



US 20050003463A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2005/0003463 A1**

Adorjan et al.

(43) **Pub. Date: Jan. 6, 2005**

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

(30) **Foreign Application Priority Data**

Feb. 27, 2002 (EP) 02004551.4

(76) Inventors: **Peter Adorjan**, Berlin (DE); **Matthias Burger**, Berlin (DE); **Sabine Maier**, Berlin (DE); **Ralf Lesche**, Berlin (DE); **Susan Cottrell**, Seattle, WA (US); **Theo De Vos**, Seattle, WA (US)

Publication Classification

(51) **Int. Cl.⁷** **C12Q 1/68; G01N 33/574**

(52) **U.S. Cl.** **435/7.23**

Correspondence Address:
DAVIS WRIGHT TREMAINE, LLP
2600 CENTURY SQUARE
1501 FOURTH AVENUE
SEATTLE, WA 98101-1688 (US)

(57) **ABSTRACT**

The invention provides methods and nucleic acids for detecting, differentiating or distinguishing between colon cell proliferative disorders as well as therapy thereof by analysis of the gene EYA4 and its promoter and regulatory sequences. The invention further provides novel nucleic acid sequences useful for the cell proliferative disorder specific analysis of said gene as well as methods, assays and kits thereof.

(21) Appl. No.: **10/482,823**

(22) PCT Filed: **Feb. 13, 2003**

(86) PCT No.: **PCT/EP03/01457**

FIGURE 1

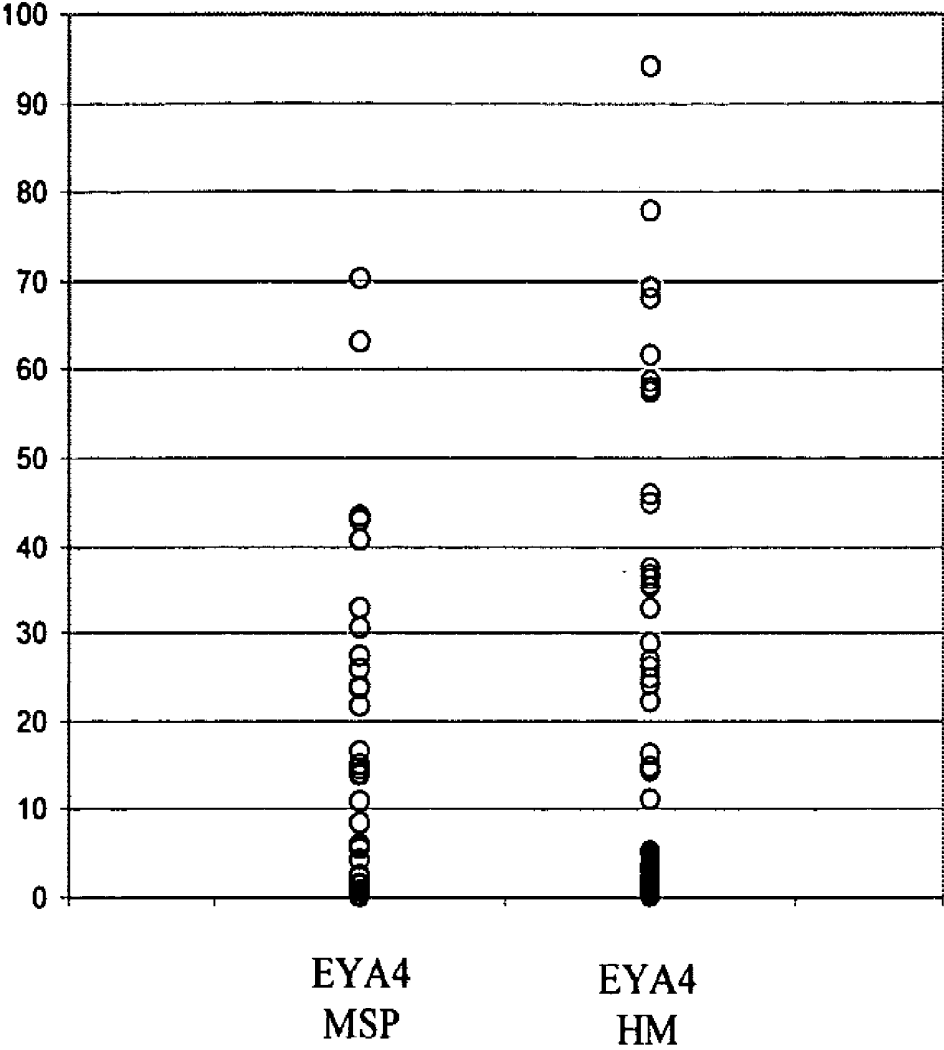


FIGURE 2

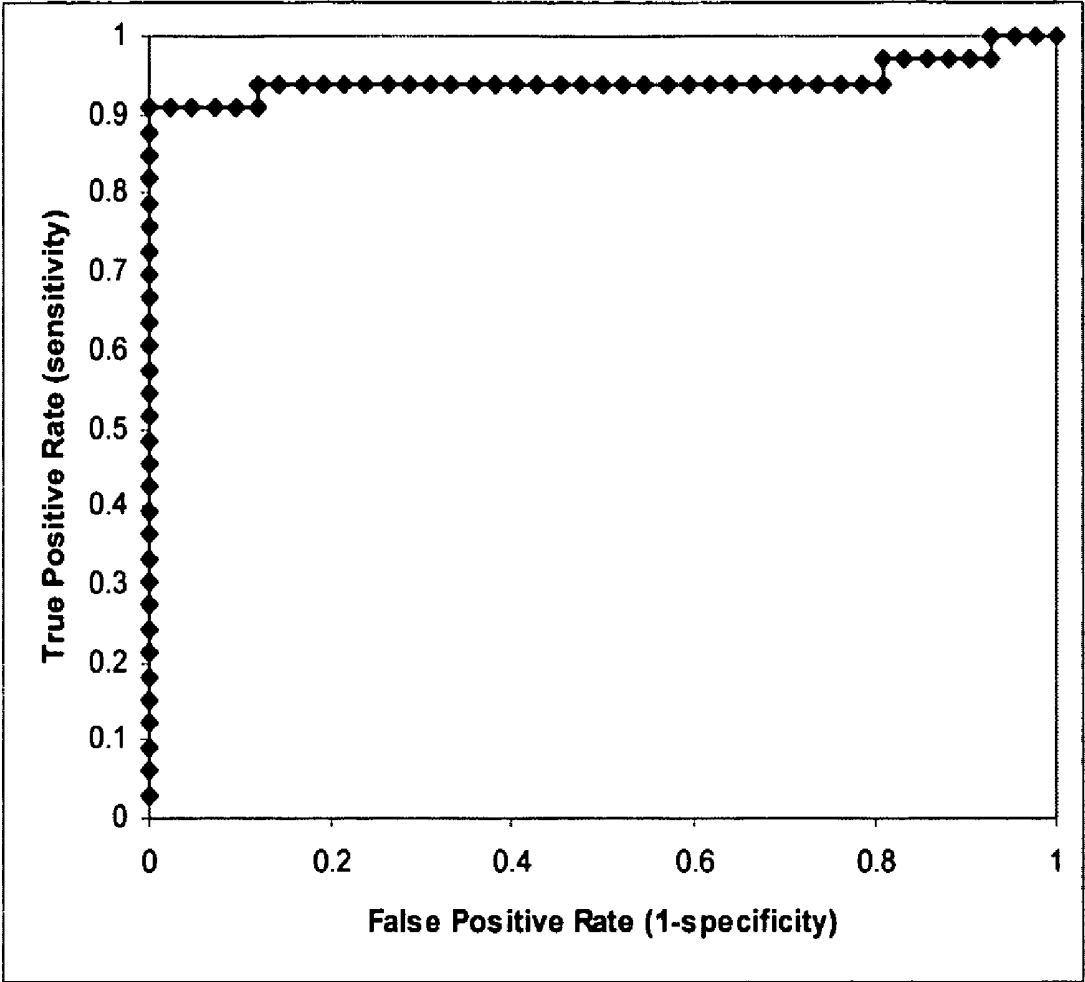


FIGURE 3

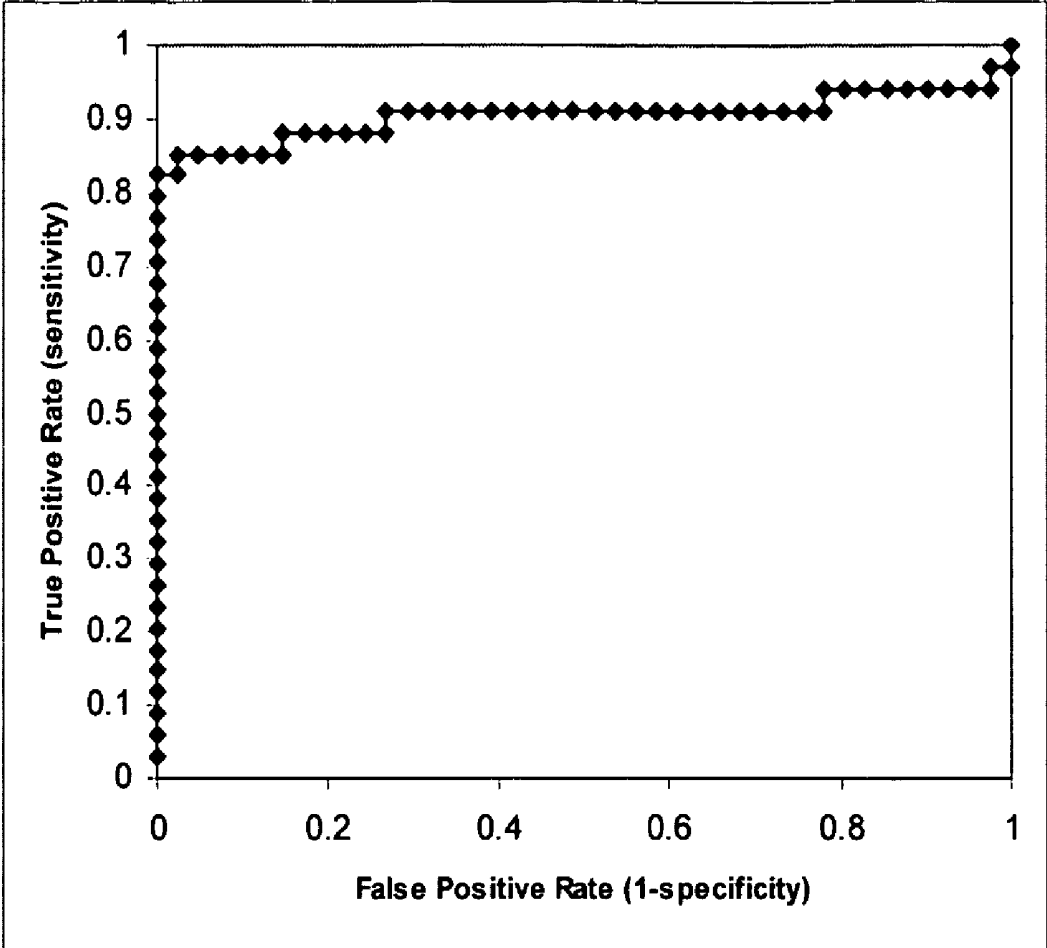


FIGURE 4

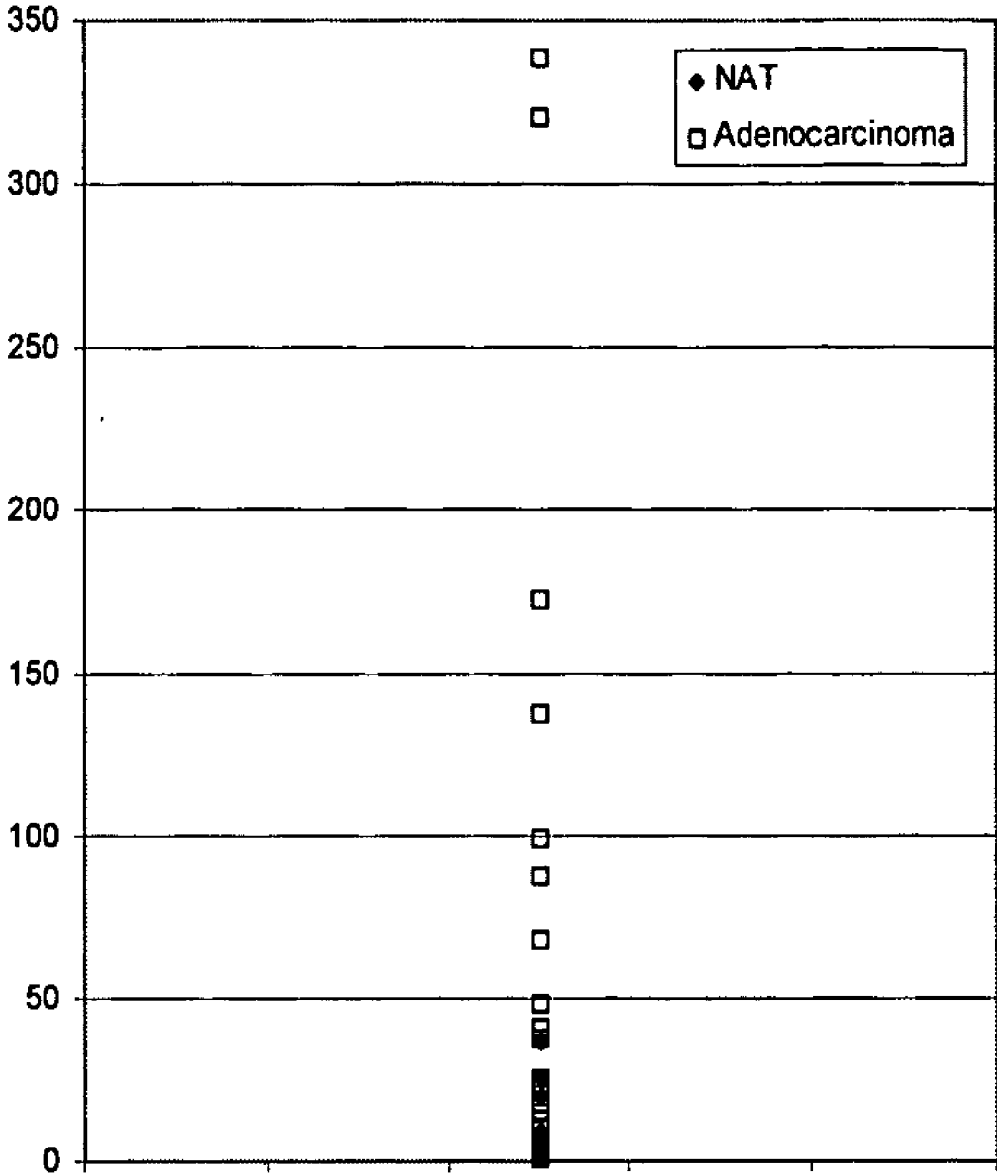


FIGURE 5

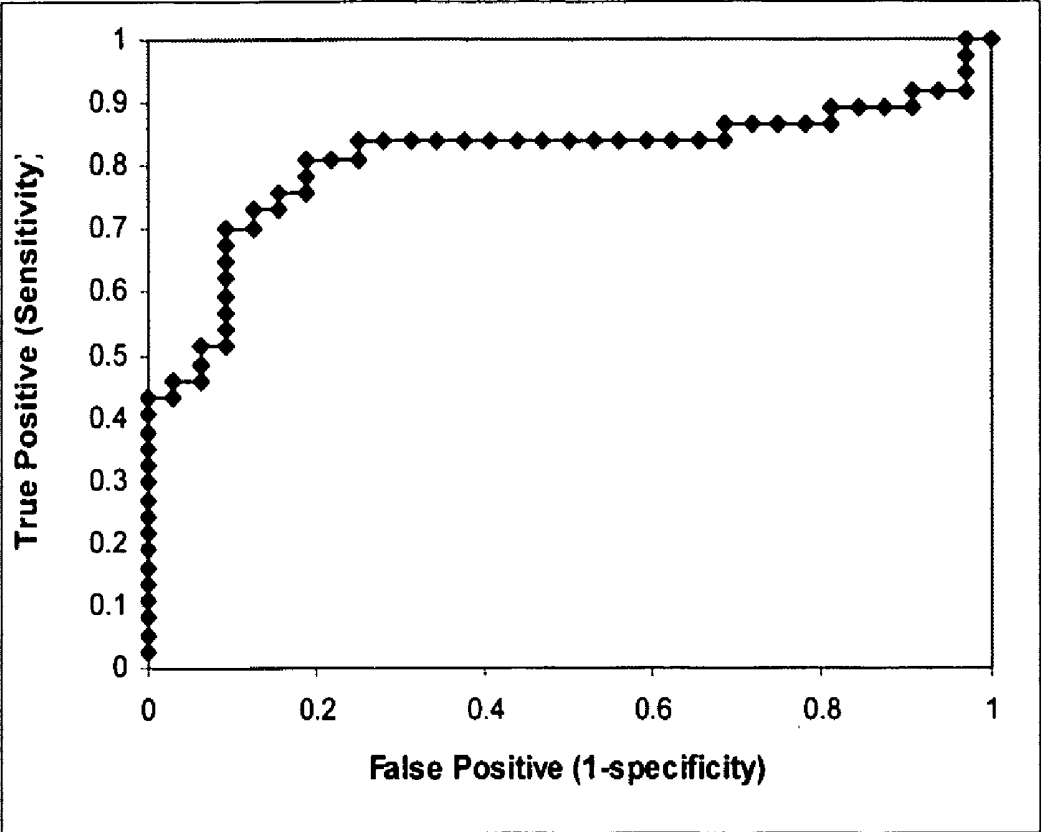


FIGURE 6

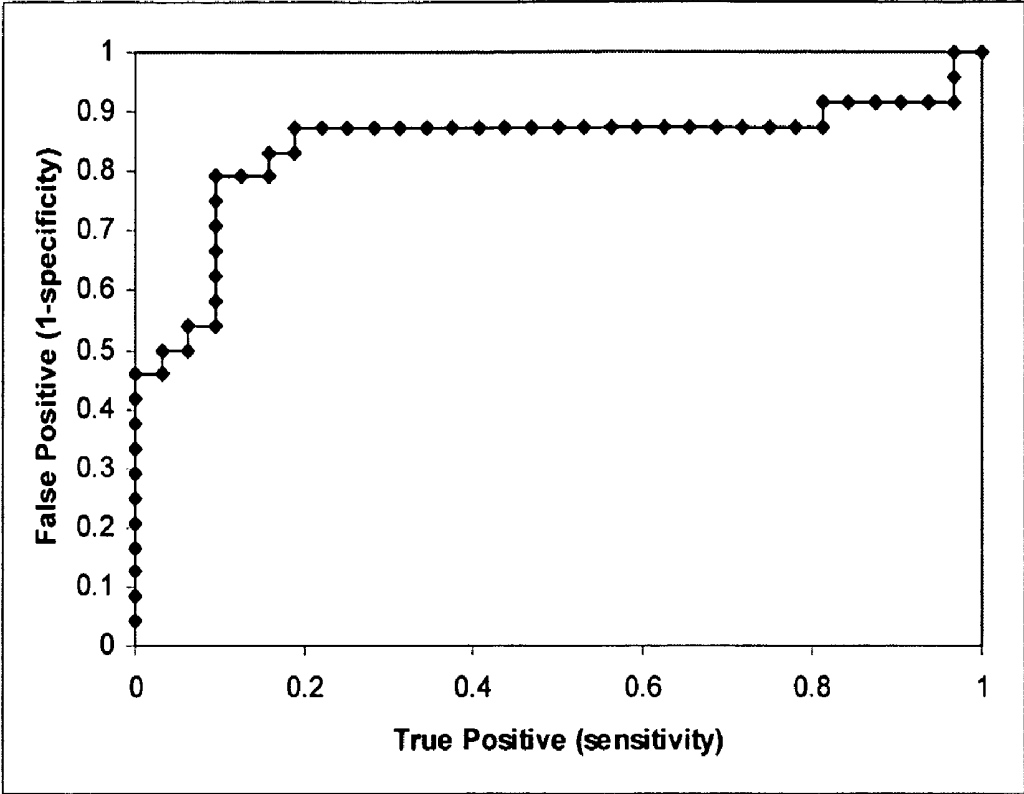


FIGURE 7

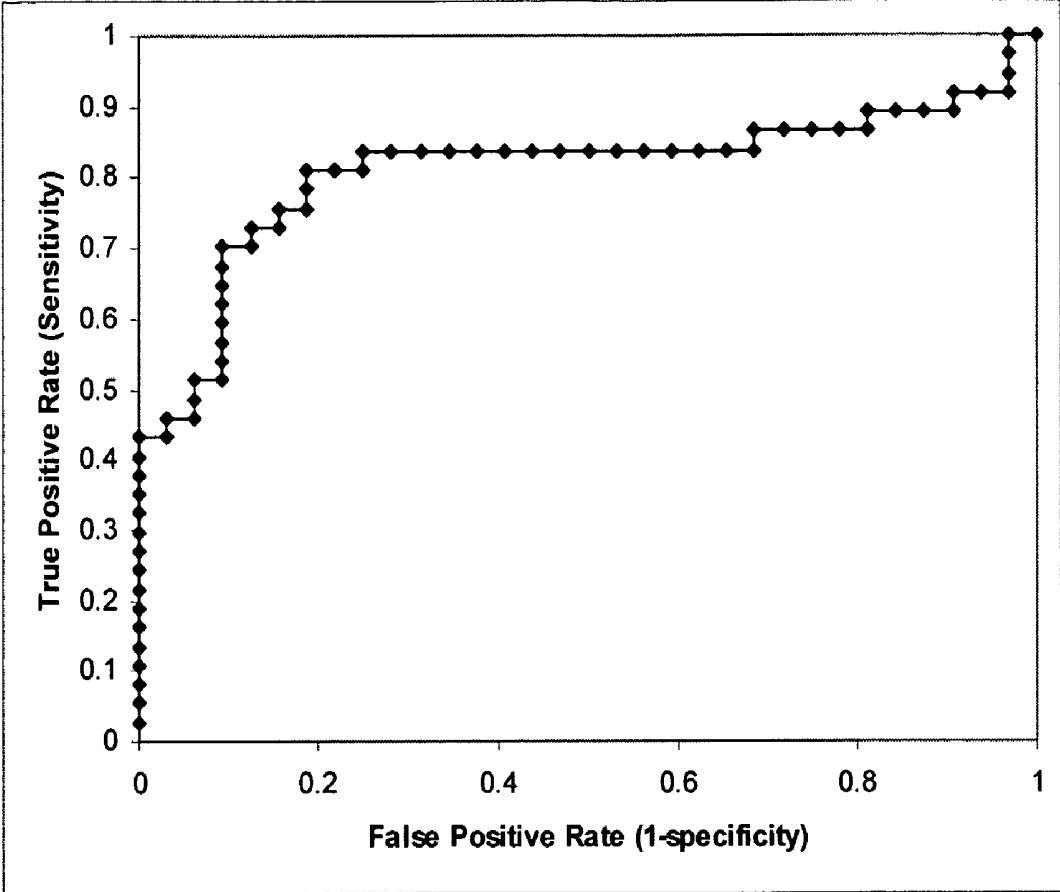


FIGURE 8

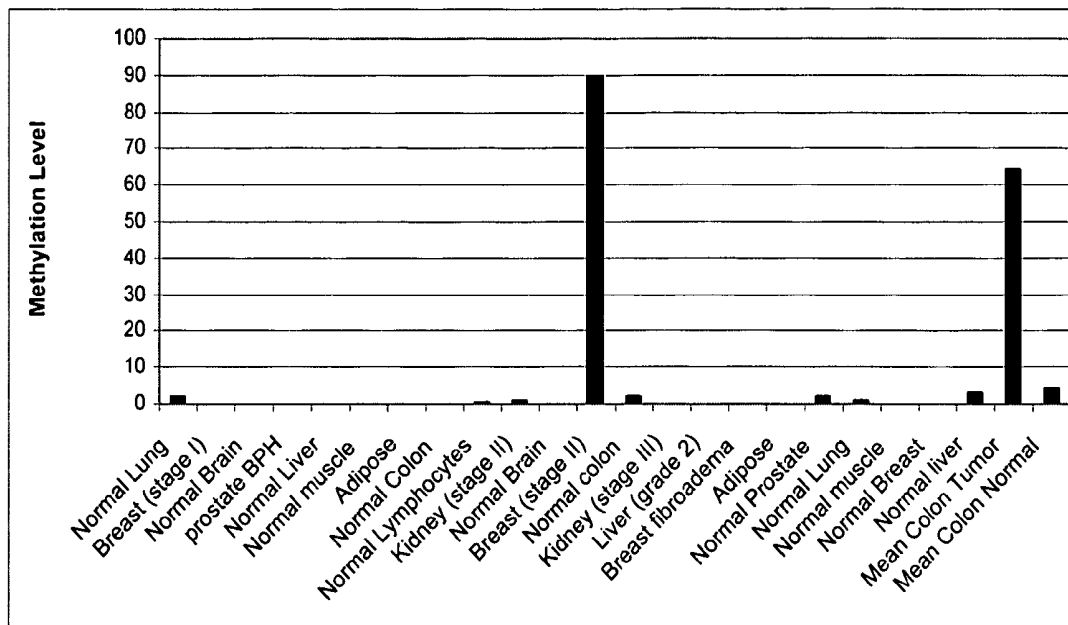


FIGURE 9

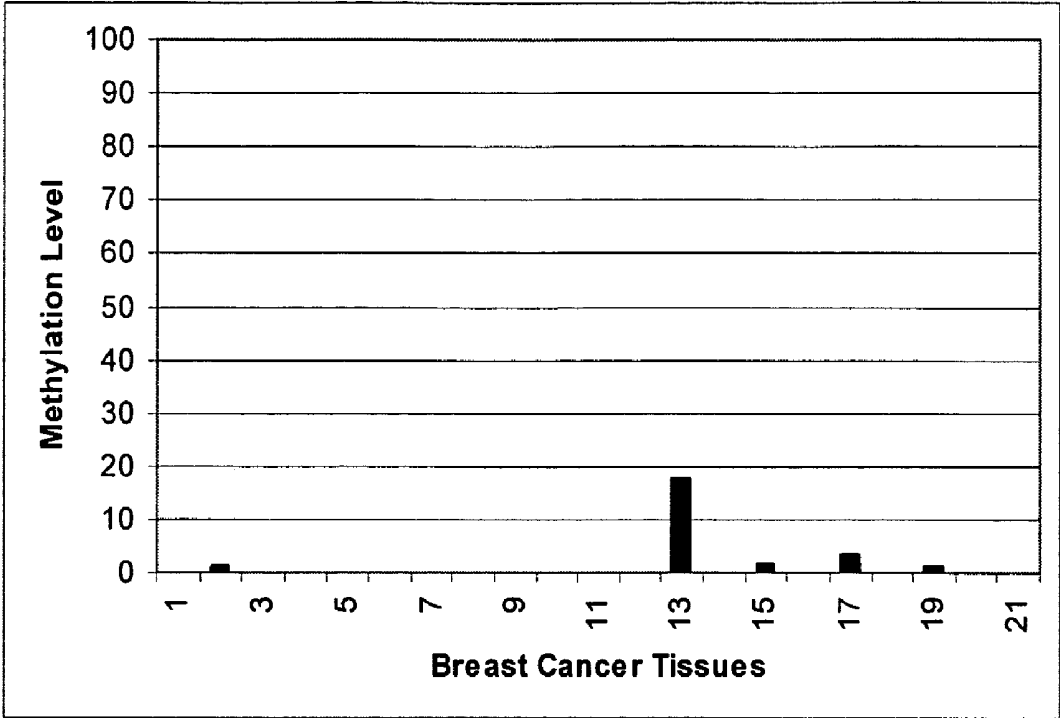
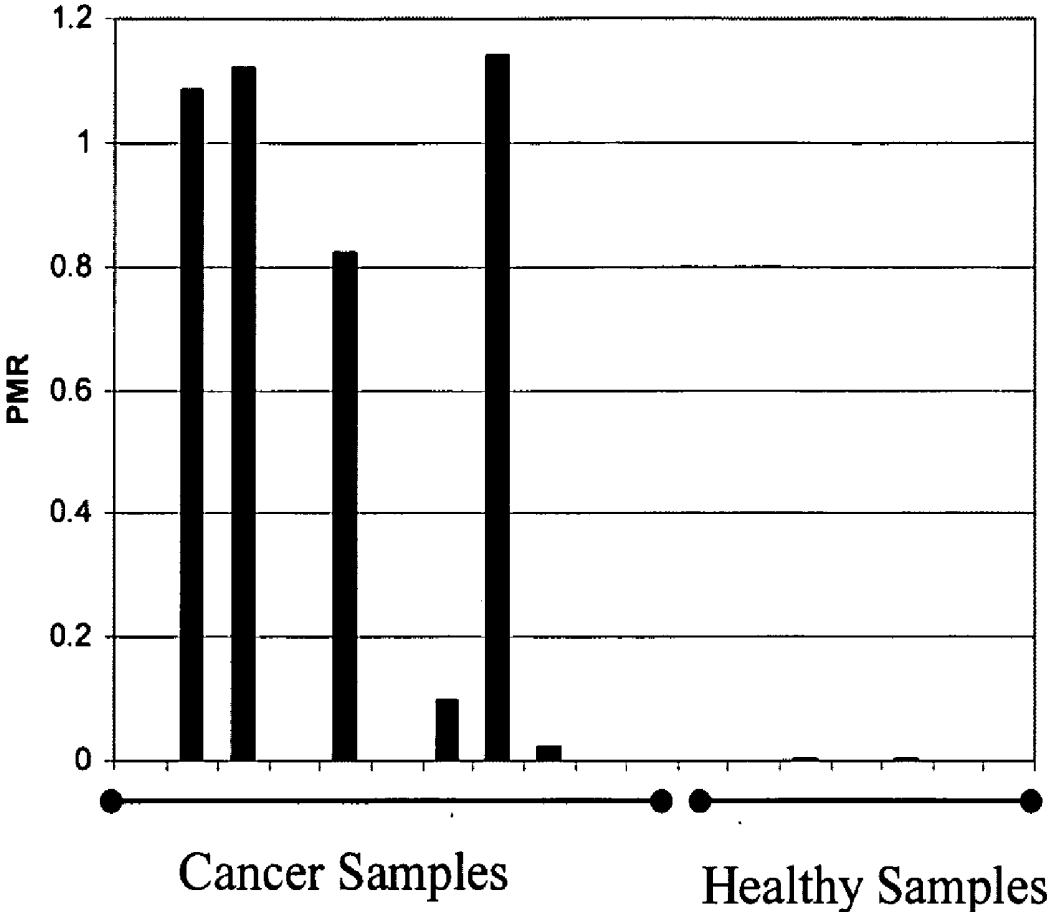


FIGURE 10



**METHOD AND NUCLEIC ACIDS FOR THE
ANALYSIS OF COLORECTAL CELL
PROLIFERATIVE DISORDERS**

FIELD OF THE INVENTION

[0001] Colorectal 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.

