biopsies, paraffin-embedded tissue, bodily fluid, stool, blood, serum, plasma, urine, sputum and combinations thereof.
- 58.** A kit useful for detecting, differentiating or distinguishing between colon cell proliferative disorders, comprising:
- a bisulfite reagent; and
  - at least one nucleic acid molecule or peptide nucleic acid molecule comprising, in each case a contiguous sequence at least 9 nucleotides in length that is complementary to, or hybridises under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NOS:1 to SEQ ID NO:25, and complements thereof.
- 59.** The kit of claim **58**, further comprising standard reagents for performing a methylation assay selected from the group consisting of MS-SNuPE, MSP, MethylLight, HeavyMethyl, COBRA, nucleic acid sequencing, and combinations thereof.
- 60.** The use of a nucleic acid according to claim **1**, of an oligonucleotide or PNA-oligomer according to one of the claims **2** through **7**, of a kit according to claim **58** or **59**, of an array according to one of the claims **11** through **12**, of a set of oligonucleotides according to one of claims **5** through **7**, or a method according to claims **13** to **57**, for the classification, differentiation and/or diagnosis of colon cell proliferative disorders or the predisposition to colon cell proliferative disorders.
- 61.** The use of a nucleic acid according to claim **1**, of an oligonucleotide or PNA-oligomer according to one of the claims **2** through **7**, of a kit according to claim **58** or **59**, of an array according to one of the claims **11** through **12**, of a set of oligonucleotides according to one of claims **5** through **7**, or a method according to claims **13** to **57**, for the therapy of colon cell proliferative disorders.

\* \* \* \* \*

专利名称(译)	用于分析结肠细胞增殖性疾病的方法和核酸		
公开(公告)号	<a href="#">US20110046005A1</a>	公开(公告)日	2011-02-24
申请号	US12/794449	申请日	2010-06-04
[标]申请(专利权)人(译)	埃皮吉諾米克斯股份公司		
申请(专利权)人(译)	AG EPIGENOMICS		
当前申请(专利权)人(译)	AG EPIGENOMICS		
[标]发明人	ADORJAN PETER BURGER MATTHIAS MAIER SABINE LESCHE RALF COTTRELL SUSAN MOONEY SUZANNE		
发明人	ADORJAN, PETER BURGER, MATTHIAS MAIER, SABINE LESCHE, RALF COTTRELL, SUSAN MOONEY, SUZANNE		
IPC分类号	C40B30/04 C07K2/00 C07H21/04 C40B50/14 C40B40/06 G01N33/53 C12Q1/68 G01N27/62 A61K48/00 A61P35/00 A61P43/00 C12M1/00 C12N15/09 C12Q1/02 G01N21/78 G01N33/543 G01N33/574 G01N33/58 G01N37/00		
CPC分类号	C12Q1/6886 Y10T436/143333 C12Q2600/154 A61P35/00 A61P43/00		
优先权	2002004551 2002-02-27 EP 10/506089 2005-09-06 US PCT/EP2003/002034 2003-02-27 WO		
其他公开文献	US9988683		
外部链接	<a href="#">Espacenet</a> <a href="#">USPTO</a>		

摘要(译)

提供了通过分析Versican, TPEF, H-钙粘着蛋白, 降钙素和EYA4中的一种或多种基因来检测, 区分或区分结肠细胞增殖性疾病的方法和核酸。还提供了可用于所述基因的细胞增殖性疾病特异性分析的新型核酸序列及其方法, 测定和试剂盒。