[0002] 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.

[0003] 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 tumour suppressor genes accumulate.

[0004] 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.

[0005] In the molecular evolution of colorectal cancer, DNA methylation errors have been suggested to play two distinct roles. In normal colonic mucosal 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 characterised by microsatellite 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.

[0006] 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.

[0007] The identification of markers for the differentiation of colon carcinoma as well as for early detection are main goals of current research.

[0008] 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 embryonic 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 (Borsani G, et al., EYA4, a novel vertebrate gene related to *Drosophila* eyes absent. *Hum Mol Genet* 1999 January;8(1):11-23). 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 NG, DeClau F, Chen N, Verhoeven K, Prasad S, Tranebjerg L, Morton CC, 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.

[0009] 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.

[0010] 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.

[0011] 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.

[0012] 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 ML, Jones PA. 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 hybridisation has also been described (Olek et al., WO 99/28498).

[0013] 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.

[0014] 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.

[0015] Fluorescently labelled probes are often used for the scanning of immobilised 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 hybridised probes may be carried out, for example via a confocal microscope. Cy3 and Cy5 dyes, besides many others, are commercially available.

[0016] Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-TOF) is a very efficient development for the analysis of biomolecules (Karas M, Hillenkamp F. Laser desorption ionization 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 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.

[0017] 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 Ionization 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 IG, 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.

[0018] 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.

DESCRIPTION

[0019] The present invention discloses novel methods for the detection of cell proliferative disorders. Said invention discloses the use of the gene EYA4, as well as its promoter and regulatory elements as a marker for colon cell proliferative disorders. More specifically, the disclosed matter shows the applicability of said gene to the detection of colon cell proliferative disorders, distinguishing between different classes of colon cell proliferative disorders as well as the differentiation of colon cell proliferative disorders from cell proliferative disorders originating from other tissues.

[0020] 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, as well as therapeutic and diagnostic methods thereof.

[0021] In one embodiment the method discloses the use of the gene EYA4 as a marker for the differentiation, detection and distinguishing of colon cell proliferative disorders. Said use of the gene may be enabled by means of any analysis of the expression of the gene, by means of mRNA expression analysis or protein expression analysis. However, in the most preferred embodiment of the invention, the detection, differentiation and distinguishing of colon cell proliferative disorders is enabled by means of analysis of the methylation status of the gene EYA4 and its promoter or regulatory elements.

[0022] To detect the presence of mRNA encoding EYA4 in a detection system for colon cancer, a sample is obtained from a patient. The sample can be a tissue biopsy sample or a sample of blood, plasma, serum or the like. The sample may be treated to extract the nucleic acids contained therein. The resulting nucleic acid from the sample is subjected to gel electrophoresis or other separation techniques. Detection involves contacting the nucleic acids and in particular the mRNA of the sample with a DNA sequence serving as a probe to form hybrid duplexes. The stringency of hybridisation is determined by a number of factors during hybridisation and during the washing procedure, including temperature, ionic strength, length of time and concentration of formamide. These factors are outlined in, for example, Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 2d ed., 1989). Detection of the resulting duplex is usually accomplished by the use of labelled probes. Alternatively, the probe may be unlabeled, but may be detectable by specific binding with a ligand which is labelled, either directly or indirectly. Suitable labels and methods for labelling probes and ligands are known in the art, and include, for example, radioactive labels which may be incorporated by known methods (e.g., nick translation or kinasings), biotin,

fluorescent groups, chemiluminescent groups (e.g., dioxetanes, particularly triggered dioxetanes), enzymes, antibodies, and the like.

[0023] In order to increase the sensitivity of the detection in a sample of mRNA encoding EYA4, the technique of reverse transcription/polymerisation chain reaction can be used to amplify cDNA transcribed from mRNA encoding EYA4. The method of reverse transcription/PCR is well known in the art (for example, see Watson and Fleming, *supra*).

[0024] The reverse transcription/PCR method can be performed as follows. Total cellular RNA is isolated by, for example, the standard guanidium isothiocyanate method and the total RNA is reverse transcribed. The reverse transcription method involves synthesis of DNA on a template of RNA using a reverse transcriptase enzyme and a 3' end primer. Typically, the primer contains an oligo(dT) sequence. The cDNA thus produced is then amplified using the PCR method and EYA4 specific primers. (Belyavsky et al, *Nucl Acid Res* 17:2919-2932, 1989; Krug and Berger, *Methods in Enzymology*, Academic Press, N.Y., Vol.152, pp. 316-325, 1987 which are incorporated by reference)

[0025] The present invention may also be described in certain embodiments as a kit for use in detecting a colon cancer disease state through testing of a biological sample. A representative kit may comprise one or more nucleic acid segments as described above that selectively hybridise to EYA4 mRNA and a container for each of the one or more nucleic acid segments. In certain embodiments the nucleic acid segments may be combined in a single tube. In further embodiments, the nucleic acid segments may also include a pair of primers for amplifying the target mRNA. Such kits may also include any buffers, solutions, solvents, enzymes, nucleotides, or other components for hybridisation, amplification or detection reactions. Preferred kit components include reagents for reverse transcription-PCR, *in situ* hybridisation, Northern analysis and/or RPA

[0026] The present invention further provides for methods to detect the presence of the polypeptide, EYA4, in a sample obtained from a patient. Any method known in the art for detecting proteins can be used. Such methods include, but are not limited to immunodiffusion, immunoelectrophoresis, immunochemical methods, binder-ligand assays, immunohistochemical techniques, agglutination and complement assays. (for example see *Basic and Clinical Immunology*, Sites and Terr, eds., Appleton & Lange, Norwalk, Conn. pp 217-262, 1991 which is incorporated by reference). Preferred are binder-ligand immunoassay methods including reacting antibodies with an epitope or epitopes of EYA4 and competitively displacing a labelled EYA4 protein or derivative thereof.

[0027] Certain embodiments of the present invention comprise the use of antibodies specific to the polypeptide encoded by the EYA4 gene. Such antibodies may be useful for diagnostic and prognostic applications in detecting the disease state, by comparing a patient's levels of colon disease marker expression to expression of the same markers in normal individuals. In certain embodiments production of monoclonal or polyclonal antibodies can be induced by the use of the EYA4 polypeptide as antigen. Such antibodies may in turn be used to detect expressed proteins as markers for human disease states. The levels of such proteins present

in the peripheral blood or prostate tissue sample of a patient may be quantified by conventional methods. Antibody-protein binding may be detected and quantified by a variety of means known in the art, such as labelling with fluorescent or radioactive ligands. The invention further comprises kits for performing the above-mentioned procedures, wherein such kits contain antibodies specific for the EYA4 polypeptides.

[0028] Numerous competitive and non-competitive protein binding immunoassays are well known in the art. Antibodies employed in such assays may be unlabeled, for example as used in agglutination tests, or labelled for use a wide variety of assay methods. Labels that can be used include radionuclides, enzymes, fluorescers, chemiluminescers, enzyme substrates or co-factors, enzyme inhibitors, particles, dyes and the like for use in radioimmunoassay (RIA), enzyme immunoassays, e.g., enzyme-linked immunosorbent assay (ELISA), fluorescent immunoassays and the like. Polyclonal or monoclonal antibodies to EYA4 or an epitope thereof can be made for use in immunoassays by any of a number of methods known in the art. One approach for preparing antibodies to a protein is the selection and preparation of an amino acid sequence of all or part of the protein, chemically synthesising the sequence and injecting it into an appropriate animal, usually a rabbit or a mouse (Milstein and Kohler *Nature* 256:495-497, 1975; Gutfre and Milstein, *Methods in Enzymology: Immunochemical Techniques* 73:1-46, Langone and Banatis eds., Academic Press, 1981 which are incorporated by reference). Methods for preparation of EYA4 or an epitope thereof include, but are not limited to chemical synthesis, recombinant DNA techniques or isolation from biological samples.

[0029] The invention provides significant improvements over the state of the art in that there are currently no markers used to detect colon 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 methods, 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 highly sensitive assay techniques.

[0030] The objective of the invention can be achieved by analysis of the methylation state of the CpG dinucleotides within the genomic sequence according to SEQ ID NO: 1 and sequences complementary thereto. SEQ ID NO: 1 discloses the gene EYA4 and its promoter and regulatory elements, wherein said fragment comprises CpG dinucleotides exhibiting a disease specific methylation pattern. The methylation pattern of the gene EYA4 and its promoter and regulatory elements have heretofore not been analysed with regard to cell proliferative disorders. Due to the degeneracy of the genetic code, the sequence as identified in SEQ ID NO: 1 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 EYA4.

[0031] 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 NO: 2 to SEQ ID NO: 5 and sequences complementary thereto.

[0032] The sequences of SEQ ID NOS: 2 to 5 provide modified versions of the nucleic acid according to SEQ ID NO: 1, 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 NO: 1 as follows. (see also the following TABLE 1): SEQ ID NO: 1, sense DNA strand of EYA4 gene and its promoter and regulatory elements; SEQ ID NO: 2, converted SEQ ID NO: 1, wherein "C" or "T," but "cp." remains "cp." (i.e., corresponds to case where, for SEQ ID NO: 1, all "C" residues of cp. dinucleotide sequences are methylated and are thus not converted); SEQ ID NO: 3, complement of SEQ ID NO: 1, wherein "C" to "T," but "cp." remains "cp." (i.e., corresponds to case where, for the complement (antisense strand) of SEQ ID NO: 1, all "C" residues of cp. dinucleotide sequences are methylated and are thus not converted); SEQ ID NO: 4, converted SEQ ID NO: 1, wherein "C" to "T" for all "C" residues, including those of "cp." dinucleotide sequences (i.e., corresponds to case where, for SEQ ID NO: 1, all "C" residues of cp. dinucleotide sequences are unmethylated); SEQ ID NO: 5, complement of SEQ ID NO: 1, wherein "C" to "T" for all "C" residues, including those of "CpG" dinucleotide sequences (i.e., corresponds to case where, for the complement (antisense strand) of SEQ ID NO: 1, all "C" residues of CpG dinucleotide sequences are unmethylated).

TABLE 1

Description of SEQ ID NOS: 1 to 5		
SEQ ID NO	Relationship to SEQ ID NO: 1	Nature of cytosine base conversion
SEQ ID NO: 1	Sense strand (EYA4 gene including promoter and regulatory elements)	None; untreated sequence
SEQ ID NO: 2	Converted sense strand	"C" to "T," but "CpG" remains "CpG" (all "C" residues of CpGs are methylated)
SEQ ID NO: 3	Converted anti-sense strand	"C," to "T," but "CpG" remains "CpG" (all "C" residues of CpGs are methylated)
SEQ ID NO: 4	Converted sense strand	"C," to "T" for all "C" residues (all "C" residues of CpGs are <u>un</u> methylated)
SEQ ID NO: 5	Converted anti-sense strand	"C" to "T" for all "C" residues (all "C" residues of CpGs are <u>un</u> methylated)

[0033] Significantly, heretofore, the nucleic acid sequences and molecules according to SEQ ID NO: 1 to SEQ ID NO: 5 were not implicated in or connected with the ascertainment of colon cell proliferative disorders.

[0034] The described invention further disclose an oligonucleotide or oligomer for detecting the cytosine methylation state within pretreated DNA, according to SEQ ID NO: 2 to SEQ ID NO: 5. Said oligonucleotide or oligomer comprising a nucleic acid sequence having a length of at least nine (9) nucleotides which hybridises, under moderately stringent or stringent conditions (as defined herein above), to a pretreated nucleic acid sequence according to SEQ ID NO: 2 to SEQ ID NO: 5 and/or sequences complementary thereto.

[0035] 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: 2 to 5, 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.

[0036] 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: 2 to 5, or to the complements thereof.

[0037] 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.

[0038] For target sequences that are related and substantially identical to the corresponding sequence of SEQ ID NO: 1 (such as EYA41 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.

[0039] Examples of inventive oligonucleotides of length X (in nucleotides), as indicated by polynucleotide positions with reference to, e.g., SEQ ID NO: 1, 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:

[0040] n to (n+(X-1));

[0041] where n=1, 2, 3, . . . (Y-(X-1));

[0042] where Y equals the length (nucleotides or base pairs) of SEQ ID NO: 1;

[0043] 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

[0044] 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 NO: 1, where X=20.

[0045] Preferably, the set is limited to those oligomers that comprise at least one CpG, TpG or CpA dinucleotide.

[0046] The present invention encompasses, for each of SEQ ID NOS: 2 to 5 (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.

[0047] 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 NO: 1. Preferred sets of such oligonucleotides or modified oligonucleotides of length X are those consecutively overlapping sets of oligomers corresponding to SEQ ID NOS: 1-5 (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 NO: 11 to SEQ ID NO: 15.

[0048] Particularly preferred oligonucleotides or oligomers according to the present invention are those in which the cytosine of the CpG dinucleotide (or of the corresponding converted TpG or 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.

[0049] 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), palmityl 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 hybridisation-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.

[0050] The oligonucleotide may also comprise at least one art-recognized modified sugar and/or base moiety, or may comprise a modified backbone or non-natural internucleoside linkage.

[0051] The oligomers according to the present invention are normally used in so called "sets" which contain at least

one oligomer for analysis of each of the CpG dinucleotides of a genomic sequence comprising SEQ ID NO: 1 and sequences complementary thereto or to their corresponding CG, TG or CA dinucleotide within the pretreated nucleic acids according to SEQ ID NO: 2 to SEQ ID NO: 5 and sequences complementary thereto. Preferred is a set which contains at least one oligomer for each of the CpG dinucleotides within the gene EYA4 and its promoter and regulatory elements in both the pretreated and genomic versions of said gene, SEQ ID NO: 2 to 5 and SEQ ID NO: 1, 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 more-over relates to a set of at least 3 n (oligonucleotides and/or PNA-oligomers) used for detecting the cytosine methylation state in pretreated genomic DNA (SEQ ID NO: 2 to SEQ ID NO: 5 and sequences complementary thereto) and genomic DNA (SEQ ID NO: 1 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 NO: 2 to SEQ ID NO: 5, and sequences complementary thereto) and genomic DNA (SEQ ID NO: 1, and sequences complementary thereto).

[0052] 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 NO: 1 to SEQ ID NO: 5 and sequences complementary thereto, or segments thereof.

[0053] 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.

[0054] 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 characterised 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.

[0055] 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.

[0056] 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.

[0057] The described invention further provides a composition of matter useful for detecting, differentiation and

distinguishing between colon cell proliferative disorders. Said composition comprising at least one nucleic acid 18 base pairs in length of a segment of the nucleic acid sequence disclosed in SEQ ID NO: 2 to 5, and one or more substances taken from the group comprising:

[0058] 1-5 mM Magnesium Chloride, 100-500 μ M dNTP, 0.5-5 units of taq polymerase, bovine serum albumen, an oligomer in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer, said oligomer comprising in each case 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 NO: 2 to SEQ ID NO: 5 and sequences complementary thereto. It is preferred that said composition of matter comprises a buffer solution appropriate for the stabilisation of said nucleic acid in an aqueous solution and enabling polymerase based reactions within said solution. Suitable buffers are known in the art and commercially available.

[0059] The present invention further provides a method for conducting an assay in order to ascertain genetic and/or epigenetic parameters of the gene EYA4 and its promoter and regulatory elements. Most preferably the assay according to the following method is used in order to detect methylation within the gene 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 comprising contacting a nucleic acid 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.

[0060] 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, plasma, serum, stool, urine, blood, and combinations thereof. Preferably, the source is biopsies, bodily fluids, urine, or blood.

[0061] 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.

[0062] 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 behaviour. This will be understood as 'pretreatment' herein.

[0063] 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.

[0064] 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 amplificates 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 six 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: 2 to SEQ ID NO: 5 and sequences complementary thereto.

[0065] In an alternate embodiment of the method, the methylation status of preselected CpG positions within the nucleic acid sequences comprising SEQ ID NO: 2 to SEQ ID NO: 5 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 hybridises 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 hybridises to a pretreated nucleic acid sequence according to SEQ ID NO: 2 to SEQ ID NO: 5 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 4 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 hybridising 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.

[0066] In one embodiment of the method the primers may be selected from the group consisting of SEQ ID NO: 6 to SEQ ID NO: 10.

[0067] 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 amplificates 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).

[0068] 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.

[0069] 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.

[0070] 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 hybridise 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: 2 to 5, 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 NO: 15 and SEQ ID NO: 16 and sequences complementary thereto.

[0071] 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 derivitised 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.

[0072] 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'-termini 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 hybridised blocker oligonucleotide.

[0073] 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.

[0074] 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.

[0075] 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.

[0076] It is particularly preferred that the blocking oligonucleotides are present in at least 5 times the concentration of the primers.

[0077] 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.

[0078] 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.

[0079] 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.

[0080] 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 NO: 2 to SEQ ID NO: 5; and the segment comprises at least one CpG, TpG or CpA dinucleotide.

[0081] 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 NO: 1, and the equivalent positions within SEQ ID NOS: 2 to 5. Said oligonucleotides may also be present 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.

[0082] 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).

[0083] 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 non-extendible 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 MethylLight 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.

[0084] 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 NO: 2 to SEQ ID NO: 5.

[0085] In yet a further embodiment of the method, the fourth step of the method comprises sequencing and subsequent sequence analysis of the amplicate generated in the third step of the method (Sanger F., et al., *Proc Natl Acad Sci USA* 74:5463-5467, 1977).

[0086] Additional embodiments of the invention provide a method for the analysis of the methylation status of genomic DNA according to the invention (SEQ ID NO: 1) without the need for pretreatment.

[0087] 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.

[0088] 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.

[0089] 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.

[0090] 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.

[0091] 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.

[0092] 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 EYA4 gene and its 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.

[0093] Specifically, the present invention provides for diagnostic and/or prognostic cancer assays based on measurement of differential methylation of EYA4 CpG dinucleotide sequences. Preferred gene sequences useful to measure such differential methylation are represented herein by SEQ ID NOS: 1 to 5. 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 EYA4-specific CpG dinucleotide sequences derived from the tissue sample, relative to a control sample, and making a diagnosis or prognosis based thereon.

[0094] In particular preferred embodiments, inventive oligomers are used to assess EYA4 specific CpG dinucleotide methylation status, such as those based on SEQ ID NOS: 1 to 5, including the representative preferred oligomers corresponding to SEQ ID NOS: 11 to 15, 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.

[0095] 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 5 is used for manufacturing it, possibly together with suitable additives and ancillary agents. In one embodiment, the EYA4 polypeptide or a fragment or derivative thereof may be administered to a subject to treat or prevent colon cancers.

[0096] In another embodiment, a vector capable of expressing EYA4, or a fragment or a derivative thereof, may also be administered to a subject to treat or prevent colon cancers.

[0097] In another embodiment, agonists which are specific for EYA4 may be used to stimulate or prolong the activity of EYA4 and may be administered to a subject to treat or prevent colon cancers.

[0098] In other embodiments, any of the therapeutic proteins or vectors described above may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of colon cancer. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

[0099] Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population.

[0100] An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier. Such pharmaceutical compositions may consist of EYA4 or agonists of EYA4. The compositions may be administered alone or in combination with at least one other agent, such as stabilising compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

[0101] The pharmaceutical compositions utilised in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

[0102] In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration

may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

[0103] 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 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.

[0104] 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.

[0105] 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); TaqMan® probes; optimised PCR buffers and deoxynucleotides; and Taq polymerase.

[0106] 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.

[0107] 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), optimised PCR buffers and deoxynucleotides, and specific probes.

[0108] Definitions:

[0109] In the context of the present invention, 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.

[0110] In the context of the present invention, 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.”

[0111] In the context of the present invention, 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'-CCMGG-3' (top strand); 3'-GGCC-5' (bottom strand)).

[0112] In the context of the present invention, 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.

[0113] In the context of the present invention, 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.

[0114] In the context of the present invention, 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.

[0115] “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).

[0116] “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.

[0117] In the context of the present invention, 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.

[0118] In the context of the present invention, 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.

[0119] In the context of the present invention, 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 Gonzalzo et al., *Cancer Research* 57:594-599, 1997.

[0120] In the context of the present invention, 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.

[0121] In the context of the present invention, 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.

[0122] The term “Ms-SNuPE” (Methylation-sensitive Single Nucleotide Primer Extension) refers to the art-recognized assay described by Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997.

[0123] 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.

[0124] 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.

[0125] 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.

[0126] “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.

[0127] “Background DNA” as used herein refers to any nucleic acids which originate from sources other than colon cells.

[0128] The invention will now be described in more detail based on the following examples, SEQ IDs, and Figures, without being limited thereto. In the sequence protocol and the Figures,

[0129] SEQ ID NO: 1 shows the sequence of the human gene EYA4,

[0130] SEQ ID NOS: 2 to 5 show chemically pretreated sequences of the gene EYA4,

[0131] SEQ ID NOS: 6 to 10 show the sequences of primers used in the examples, and

[0132] SEQ ID NOS: 11 to 15 show sequences of probes used in the examples.

[0133] FIG. 1 shows the level of methylation determined by a MSP MethylLight assay and by a HeavyMethyl MethylLight assay according to examples 1 and 2. The Y-axis shows the degree of methylation within the region of the EYA4 gene investigated. Tumour samples are represented by white points, and normal colon tissue samples by white black points. A significantly higher degree of methylation was observed in tumour samples than in healthy tissue samples. The level of significance as measured using a t-test was $p=0.00000312$ (MSP-ML, Example 1) and $p=0.000000326$ (HM-ML, Example 2).

[0134] FIG. 2 shows the Receiver Operating Characteristic curve (ROC curve) of the MSP-Methyl-Light-Assay for Adenocarcinomas according to Example 1. A ROC is a plot of the true positive rate against the false positive rate for the different possible cutpoints 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) The AUC for the MSP-Methyl-Light-Assay is: 0.94.

[0135] FIG. 3 shows the Receiver Operating Characteristic curve (ROC curve) of the HM-Methyl-Light-Assay for Adenocarcinoma according to Example 2. 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: 0.91.

[0136] FIG. 4 shows the level of methylation determined by a 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 tumour samples than in healthy tissue samples. The level of significance as measured using a t-test was 0.00424.

[0137] FIG. 5 shows the Receiver Operating Characteristic curve (ROC curve) of the 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.

[0138] FIG. 6 shows the Receiver Operating Characteristic curve (ROC curve) of the 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.

[0139] FIG. 7 shows the Receiver Operating Characteristic curve (ROC curve) of the 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.

[0140] FIG. 8 shows the level of methylation in different tumour and healthy tissues determined by a 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.

[0141] FIG. 9 shows the level of methylation in different breast cancer tissues determined by a HeavyMethyl MethylLight assay according to example 3. Only one tissue was methylated.

[0142] FIG. 10 shows the level of methylation in serum samples determined by a HeavyMethyl MethylLight assay according to example 4. The Y-axis shows the degree of methylation within the region of the EYA4 gene investigated.

EXAMPLE 1

[0143] Analysis of Methylation within Colon Cancer Using an MSP—MethylLight Assay

[0144] 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.

[0145] 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.

[0146] Methods: The EYA4 gene CpG island assay was performed using the following primers and probes:

Forward Primer:
CGGAGGGTACGGAGATTACG; (SEQ ID NO: 6)

Reverse Primer:
CGACGACGCGCGAAA; (SEQ ID NO: 7)

and

Probe:
CGAAACCCCTAAATATCCCGAATAACGCCG. (SEQ ID NO: 12)

[0147] The corresponding control assay was performed using the following primers and probes:

Primer:
TGGTGATGGAGGAGGTTTAGTAAGT; (SEQ ID NO: 8)

Primer:
AACCAATAAAACCTACTCCTCCCTTAA; (SEQ ID NO: 9)

-continued

and

Probe:
ACCACCACCAACACACAATAACAAACACA (SEQ ID NO: 13)

[0148] 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 μ M dNTPs; 7 μ l of DNA, in a final reaction volume of 20 μ l);

[0149] Cycling conditions: (95° C. for 10 minutes; then 50 cycles of: 95° C. for 15 seconds; 60° C. for 1 minute). The data was analysed using a PMR calculation previously described in the literature (Eads et al 2001).

[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 ($p=0.0000312$). The results are shown in **FIG. 1**.

[0151] 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 cutpoints 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, Academic Press, New York, 1975). The AUC for the MSP-Methyl-Light-Assay is: 0.94 (**FIG. 2**).

EXAMPLE 2

[0152] Methylation within Colon Cancer was Analyzed using a HeavyMethyl MethylLight Assay.

[0153] 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.

[0154] Methods. The CpG island assay (methylation assay) was performed using the following primers and probes:

Forward GGTGATTGTTTATTGTTATGGTTTG (SEQ ID NO: 10)
Primer:

Reverse CCCCTCAACCTAAAACTACAAC (SEQ ID NO: 11)
Primer:

Forward GTTATGGTTGTGATTTTGTGTGGG (SEQ ID NO: 15)
Blocker:

-continued

Reverse AAACTACAACCACTCAAATCAACCCA (SEQ ID NO: 16)
Blocker:

Probe: AAAATTACGACGACGCCACCCGAAA (SEQ ID NO: 14)

[0155] The reactions were each run in triplicate on each DNA sample with the following assay conditions:

[0156] Reaction solution: (400 nM primers; 400 nM probe; 10 μ M both blockers; 3.5 mM magnesium chloride; 1 \times ABI Taqman buffer; 1 unit of ABI TaqGold polymerase; 200 μ M dNTPs; and 7 μ l of DNA, in a final reaction volume of 20 μ l);

[0157] 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)).

[0158] 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 ($p=0.00000326$). The results are shown in **FIG. 1**. A ROC curve of the assay was also determined. The AUC for the MSP-Methyl-Light-Assay is 0.91 (**FIG. 3**).

[0159] 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**.

EXAMPLE 3

[0160] The 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 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**).

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 15

<210> SEQ ID NO 1

<211> LENGTH: 29993

<212> TYPE: DNA

<213> ORGANISM: Homo Sapiens

<400> SEQUENCE: 1

```
ctggtttttc ccttaaggtt gcttttaaat gaaattaaaa caaaaattat gttttggatt    60
atgagctcct agagcgcttt gtttaatttg acttatacaa aatctaagtc caagacagtg    120
ccatgcacgt ttgaaaaata acaaaatagt aaatggtaac tgcagaggca tcccattcca    180
gaggagtggg ttaaggaat  tgctaataag tgaatcttat atttcactg atttaccta    240
taaagtctta tattttttat ctccccactc ctccccaaac acatcccctt tcagactgag    300
ttgaacatgt acaaatattc tgagagcctg actcatgtca tgcagggccc actggtggtt    360
ttgaatcact tcaagaatct ggcagcaggt ctctgacttc aatttcccc atgtgtcttt    420
ctaggccctt tgcgtttctc ccaactgttct actttccctc tcttcttcat tgtctgtact    480
gtattgcact actgggctaa tagatgctag ggaagggtaa tgagtaaccg gtacaatatt    540
tcaggaaaag gatttgcctt ttaattgtgc atttatgaat tgtttgcatt ttaccacaaa    600
taagatagta agtaaaacaa taaataatac atgaatttta aaaagacgcc tcaacttcag    660
agacttcaag tatattttag acaagaaaaa caataattaa ataagcacag aatttaaaaa    720
gtaatacctc aaagtaaatt ggacacaaca ttttgttccg gttatgaaaa ggggaggcct    780
gaaaaatctc ctgtgtgagc ctgagggaca tatggaatgt aattctttaa accaatgttt    840
ctcaaacctg atgtgagata attttaggtg gtacatgaga aataactttt atttctacaa    900
tataggttcc tttgcaagat catcttctat tcatggcggg ttttattagt tttccattta    960
tggtgactat aaaaaatgct ctttatcaga taaacttatt taataaaaaat agattattta 1020
aagatacata ttaagtcaat aatataggcg atatggctat aattataaaa gtgatattca 1080
atgtaaagtc tggaaataag gcctcattta agcaacaat acatttttgc ctttataacc 1140
ctaaaattct gtttcotatg ttaaacttat ttcaaagtgt tactgaaaaa tgacgtttta 1200
tctgacacct agtgggtgcc atgttaatat atgtatattt aagacaacta ttagaaggta 1260
gtaatatgaa gtcaaaatac taatatatta tcctggattt taaaaataat ttctgtacat 1320
gtattcttat ctagatgtac aaaaaactat gcaggagcaa agtgaccttg aaagcaaggt 1380
cagtgtactt gaaaaatctg actttagaag ggtacaaaga aacatgctaa gaaaaccac 1440
cctatattta tgttaacaat atgctaaagt atgaattttt atttatttca aatcatacat 1500
ctcccacctt tattttagtg tgaaactagc acataatgtt ttatgccaaa atgaaattca 1560
aaatgttgtt gccttttgag gacagtatta aattcatcat tgaacttaat ccttacatga 1620
tatttgctat agccacagaa cattcttaga cactatgatt cacaaaacat ttaaattgca 1680
tatttggcctt gtaatatatt acatatttaa aagaattact atttttctga ataccattaa 1740
tactaataat agcatcacag tgaactaagag tggacttaaa taaaggctgg cttgaaactc 1800
aggctctgca tttatttagc agcttctaaa aattctgagc ccttagtttt ctccacatgtg 1860
aaatggagaa aataaatctg caaaattaa taatctgttc cattcatatt tttctctcca 1920
```

-continued

gtgtatatta actggcattc ctcgtaggc cagaatgtgc totcaacccat gctccaaatc	1980
cgctttgtgc caacccccact gccagaaccc tttctacctt gagaaccaga aaaggaaca	2040
ttatgcctgg caatgcctac accctccaaa ataaatctgc aggaagaac acccagtaag	2100
tgatgagagc agcaacgact gcctttatca ttttaaattt acaacaccac cttttctaga	2160
gcctcttaag cattgtagat aattccccac tcattaataa ataaattgta accataagta	2220
ttcagggttg atactgcttt tgaattagac agtgctcata tcagttgcat aagaccaacc	2280
taaagtagag gatgaaatct tttttctgaa cttttttcag aacgtaactt agtgaatata	2340
ttaaaactaa actttctttg aatgggagta attttctacg attaatctgt aatctcttag	2400
accacaccta aggtaatgta gaggtgttg tatcataggc tttgtgatta gagaccactg	2460
gatttgcctt tggaaaagtc actcatgatt ctttgggctt tggtttcctc atcttttata	2520
ccggcttaat aatgaccacc gtagagttgt catggaagtt aatgaactc atgtaccata	2580
agtgaccaat acaatgattg acactcagtg atttatcaat aaattaaac atttattatc	2640
aatatgacag agaaggtgac gctaaaatag acaataggtt tttggaagag gtgattaaat	2700
ggatgcaaaa tttatggatt gtttattccg tctaccttg ctgtgtcccc tggttgtggc	2760
atacacacgt gtgggtataa aatcgtaaat cctatgtagt cgcgtagtgc atgcccagaa	2820
ggcttagaca cgaatgtca tttcagcaat gtgcctagag aagctctgac gccgccttg	2880
aagtaagtgc ttgctgcctg acctttgggc gtctgggacg gatgcctata cctgcacca	2940
gcagcactgg aaggggccca ggcccttcgc agcacagcct atccccagac cgcttagtcc	3000
ttcataacat atatctccc ggaaaagggg atttctctcc gtcagaaaaa gcgccccagt	3060
ctggtctggg ttggttttta tttcacgttg ttgcaagtag gcgaagtccc ttctgtctcc	3120
tcccttgggg taagtggaaa ggagtccggc agggggcccc cagtggcctg cacaggggaa	3180
ctgggtagcg agagagtcc aggcaattcc gggggctgcc ccacagaagc aggtgggat	3240
cgacagtggc tctccggccc agggaggaga gcgcggtcgc gggtcctcc cctcagcctg	3300
gaggctgcag ccgctcgagt cggcccgggt gggggcgggg tggggcggc gcggagggca	3360
cggagattac ggcggcgcca cccgggacat ccaggcccc gaggccctgg gcggtccca	3420
cgcgagatcg caaacctgca caataggcag tcaccgaggt tcaataaaaa acggagtggg	3480
tccccgcgc gcccgcccc ccccgctccc tggcgccctc ccccaggcc cccggcggcc	3540
tcacgagccc gcagtgcggc gtggcgagct cgcctccccc ccacctcct gcgcaagtgc	3600
gaggctgccc gcagcgcggc gcacgctccg gccgttcccc gcttccgcgc aaaacttcca	3660
tcctgtccac gtgaagttgt cgctgcctta gagaggggga aagagctgcg gaaaaagccg	3720
gggagtgcag actgcggcgg ctggggcgcg tctctcattt tctttcttc tcctttcccc	3780
cctgtgcgag tccggagttt tggctcctct cctttcctcc tccccctcg agccggcttc	3840
tccctccgcc ccgcttctcc cccgottgtg tacgctattt gttgtggggg gccggaagg	3900
gatgtcctgt tttcaccaga ggcacagcgc gaaggggaaa cttcgacact ggaaggaacg	3960
agaataaata cttaattacg gacgcaactga accgcgctg ggacagacac ttcgggaacc	4020
cgaggcggac cgggcgacga ggtgagtgc ccttctctcc aacccccgcc ccagggtcc	4080
cgggggagcc tgagttgaga gaacccccaa actttccggg aaagtgcgcg aggctccgcc	4140
ggggacgccc agcgtgggt actgaggacg cgcagctgga cggtgctgg gcgcctgct	4200

-continued

ccccggggg	cgcttgag	ccgggtgcc	cacgcctgag	ggccggggc	gctcggaccg	4260
cagcgggtgct	ctctgccc	gaagacgtcc	ccaagcccca	aggggtccctt	ccgagcctgc	4320
ctgtccccttc	cggggtcggc	gcggagcctg	cgcgtaacgg	agttcatcca	gcagtcaccg	4380
gcgcggettc	tacctgcacc	ccgcctccac	ctggcagagg	cgcgagcacc	ggggctctcc	4440
ccacatcttt	cttatgacgt	gtattacttt	ctgatgacct	cctagatggg	ccaggcgcga	4500
ggatgctgac	ccagagtctc	tccgagggtc	acaggcgcct	gggctttccc	ggtgccgggt	4560
gcgtgtgtac	tttaaaggct	cgcgttctaa	tctccaggca	ctgatcgggc	ttttcaactg	4620
cgcgatccc	actttaatag	tttttatgtg	gcgtggactg	aatgtctcct	gcagtttgcc	4680
agggctcggg	aaattagag	cgcttctgca	gagcagtcgc	gttcattggc	tcgagtagcg	4740
ggtgccatgg	aagcttata	acttctccaa	aggaaaggac	ctggctgggt	agagcaggtt	4800
tttctctcct	tccaagcctg	ctgggtctgg	ggaggcagtg	gaacttgaaa	tggtctcgat	4860
ttttagcgtg	gtgaagcgag	gtttggaagt	agacgtgtgt	gtgcttgttt	tattctgcgc	4920
cgcacagcaa	ccgaacttt	cgtttgtag	cacttgaaag	agttttctcc	ctttgtttgc	4980
gagattctga	acagctcgg	gcgattagg	aatttgccga	ccgagtcgg	tgccagagct	5040
ggggcgaaag	cagagagcgc	aatttaattt	ttgtcatctc	ttccctgctt	gggaggatag	5100
tgtttccctt	caccaccacc	cctctttctc	cttctatga	agacaacgga	tttgcgcctg	5160
gggtgagagt	gtgtgcggga	gagtgggtg	gagactgtcc	totctcaccg	cgctctcctg	5220
gcctctcccc	gccatcccga	gcgggcctag	agagtcattc	atgaatctta	acctgagggc	5280
aggggaggaa	ggtgcaggtc	cctctgccc	ttctgccaag	gtgcagaata	gcgcccgggc	5340
gtgtgttttg	gttccagagc	agttccacgt	ggagcaactt	cggtgtgtg	tgtgtgtgtg	5400
tgtgtgtgtg	ttgtgtactt	gatctgtgag	gaggtaacag	gactctgggt	ttcaaaccga	5460
gtgggccggt	ggccattagt	ttgctttcct	ggctgtcatt	acagacactt	ccaaaactctg	5520
atacctaaga	gaaccaacag	gttaggtttc	acattaaggg	ctcactacta	acagttttct	5580
ttctccctta	ccccttttgg	cttggcacc	tggtatcaac	gtaattgttg	gagcgaata	5640
cacctctg	aatatggcat	ttgttctc	ttctcatctg	tgccacttt	gtgaaccctc	5700
gggtgtttgt	cagtttcagt	gcgggtcct	gcggatcta	ggtgggagtc	ttaggagcgt	5760
ttaacaaccg	cggtctccc	atcagcagct	tctgaagttt	cacttacacg	taggtgactg	5820
acaggattga	aagttgacga	tggttttttt	gtttgttgcg	ttgtttttgt	ttttaaacc	5880
ttagggaagg	gattgtactt	gaattccttt	cccgggtacg	gtttggtttt	aagcagaatc	5940
agtgcccttt	ttttttttt	tgtcttttaa	aatattattg	gcaagcttaa	acctgaagaa	6000
ccaaaaacta	gagggggtg	ggagagaatc	cccccaaaa	atatttgata	cgtgatacgg	6060
agcgttttag	gagactgcat	tcaaagacat	ttgtgtat	ttaaaaata	cattatccca	6120
agaacaaaa	agcagtagta	acaagagaca	gattgttttg	tgtggagcaa	gactgccaga	6180
atctgatttt	tatggcaaca	atatcgaaag	cagacataac	tacaccaca	tttattgtta	6240
taaaccgtaa	aaatagtttg	ttccacctga	ttaaaagttg	taagtcttc	aaagttaa	6300
ccgtatttag	gaatgactgt	acaagaatgt	taaaacctt	gacagctaca	gctttgaaag	6360
caacataatt	agattgttg	aaaatgtctc	ctttctctga	ttattagtga	tatgtacttg	6420
tctttcataa	taagtaaagg	tctacaatca	atatggtttt	caaaagcctg	ttttccttg	6480

-continued

taaatttcct gaaaatagta tgtttcaata tttaaagacc attgtaacta ccggttggtt	6540
cacaattata actaaaataa gatttactca gttgttttct ttttattagt atttgttgaa	6600
tatttgaaat tgaattttctc atttccaaaa taaaacgtct cacatatatg tatgtactca	6660
caattaacac ttottttaag tagtaggctt aagttttaaa atttttaaaa tcttaagatt	6720
tgtataaaaa ggagtgattg tataaaaagg aataataatt attaattggt aagaaaaata	6780
gatgctgaat ttcagagggt ttctagaagc tggagaaaa aattgcatag aagtttcttt	6840
tttgacgggt gactggcaga tgagttactt gtcattttgc tataatttat ttttttcttt	6900
taataaggct ctttgggta aaaaaaaaa aggaaaaaca agtactccat ttctaagagt	6960
ttccttttta tttatgttgt actttgggtt gtaaataaat tttacatagc cacgggaaat	7020
agagtatcca atttcatggt ttcagtctct gttgcaagct ctaaatgact gatgtcgtgt	7080
ctaaaaatat atttatatgg cttgatattt cattttaaca ttttatgcac tggaaattaga	7140
tttttagga ttttagaaa gatagaatga cagaaagacg atgactattg ttataaaaag	7200
ttagaatggt aggaagtcca cattctaact ttcaaataat caaacttatt ttgccccaca	7260
aaaactcaca gtttttatac agatttcaag aagaggggtg ataacttttt tccactactta	7320
gttaagtgtt aagtaattta agaatgcaag tatgtattat gctttctcat tatcattttt	7380
ttctgttagg taaatgcaag aacactggac attctgtaaa aacaggcccc tctttataag	7440
gagtttattc accatagttc tgtattgcaag ctgctgatca tatttcatgg agctgttaa	7500
gcacttaaaa cctaaaatta ggtactgtct ggttgtaaat atttcagatc acttatttaa	7560
gaaataaata gaaaagtgtg ctatcaaaaag taggagacgt tttgaatcct ctcatgaag	7620
agctgaacaa acctctatca aaacaccttc ttcttttttc agtgagaata taactctgac	7680
agtttctttt tctaaatgga tattattctt acatgtacta aatgctaaac tcttataacg	7740
tgccctccctg tagagtatct taaactaatt atattcagaa atacagttgg gggattttat	7800
tacaatgggt actagtgtaa gaaatcaac accaggggaa tgggggtgca ttgcagtgt	7860
cctgcttctc tcatgaattt ttctgctaa actaagaaat gacatgctgt ttcaggcatt	7920
tgccctggag atgggtgaga tgcaatatgt gtaatgctgc atcttatagt tagatgtgtt	7980
ttaatgaagg ggacactgca tagtcaataa atcatttttg gagccaaact tggcgtcatt	8040
tagcttgaac ttaagtggaa gaaaatgaac aagagttaca cattcaaaag aagtacaagc	8100
aactttgatt gcttttaaga ggtttgaaga ctttgtaaac attactgtca ctcaatattg	8160
cttggtggagc tgtacattaa tatatgcttt ggtgatatgt cattttacgc ttgaaatttg	8220
tcattcttag tgtttctcca tccatctttt tattagtaaa gagatactga aaatgaacac	8280
tactattctt actcccctaa cccctttcac acctccagaa aaagagatga aactgattaa	8340
tttaaatag aaaccatttt gtgttatcaa aaccacattt atatagtgat ttgagacagt	8400
ttcagagagt gcacctctga gtctcactgt aacctttttt gtcattgaaa ggtgctaatt	8460
gatcttaggg tactgacaca atagtatagt ttgatatttg aaaccttca gagttggctc	8520
ggcccttttc ttacctgag atttcagtgc atggatgatg aagaaagaca ccattctaaa	8580
ataccagaaa ttctcatttt tttccaatat gaaatgtttt aatacagtat gttcatattt	8640
ttaaagcttt tatttacata cagtaagtaa atttatttta acgtactttt ggacagtagg	8700
agaaagacct atatgttcta tcgtgttaga attttttagt tttttttttt ctgcacaggt	8760

-continued

agtttattta ggttataatt tttaggcaaa gtctgattcc tattatcaca tgaatatttt	8820
caaagtgaaa ttgctgtaaa ccaatgtgga atagcttttg taccaccaag gcatatatta	8880
atgtagatgt caaatatgag agcatatttt cttgagtata tttatatcct aaagtgtatt	8940
tttaaataaa agtggctact gtagtcttta gataattaca atttggctgt cattattact	9000
ataatattaa tcactatcac catcataaca tcatagctag catttactca ggaatttctg	9060
gcaaaacact gttttaagca tttatatgga ttagctaatt ttaatcttct taactatgcc	9120
gtaaattagg tacttttggt attcctattt tacagataag gaagctgaga cccagctcatg	9180
cagcacagtg gagccagat cctaactccc cagtatgagg ccagcacctc tacccttaag	9240
cgtgtgctgg gctcttgctt ccatcagtca tataccacca tttttaggty gtacctgaac	9300
atthttgttt aataaactt atattttatg taaggataac tagatattag aaaaatttgt	9360
taaattttgt attaaacctg taacttcatg ggcaatattg tttgagacia gaccaaacia	9420
agtattgaag tcaaagaaaa aaaattaagt atctgaagaa acgtattaag taacagtga	9480
caagaatatg gcctaaataa ccacagctcat gaaggctgga cagcaaatga ctgaatttgg	9540
agaaatgctg tattttgtaa tgtttcccaa ttatcaagaa cttatgacca gatcttttaa	9600
atatttaact aacatgtgga atcttctttt gctcttcaag ccttatccaa attggttaaa	9660
tgttatcaac tttgtatttt ctttgttttt gttttttgtt gttttaatgc tggttttgaa	9720
tctcaaatct gcacatttat gttgaaccaa ctaataaggc ttgaagagtt aaagagtga	9780
tgatggactt ctggaggcag gtttaaatga taatggagct gcccatattt tggaaataca	9840
ttcaatttat ctggttatcg catgtgtaag gttttctgta ggtaaatact tttgctgtta	9900
ataacctgtt ttacaatttt atagattttt tccactgaag cagtggtttt cattttttta	9960
tttacttata tactcagcca cggatccttt cattttaatg gaaactgatg atatgccag	10020
tagagaaatg tgccactgct ctagggtgaag caggatgtag gtagcctgag aatgactcac	10080
caacagtacc ctctcagcgt ggccgcccgc tctggagttc tctctaagtg tcttcaatgt	10140
atggaaactg ctgcaaaaaa attcaagtct tctgacaaaa ggggttaatt cagagtacct	10200
gccctaacat gttcatgtgg catcactaaa aaacagactg tcagatacgg taaaatatct	10260
cccagtgtga catatcagca gaaaggtgtg tctactctcc cttctataac ttgagtatcc	10320
gtattaacca gtcttcaaat tctgatttagc actgagaaaa ttaaaactga tcaaaatgtt	10380
ccctgtgtgt agttacaggt ctgaatgagg cacaaggac ttgtacctgc aaaggttgac	10440
tttattaatt agaacatctt tctctcttta aagactgtaa gaagaaacac cagcagtggc	10500
ctaacttgac atgactttag attttcacgt aaattattgc tactatttct gttatccttt	10560
ccccctttct ttttaaatg aaagggacat ttcttgtgaa agactacaat taaatcataa	10620
aaatttacat tcatgtgcca ttaagtttaa ttctactcac aaaagcaaca gtacagagtt	10680
tgaatttcta tcccaatca agtaggtgta ccacataccg ggagggtca ttatgcacia	10740
ggcatatat acaattcaca gacctctgca tatacccaac ggagtgatct attcattaca	10800
tttcacctct gactttgaac tccctaagt taaaagattt gaaaagaacc gaatgttctg	10860
attaagagat tgaatatttc taacttaatg ttttcagtat gttgaaagtg atgatgactt	10920
gggggaatca gcagatctct acattacctt attcttttct cttacatttg aatgcaaatg	10980
tataattcatg tgcggttatg actcaagtca ttcttgctaa atttaatgac gttgtaggtg	11040

-continued

aatcacattc agatttcctt ttgcaggttt tccagtaatc taaaacaatg cttctagtag	11100
gtaacttaag catgcaaacc tcaataaacc tgcaagaac ggcaattcta ctgttttatt	11160
ttgttttttt tttgtttgtt ttgtttttgc attaacttta gttgataaga tgatgggtact	11220
gttatttttc ttagttgact catgaagaat tttatatttag gtctagtttt ttccttaat	11280
tggtgacttt agtttttaaa ggtttcgttc atgaaaatgg ttagcaaagt tgtgggtact	11340
tggtaaatgc ttgttaaagc ctttttcctt atcagtggtg ctgaagactt gcaaaattag	11400
agtgggatgg atagatttct tttctactct gcatggcttt gaagactttg gagcttttat	11460
tgtattctta tatttttaca taccattccc aggaatattt agagagagaa tcattgtaac	11520
caaggtcaca ggtctaacc tcagggtattc aaattagctt taggtggaca attgtcctat	11580
acacacttgt tcgtattaat gttgtcataa caaatatagt tataatatct tgatgcctct	11640
cctggggggc cattctggac tgtgttgaag ctgtctctaa cactctctca cgtccctaag	11700
atattcaacc acatttgctc atgattttta tgaggtcctt cctggactta aaattcctta	11760
aaaaatttgc ctgctgccta tagggcaaat tccaaaatc tctgcaggga agataaccct	11820
tcatactgag aaccttgctc aggttgctct tctcatctct cactccccac tcacaacaac	11880
tctcttttcc agctgtatgg aaaactgcag ttctcaaata caccttgac atttccactt	11940
catttctttg tgcgtggaat gctttttctc tggaaaacat cgccctctc cccctgcaaa	12000
tatcttttct actgcttttt caatacttaa ggtgcaaaat gttagctcct ctgataagtt	12060
cctctgacct ccagagtctg agttgaacgt ttctttcttg ttactcccaa agcatccagc	12120
ttctacattt aacatagccc ttgccacatt gaaatgtaat agattacatt agtctaaatc	12180
ctgctttaag ctgagaattt actggaagga ttagctgtat ttgtcatttc tgtatcgcta	12240
atgcacagtg cactgctgga catatagtca gtaggtagtt gttatattcc cattgaatga	12300
atgagtagag gtactgggaa tgagagcaga gattgtgaca ctgaacatcc cctccttgag	12360
actgggaaca gtgagaaggg cagtgactat tacaagggtg gactacotca ataacaacct	12420
caggaggttt tggataaac ccagtagaat tgcaagctgt aatccaggct taggttataat	12480
gtatgtatac gtgaatgtat aaatgtgttt atttcttacc actgggaagt cagtgatctg	12540
gtggataagc caaggatcct gaaaatctct ggagatacgg taatttcata gtacttgaaa	12600
tctgagagac tcagttctta tagcacagtg agtaagtaga aagaatatgg gggctggcg	12660
cgggtgctca cacttgtaat cccagcactt tgggagccg aggcggatgg atcacctgaa	12720
gtcaggagtt cgagaccagc ctgacccaaa tggcaaaacc cgtctctac taaaaataca	12780
aaaattagct gggcgtggtg tggggcctt gtaatcccag ctactcggga ggctgaggca	12840
ggagaatcac ttgaaccagc aaggcagagg ttgcagtgag ctgagattgc accattgcac	12900
tccagcctgg acaacagagt gagactccat ctcaaaaata aataaataaa taaaaaggaa	12960
gaagaatatg gagccaggca cttgggttca gacctcagct ctaacattga tgagtttcta	13020
acatcaaaca cattactgaa acagattttg cttcagtttc tcatgtgtaa cagaacatgc	13080
ctcatggggt tctgatgaga attcaatgag taaatatatg taaatatatt tagagtagtg	13140
cttgatatga gtatcagttt ttattactat gattgggtat ctttcattag attacccca	13200
aatgttcagc agatatttag agctgaaggc cttttttggg gggctatgga ttggtgacca	13260
taaatggagt gtccagtagt ccaagtataa attagatact gtgggtcatc ttggcaaatc	13320

-continued

tgaggttgta tttggagcct taagagaaga cactataggt gotcaacagt tggcattggc	13380
catttacggt gtgcagtaat atttctttta gtaaatattht cgcatacag taatctagat	13440
tgttttagag ctttaactctg tgtggttcct taactacttht aagcaattat aagcacgthg	13500
tataatggta gtactttata gccaatattht aagtttcctc agttgcttht ttttcagccg	13560
taaggtacca attatgagaa tttgtataat tttcaatagt cagtgggcct gatgaactca	13620
atgtaaactt tttaaaaggt agctttacat attgctatga ttcttacctt tagagagaca	13680
gtaggaggta attagtagct ctacttacag aatttatcac ttggattcag atataatggc	13740
tatgtggcag gatggggctg aaggaaaaag gaaagttaaa atattaattg tcaaggctc	13800
atthttacag cttgtctgga actgctacca ggtgtatctg tatagthttht aaaatgataa	13860
cgthgtacct gaattcttca gtatathttht ataagagtht ttcaagctct ggtthattgac	13920
atthttggct ggataattht ttcaggggat actgtcctgt gctggtagg atgcttagtg	13980
gcatcgctgg atgccaggga ccctccctcc cctggtaatg atgggtgctc cagacattgt	14040
caggtagctc ctggggagcaa aattgcctct tcttgagaac ccataaccta agcagatctt	14100
aatatcattt gatatagtha aaaaactccc agggctctgc ataagaggat tgttctattht	14160
ccagtcaagt gtggagaatc ctaatacctt cctcccaagt ttaaacatga agtcaaaaa	14220
ttctcattag tctgtgttga tagattaact ttgcacagag gggatttaca gatagcttc	14280
acacagattg accatctcca gcatathttht cthttctgga aataatataat gagtgggagt	14340
aacagaatac ctgagagaga gtatgtagga agtaaatatt tathttgaatc atctatgtgt	14400
ctthttctcc ggtthattgt cagtagatag tttgatttht tthttaaata atgaccattg	14460
ctthttccatt tctgtgggta aatattcagt aataagaaac tthttathttht tttagctctgt	14520
agtgthtagaa aaggtaaaat tactgataat cacaactgct gaagattaaa atacttagtg	14580
agthtaatta tttgtthtga atgagaaatt tthcaaaagaa ttatgtgtag cthttcagthg	14640
taaccacaag thcaagatct tgagttaata acatttgtac agctagaaga aagtaaaaat	14700
aaatatatha catgtaaaaa ctctctthttg tgaattctca gatccaaact tthtaagthct	14760
cattctatgg gthttgatgt gaccaataca tthttatctat gaaatgatat ataattgatg	14820
taagtaaaat tgactgttgt ttgggcacca ccottacata atthtaagcat atatgaaaca	14880
cattthcaaaa atcacttgaa caaaatggga acaataatgt thcacataaag tgagagagga	14940
taactthttcc catatgagth ttgggthtga tthcaaaatt atthctgaaat atgaaacatt	15000
aatattataa atatatgatt agactaatht tcttaagagt thcaaacagt agataathtt	15060
tatcttatta gaaattgcaa agatacttga aattccacac thtatgcattg thcaatttag	15120
tggtgtact thaaattgtat taggtagata atcatggaag cagthcagact aacctggaga	15180
thttcacatca thttactaatt gatgcctaaag taggcagggt agaggtactt gaaaacacac	15240
acacacacac acacacacac acacacacac acacacacac actactctc thctctcac	15300
thctthtgat thgtgcttht atthttgaaat gtgtgcttag gtggcaaagg gctgcaaata	15360
cagcactctg thtttgacca aaatathttth ctgthttccta ggtggagaag gacctgcat	15420
agthctctgt tagcagagag ataaagtggg thctgggcaa cccaatgaca ggtggaggcc	15480
gthctcgggg agthttgctg agaathttgag acaggggtht aaatcacaga tagtggagga	15540
gagthctgct thattcatgg acaggaagtc caaagaaggc acagcgaacc acagthccat	15600

-continued

```

ccaggcactt tatgccagat attgaacggg aagcctagta gccctttagt tactggagaa 15660
cggccaaaaa cccagagggt ctatttttcc aaatgagcac ttcctggggt cataaatcaa 15720
ggaacccccc agagtattca agtcacatat ctccctttctc aggttgattc tttgtatagt 15780
gggtcctaata gtggaagac ttacttatga acattttatt tcttataatg cgtcagtc 15840
cttccttttt tctccctggc atattctcac cacccttacc ctccatggcc ccaactccct 15900
taacctactt gcctagtgtc ttttgctgc ctacctggac aaaataataa taataataat 15960
aatctactca tctttcaagg cttagctcaa atgtcacctc ctctgtgaag ccttccttaa 16020
ggctcaggca gaattagttc ttttctctgt gtaataccat agcacttttt ttggactgtt 16080
aatataacag cagattgtgt tggggctcat tgtatgtgtc tgtcttctag agagactatg 16140
aactcacaga ggggagagat catatctttt tccctccgta acctcaacac ctggagatgt 16200
taatgaaat taagtactgc cactattcca tgcatactt taaagggcat taggtccaa 16260
accagacatc tgcaaaggat gaactcctgt aactttttgg aaaactgggt ctcttttgct 16320
tgtggctcatg aaagctgggt tattgggtat tttctactcc aacttattta aagcctcatt 16380
tgtaccaaag tattactgat tttctagtgg aaaaacaggt atgtccggcg tacattggta 16440
ttttcaataa actgacattt aaaagacaga tttctcgttg aggtttatgg tcttcataat 16500
tctagttttc tccagaaaca gaccttgatc tctttggttg ccttctgcta agatcgattt 16560
catgttattt tgaagaattt ttaaccctca gaattataga tttcattga gagaaagcat 16620
gttctgtatt tgtttgtgtc agtgtttcaa aatgtgagga ccaaaaatcc cagatgtatc 16680
tattaaaaat atacgtggcc ttttgctgag gaaattaca tctttttttt caggcacaat 16740
tgaggaagta aaaaaatttt gtttagtttt tagtggtttg tttttaggca ttttcatac 16800
aaagatgagc aagacccttc aaaaaccacc aatttgctta tttagggggg aaagtcttcc 16860
tatgtccaga aataacatta aattaagatt attgtttcca atagtttcaa aaattgtggt 16920
tttattttct tgatgatgac attttttagc atagtgcagt acatgattgc actacatcat 16980
gacaaagcat atttgctctg tgtttcagtg ggtcacttta tttatcagcg tcagatgatc 17040
agatcagaag agagatctag cctaaccctt ccattttgca gatgagatgt tgtattatgt 17100
gaccttaaga ccttaagtca tttactcaga acaaacaga aaaagagaac ttcattgctg 17160
ggacacttca tgcctccaac tctacatctt accactttat ttttatttgt ttgtttctga 17220
tgaggagtgt agaagctgtg aatcagacta cctttgtgtg tatccaggct ctgccactca 17280
tcagcgatgt gactttggtc atgtaatttt ctgagtggct gtaagccata gttccagcat 17340
gtgtacatgg ggaataataa cagtactcgc ctcatggagt tgtaaggaat aattgaattc 17400
atgcagaaa gacatttggc tcagtgctgt gccactata aagctcagg ttatagtttg 17460
aattgatatc ttggaataaa atgctgagtg tatttgatat tcagagaagg aaattcatgc 17520
tgctaaagca agtttttttt tttaatctgt tacttaaaaa atacaataa gtcatacttt 17580
gtataacata taaaacaga gatgatagag gttacttttg agaggaatga ccagaggag 17640
caagagagag gctttggaat gctgattatg tcttttatct tggtttggtt gctggtcatg 17700
tgaaagtgtt tattcaataa aaattcatca gtgatcaact tttgatttat gtattttatt 17760
ttactgtgca tatgttatac tttcttaaaa agttttaaaa atacggatat atttaagta 17820
aaaagtggat ttactcattc ttattctctg caaatcagga taagttgtta tacttagatg 17880

```

-continued

tgcatctttt caggcatttt tttttctata tatgtgtaaa tgtaataact atgtgaggat	17940
gtaatttatt ttcattttac atttttaact ccaataggat cattgatttc cttaaaactt	18000
gcttttgatc tgctgaatct tgaatatttt tccacatcag tacatatagt tttgggtgct	18060
tctatttaat gtatgcagat atttcataat ttcactgttc ccctottggt gtatgtttac	18120
attgcttcca tatttttact attactaagt atgttgcaact gatgattttt atgtctcttc	18180
gtacatctgc caaaaatc tgtagataaa actctagatg caatattaat agataaatat	18240
gtatgtgcat tttccataga aaatacaaaa atgcttttta aaaagattta tgaatcttgc	18300
atctttaagg taatattctc atctttctca caaaaaccag gtgcaaaaat ttcaacttaa	18360
ttgaaccttg taattatttg tgctctggtt aagaaaaatg gttatagtag taaattcctt	18420
gaaactcata tagtaatcgg aatactccac tatgattagg ataaataatt ctattaacga	18480
tatatattta atatacatc ataatttata taattagttt ctcaactaag aatatttagt	18540
attcattttt atgaatgcat attatattta cttatggttc gtaagcccaa atgttaggaa	18600
gaaataaatc agtaatatta tattcactac attttttttt acttagtacc atcaaatatt	18660
taccagaatt attaggagaa atggttctaa ctctcatgtg attggtgaga acttatgagg	18720
aatgatataat tgttattcta gtaagagcct actctgtgcc aggcacagtt ctaagcacct	18780
tacatatatt acgatagctt tttaacatag agatgaaagg tatagacatt gtataagtaa	18840
taaatgacat tactccatta taaatgtgtt gataaatatt attatacatg tcacaatc	18900
agttgtgtca tatcaatgac taatttataa atctgttact tattataatc tgtttaatgt	18960
ggaaaataaa gttaatgcat caaaatccag aaaattgcta ctcaaatgct ggctaaagta	19020
atctgataga tacagtttta aatgaagta tattaatcag ttgattttgt tgttcgattt	19080
ttatatgtag aaaattctgt ctccagaata ttcagttctt gtatgttttg gttaaagatt	19140
tgattaatct tcaatggtta tctagtgcct ttaaaaagta gtccaactca gaatagctaa	19200
tgtattcctc ttgcattgaa taaatatggc tatcaatttt tgtggggttt tttttcctgc	19260
agaatttcct gttattctaa gggattcaat aggacttaca catataaac tgaataattat	19320
attactatgg ggaagctgtt ttgtgccttt cagaaggaac tctcttgcta acttaagtac	19380
tgtatgtgtg tagaatagct aatattttcc tcccaagaaa tttcatgtat gcagttaaaa	19440
cactcttaaa ttgattaagg atttgttata ttttccaga agcactcaag agtactttta	19500
tggttataag agtagagtgc attagaatgc cagaaaccaa tgaatacgtc agaggccttc	19560
aaaactgtgc ccaacattcc tgactttacc atataaacat gtggggtagc ctggaaaaac	19620
aaaacaactc tccttcccc aacgtagggtg aggcctgtga tgataaacact tcaactggct	19680
tgaatattaa tttgatactt ttctgttaga agtaattttt attacttagc aatgaaatg	19740
gagaggtaat aaatatgact gatatttatt tcttgagtaa agtattttt tacatcagat	19800
tcgaccagtt acotttatcc tttttgaaa atgtattata caaactacca cagtattttg	19860
cctattaatc agggaagcta aaaacgttt tataaacggt gaatcaaac tctcactgct	19920
gtgaggaat taaatttcaa gcaatatact ccttttatag ggtgagaaac tgaagtatgg	19980
agaaattaca aggttttccc aaagctatac actgagtcac ggcagacata gaaatctaca	20040
attctgtttg cctgattcaa tctgaaactt catgaaactg aatttatcca tttcctgaca	20100
tgtgctttag aagtatattt ccttatttgc ttatttatgc tctcattcat tcattcatgc	20160

-continued

atgcatttat gcoactattga ctatttatgc actcattcac tcattcatgc attttaaaag	20220
aatcgggtgtt tttgtatgat acagtcttga tttagggact atgaatattc tgaatttata	20280
tatcttagat atttcttcaa agaaactcat tgcatttttt tcacgatcat gaaaaagaaa	20340
ttaagacatt aagaaaacca agctggattg tccaagtgtg ggccatagca gtgtctatgg	20400
gcagcacctg aagctaattc aggagtgaat ctagtttgtt caagtgcctg aacaagaact	20460
tggctgaagc cctggaggag atgcctgggg catagagtgg ggacgaggaa aaacctagc	20520
actcttagag agtactgtaa ggactagtat ccacatctct accacctgat gtctttacat	20580
ttggttcagt caagagagag ggggcaaatt attctgcttc tcattctgcc atcttctga	20640
tttcagttct cagcgtttaa attctttttt attgcgatct ttttcttact agtttaacca	20700
gtattttccc gcgctttctt caccaaaattc gtttctttcc atcagcttgc ggagttatat	20760
tttttgatcg aactgctggg gttacaagta atttgaaatg aggaaagtgt ctagcaactt	20820
ccctcagtcc agtaccctgt aacagagact tcagagtacg tagtgatgac acctgccctc	20880
ctccccttct atagagagtt tagtgggcaa cttttttggc ttgttctctt gttcttttct	20940
tgcagcccca agacttagga ctacagacttg gtttctgctt gtatggctct ccatctcttc	21000
ccttgaaga gttctttccc gggggctttt taggtccaag agcatttggg ccctgggag	21060
atgaacttgt agattctaca aagaacatcc aagtaccagg caaaccacat cttcctgtag	21120
cttaaagctt agtctttttt ccccttcaa ttttactgaa gaataattta gtctccaaca	21180
aacaagcaaa cagaacccta gggcttttct caaaggtgca ttcttgctgc tttgaagatt	21240
atcatcatta atttctattt cagctaattg gttttccttt acatttttat ttgggtataa	21300
atgtgatact ttggatatac ttaatgttga tctatattaa tacgtcaatt tataattgag	21360
tgatttaaaa ataattgatc actatgacct agtttatatc tgccttccac ttaatgcatt	21420
tcctaagcaa ttatgctttt gtggctgact gaattccctt ataattccct gttgggaaac	21480
aaagattttt ttttaaattg aagaagtggg cttgttaacc aggggcaccg tacttaactt	21540
gagagaaaaa gattaggaat gacaggaatg acagcatact cctaatatct aacagcctgc	21600
catacatgtg aaagcaggac acatagcagg gaagcataac gttgagttga ctcaaaacac	21660
caaatgtatt tgatacatta ttttctgaat tcttacaact tcagcccttt taaaggtctt	21720
tggattttta gaagctataa gggcagtttt tggtaaaggt ggcctactt ctaattcagt	21780
tacagagtca gtggaatgat ttttatattc aaatgctaca ggggaattta ttgagaggaa	21840
aaaggttact caaactttcc tggatgcctc aattaaaaat ctggatcaat aaaatttcct	21900
caagtgtccc aaatttaagg aaagcaaaaca gcatcttatt taattacatt cttaattagt	21960
attcactata gacacaattc aaatcagtaa cttggcatta gatgaatcag atttattgcc	22020
atattatata actctcatgt ttactttgtt gtgtttggga cttctccctc cccataataa	22080
agaatagac tcacaggtgt cacacctga ttctctgagaa attaatgtta tagaaaagtt	22140
gtttgaaag atataattgt gattttgtta ttgtttttgt tgtcattatt ttcttttatg	22200
gggatgggca cacacttctt gagatttacc agatttacct tgtagttcat aaattatctc	22260
tttagatgta tctaatttgc tgtttagtgt gtttcttcta aaaaaattca gtgaatatat	22320
ttttaaaaac atttttatgg ttctcttcca ggtttgtgtg tatatttttt tcgtgtgtgg	22380
cttttcaatt ttatttctgt ttcttttgtg tctgttttaa gtatacctat ctagtttttt	22440

-continued

ccacattgct ctgtagtttc acattcttag attcctaatt tgttgttggtg gctactctct	22500
actgtgtttt gtttctttat gtagcttata attttttttt tctttttgag atggagttgt	22560
gctctgttgc ccaggctgga gtgcagtggc acgatcttgg cttactgcaa tctctgcctc	22620
ctgggttcag gcgattctcc tgcctcagcc tcccgaatag ctgggatcac aggtgcccac	22680
caccatgcc agctaacttt tgtattttta gtagaaaagg ggtttcgcca tgttagccag	22740
gctggcctca aactcctgac ctcaagtgat ctgcctgcct gagcctccca atgtgtctgg	22800
attacaggca tgagccaccg cacctcgcct gtagtttata acttttgggtg gatttctctc	22860
ttagaaccct ttgtactctg tggagcaatt ttgcattccc ctggagtcta gggatttcaa	22920
acctctaggt cagtctttat tttattttat ttttttgtct tgaggtttcc ataccacaga	22980
tgcaagtaac atttagaatt tacatcctgc aaggctgaag ttttgattta tctcaactca	23040
gaaattatth tcttttattc tcttcccaaa ggctccatca taggacaggc ttctttgctg	23100
aatctgatgg tgggcccaggt ttttctatcc ccctccaat aaatgtgaca aactttttag	23160
actcctcct ctgtggattg cctcttgctt ttgcttctctg cccatcgcca taaaagaaga	23220
taatgacaac aaaacgataa caacacaaat gtaaacgtac ctttctctct taacttttcc	23280
atatcatcaa tttctcatga attaaggtgt gatacctgaa agtcacattt gccttatcag	23340
ttatattcat atagttcctt tgtcttattg atgtatctct catattatcc ctcatatcca	23400
cttctccagg cctttttcca ccccatgtca gtctagattc tottggttc tcaactggag	23460
cgtcacagca gcctttaagt tgccacttaa ctccacttct tgttcatctt gtttgcaga	23520
aatctgaagt agatgtgtg ttgcaatgat atcacatgtg ccattctttg ttttgggggt	23580
tttaaggctt ttaaggttat caaaatattg taataaccaa ttgtottaat tgtattttgt	23640
acttattctc ctgcttttgc caacttaaaa tgtcttaaca tctgttaat actgaatgct	23700
gatgctttcc tattcccttt tctttcacat gtccgtctat aattgtttat tgattatctc	23760
agtctctcaa cactaaaggc accagtcaca actttgatgt tatacagaat gacaactaaa	23820
aggtgacaaa tgctaattgc ttaatcgagg agttgttcat ttttacaagc ttacctgagt	23880
tcagagaatt ggtaaaagc tggtgacttt tggctactat ctaaatataa ttggttaatg	23940
tagctcctgg agatcagtgg ttagtataat actcttagtt gagcagggtt aaaaataatt	24000
acattttggtt aagttgcaga tactactcta ctgtacagca aagaatgtta aaatgtttta	24060
attcagtata tgaagagcaa aatataaaaa caggttttct tccactctca ctagcaagca	24120
aagtggctctg tatttttgaa atatattttt tgtttaatca taacaaaata taatacaaat	24180
gttattaaaa gccctattht atggatgaag acagtgatat ttagagatat taattagctt	24240
ttccaagatc atacatttat ccagtagtaa aagcaagttt tcacaagctg atctatthta	24300
ttccagctt tatttggaa cactaaactt acattccctt tagcaaaact atcgtttaca	24360
tttctaggg ttoaaagatt caatgctaaa ctgggagat ctattcttaa attctctcaa	24420
agtgcaacac attttctca ggatgatcat aaagtcacat gatgtgtgtt aagcaatata	24480
tatagtcata tactcttgat caaaaaagat ttccaggtta cacagaagag agatttatgg	24540
ctgagaaaaa ggcagattgc ttctaagaaa ggcttcaaat atattctctt ctgacaggct	24600
acaggagagc cagctagatg aaccattaca gtgtgtacaa aatgctcagc agaataatth	24660
ctcagtactg gttagggaag gcaataggag taacacaaga gaaggaaagt tggctctgag	24720

-continued

agtctgcctt agtatccgta ggggtgatga ggttgcttgg gotaagaatt ctttcacatg	24780
tctccctggg acaattcact cagaattact tctgtgtttt ctaggctatc taggcaacca	24840
tcaccaatta gcaaacacag acaaaataga acaccttgtt tggacctgct tgtcaaaaat	24900
caaatgatca taatttataa ttccaggaag acttgccatc aaatgaactt atgacaaatg	24960
acccaaatga ttctcacaaa gtagagcttg ttgtaacttt tagcatgaga agtatgaacc	25020
ccattatcat tttttaaaca ttttgagtta aatggaccgt tttgcaaagt ccatgttcta	25080
tctttataga cagctcaaaa atagctcagc atctagtttg taggttggtt ccatthaagg	25140
agcttgaagt caggcaaggt gctcattact attcatccag atctgtttct tggacacctc	25200
agctgttttt cttttctacg gccagggagg agagtcatgt gcatagtcca agaattgtca	25260
tctccatgtc tctgaacagt tggccataga attggagcta ggtgttcttt gtgccattgg	25320
ctttcctctc tggcatgctt cccgggtggt tggagccagc cagtgcctga tccagcttca	25380
cagaccaggt aatgaaatth ggagggcaga gtaggtaact aggtacacaa gaattgtaa	25440
gcaatgcca tctggacagt tgatcacaac tataacttct caataatgta acaggaatat	25500
aaactgttag acttctata ccattgcaat ataaagaaga taattgaagg cctggaagat	25560
agctcatgaa atgcaaggaa gaattggaca actaagcaac aggaagagca gagaaacagt	25620
gagactttgg gagtcaaac atctgggtgc ttaaattctg ccaggacttt cttgccatct	25680
tcatttttgc tcctctttct cggtcagttc tcttctgtct ttccctacag accagctttt	25740
ttccacatgg atatgaaacg tttccaagtt ttatataccta ggggttttgc cactgaagga	25800
ggactagcct gatattttcc cagactcata tcctaaggaa gatgaatcat tagatccaag	25860
taatttcgtg caactggaga agacttgca accctaccta atgggttaag ctggatatgc	25920
atctgtggga gggagatatg ttttgtgggt ataatacaca tgcaagcaaa ccttaatat	25980
gtacaagata cgcttctgaa aatttacttg cagtttgaaa atctggaagc caaacagtag	26040
accaactggt gtgagtgtta gaatttagtg taatggcact ttctgtgacc caggcaactc	26100
tgcataggca gaaagattct tgaagcccc acccaggttg gaaaggtct gtgtcagagg	26160
tcatgagtgg gactcttccg gcacctgttt ccccatccca aggtctgcac atggactgga	26220
acaaaagttt tggcagctgc aatttttaaa tcttttccca aggtatcta gggatttaaa	26280
gcagtaggga atgtactccc ccactttccc ttgttttctt tttcactgca gcaagaggca	26340
agtatagccc tcagccctgc ttgtaccctc acaggagact catgtacagg tcccgggctc	26400
tggctgcagg acagtgtgag gctgctgaca gggcgagaga atggcgagaa caatggaaca	26460
ggactgaggg agaaggaaga ggaatcgtcc ccagaactac tacttgggac ggtggtccca	26520
gggagcaaaa aacaccctcg gagcccttc aaggggcctt ttggcagagc ggtactaccg	26580
ctagacactg ccatggctgc aggtaggact gattaaagaa tttctgattt cttctgagga	26640
cagggtaaag ctctgaattg tttttagaga ctgcattcag atctatattt cggaaacctc	26700
actctggttg cactgtgaac ggtaacgtcc agggagcagg actggaggga aggagaccat	26760
tacgtggctg ttgctcctgc ccaagcaggg aacaatgatg gcggaacgcc tgtactggca	26820
gagaggagag agaggaggga tttggcgaac ttttgagggt attggtagga cttggttaata	26880
tattgaatgt ggagagcaaa tgggcagggg gtcagtgaca aagcctaggt gtttggcttt	26940
gatgactggc tatgcctgtg ccggttgtgt acagaatcca gaggtttagc gtttccatgt	27000

-continued

```

gcagcgtcgg tcctcattcc gcggcagcct ggtgtggagt tgagatgtgc cccctccact 27060
ccatgttcac gtttctttag ccagagcaaa gcatgcctgg acaacctagg tttctttctt 27120
tctttctttc ttttctctag agagagaggg tctccgctct gtcacctaga agaagtgcag 27180
tgtcgtgatc atagctcact gcagccttga actcctggcc tcaagtgatc agaggattag 27240
gtgtacttct ctacaaactt ctcaaacctc cttcagcctc tcattcctta agggacctta 27300
cagtctctcc caoctggacg gcgccaatgg gcctcattcc atttctgaga attgtcttaa 27360
aactttatat agaaaacttc tattatgtat ttaaatgaga agaaatgttg aatttgtaat 27420
tctaattttc taagacaaga gtttgtagct cgcttctagc agagttaaag aaaaagatat 27480
gatgaatggc ctatatattg aaaactgcag aagaaataat ctaatatgga atgggtgagg 27540
ttgctacagg tggtgcaaaa ataattgcgg tttttgctgt tactttcagt ggcaaaaacc 27600
gcaattgttt ttgcacaaac ctaataatac tgtttgaaat tcataatttc tggtggaaca 27660
gtcttgagat agagcgggaa cttagattatt ggctcataaa acagcatgca gttactgtta 27720
ttccttaatc ttatatcttt attatgaggt aatgcttctc ttacttgatt tggagaatgt 27780
tctcatcttt ttctcatata tttgtgaaa ctaaatgtta aaggaattaa cataacaacc 27840
attgctaaat tatgatgtaa aataaactg tgtctgcaat tagcctaggt agttttactt 27900
ggttcaataa tataaaatca cttaggctcat aatatttttag taccttagtg tgattggcca 27960
agaaaaaat tcacgttacc ataaaaagtc aaaccaagta acaattaatc taaatttcaa 28020
ggaaatataa agcaaatgat tatgttgtaa catggaacat taatgtaaat aactcccaa 28080
cagttcatga aatagattct gaactggcag ttgacaaaga aaggctagtt ttggcttaat 28140
tctcttatgc tccaatatct ttctgaagt cttatttaat ttcacttgcg ttagcaacag 28200
tactctgaga tgggtgtgca tctggaaaa tatcacttct ttcctgacaa tcagtatctt 28260
taatttaatt cttacatatg ctttctggaa atattacatt catacttga tgacgtctta 28320
aaaataatth caatgtaata attatacatg tgtgttgagt aaattgtgtt ccgtaaaaga 28380
tacttttaaa tcctaattgc tgtatttggg aatgtgacct catttgaaa cagggtcttt 28440
gcaggatga tcaagttaa atgagatttt actggattag gatgggtcct aatacaatga 28500
ctgtgtact taagaggaaa atttgtacag agacatacac acacacaaaa gaatgccatg 28560
taacgttagg cacggagaac atcatgtgac aatgaaggca gagattggag tgatgcatct 28620
atgagccaag gatggccagc aacccccaga agctggggag agacaaagga tgcttcccta 28680
ggatctccag agcaagcata aatctgccag caccttgatt taggacttct ttctccaga 28740
actgtgagaa tagatttcag ttctaagcca cccagtttgt ggtactttgt tataggagtc 28800
ctaggaaact aatgcaacat gctatgttgt aaataaacag tagttctaata tggataaaat 28860
aataatttta agtatatttt attattaggt caggtgagta ggaaattgcc tggtagacta 28920
attatcaatg ttgtattttc tgaaaagtta aaatattccc attcagtaaa taatacttgt 28980
ctaaaagttt ctatctgtat tgcctataaa caaagcaaaa attaaatgg agtcagttat 29040
gtacctgaaa aatgtgaaaa aggaaagaaa aatcttttag gaaagatttt tcctaaaatg 29100
gctatacatt tgagaaagaa ttttaaaact taaatcatct ggctgtttat tccatttttt 29160
tgtttacatt atcagataca ttatgtatga tttcatacat tatgtatgtt acattcaatc 29220
aatatttatt aaatacctaa aatttgctag gccttgagct aggtattgta gataggtaaa 29280

```

-continued

```

taagacaaac attcccttcc cctaaagaga ttatitttggc atgaaattta atctgtagaa 29340
gaagtgaag aggccatttg cagagtctac catgaaaatt cattttcact atctgctgta 29400
gttttacctg atcctgtgaa accatttttag catcattttg ttttactttg tgcttctgta 29460
ttatattgga gaaaataaaa tttacttatt gcacagacaa cataaaatta attaaagtat 29520
gtgaaaccaa aaatgtaatt gtttatcaaa ttttttttag tgtttcccat tcatcagtat 29580
ccaaatacat tcttagttgt gcatccatgt tagcataact aaaatgtggt gttcttattt 29640
tttcatctaa ttcacaatca acaatgttca cttttccact ttttttttt gagacagagt 29700
ctcactctgt cacaccagg ctggagtgca gtggcatgat cttggcacac tgcaacctcc 29760
gcctcctggg ttcaagcgt tctcctgcct cagcctcctg agtaattggg attacaggcg 29820
cccgccacca tgcccggcta attttttgta ttttagtaga gacagggttt caccatgttg 29880
gtcaggctgg tctcgaactc ctgacctcag gtgatccgcc cacctcggcc tcccaaagtg 29940
ctgggattac aggcataaac caccacgct ggcccctcat catttttata att 29993

```

<210> SEQ ID NO 2

<211> LENGTH: 29993

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

<400> SEQUENCE: 2

```

ttggtttttt ttttaagggtt gtttttaaat gaaattaaaa taaaaattat gttttggatt 60
atgagttttt agagcgtttt gtttaatttg atttatataa aatttaatgt taagatagtg 120
ttatgtacgt ttgaaaaata aataaatagt aaatggtaat tgtagaggta ttttatttta 180
gaggagtggg ttaagggaaat tgtaataaag tgaattttat atttttattg atttatttta 240
taaagtttta tattttttat tttttattt tttttaaaat atattttttt ttagattgag 300
ttgaatatgt ataaatattt tgagagtttg atttatgtta tgtagggttt attggtggtt 360
ttgaattatt ttaagaattt ggtagtaggt ttttgatttt aatttttttt atgtgttttt 420
ttagggtttt tgcgtttttt ttattgtttt attttttttt ttttttttat tgtttgtatt 480
gtattgtatt attgggttaa tagatgttag ggaagggtaa tgagtaatcg gtataatatt 540
ttaggaaaat gatttgtttt ttaattgtgt atttatgaat tgtttgtatt ttattataaa 600
taagatagta agtaaaataa taaataatat atgaatttta aaaagacggt ttaattttag 660
agattttaag tatatttttag ataagaaaag taataattaa ataagtatag aatttaaaaa 720
gtaatatttt aaagtaaatt ggatataata ttttgtttcg gttatgaaaa ggggagggtt 780
gaaaaatatt tttggtgagt ttgagggata tatggaatgt aattttttta attaatgttt 840
tttaaattgt atgtgagata attttaggtg gtatatgaga aataattttt atttttataa 900
tataggtttt tttgtaagat tattttttat ttatggcggg ttttattagt tttttattta 960
tggtgattat aaaaaatggt ttttattaga taaatttatt taataaaaaat agattattta 1020
aagatatata ttaagttaat aatataggcg atatggttat aattataaaa gtgatattta 1080
atgtaaatgt tggaaataag gttttattta agtaataaat atatttttgt ttttataaatt 1140
ttaaaatttt gttttttatg ttaaatttat tttaaatgtg tattgaaaat tgacgtttta 1200
tttgatattt agtgggtgtt atgttaatat atgtatattt aagataatta ttagaaggta 1260

```

-continued

gtaatatgaa gttaaataat taatatatta ttttgattt taaaaataat tttgtatat	1320
gtatTTTTat ttagatgtat aaaaaattat gtaggagtaa agtgattttg aaagtaaggt	1380
tagtgtaatt gaaaaattg attttagaag ggtataaaga aatatgttaa gaaaaattat	1440
tttatattta tghtaataat atgttaaagt atgaattttt atttatttta aattatata	1500
ttttattttt tattttagg tgaaattagt atataatgtt ttatgttaaa atgaaattta	1560
aatgttggtt gttttttgag gatagtatta aatttattat tgtatttaaat ttttatatga	1620
tatttgttat agttatagaa tttttttaga tattatgatt tataaaatat ttaaattgta	1680
tatttggttt gtaatatatt atatatTTaa aagaattatt attttttga atattattaa	1740
tattaataat agtattatag tgattaagag tggatttaaa taaaggttg tttgaaattt	1800
aggtttgta tttattagta agtttttaa aattttgagt ttttagtttt tttatatgtg	1860
aaatggagaa aataaattt taaaattaat taaatttgtt ttttatatt tttttttta	1920
gtgtatatta attggtattt ttcgttaggt tagaatgtgt ttttaattat gttttaaatt	1980
cgttttgtgt taattttatt gttagaattt tttttattt gagaattaga aaaggaaata	2040
ttatgtttgg taatgtttt attttttaa ataaatttgt aggaaagaat atttagtaag	2100
tgatgagagt agtaacgatt gtttttatta ttttaaattt ataattatt tttttttaga	2160
gttttttaag tattgtagat aattttttat ttattaaaa ataaattgta attataagta	2220
tttaggttg atattgtttt tgaattagat agtgtttata ttatgttat aagattaatt	2280
taaagtagag gatgaaattt tttttttgaa ttttttttag aacgtaattt agtgaatata	2340
ttaaattaa atttttttg aatgggagta atttttacgg attaatttgt aattttttag	2400
attatattta aggtaatgta gaggtgtttg tattataggt tttgtgatta gagattattg	2460
gatttgtttt tggaaaagt atttatgatt ttttgggttt tggttttttt attttttata	2520
tcggtttaat aatgattatc gtagagttgt tatggaagtt aaatgaattt atgtattata	2580
agtgattaat ataagtattg atatttagtg atttattaat aaattaaaat atttattatt	2640
aatatgatag agaaggtgtc gttaaaatag ataataggtt tttggaagag gtgattaaat	2700
ggatgtaaaa tttatggatt gtttatttcg tttatttttg ttgtgttttt tggttgtggt	2760
atatatacgt tgggtataaa aatcgtaaat tttatgtagt cgcgtagtgt atgcgtagaa	2820
ggttagata cgaatgtta ttttagtaat gtgttttagag aagttttgac gtcgttttg	2880
aagtaagtcg ttgtgtttt atttttgggc gtttgggacg gatgtttata tttgtattta	2940
gtagtattgg aaggggttta ggtttttcgt agtatagttt atttttagat cgttttagtt	3000
tttataatat atatttttac ggaaaaggtt attttttttc gttagaaaa gcgtttttagt	3060
ttggtttggg ttggttttta ttttacgttg ttgtaagtag gcgaagtttt tttgttttt	3120
ttttttggg taagtggaaa ggagttcggg aggggttcg tagtggtttg tataggggaa	3180
ttggtagcg agagagtttt aggtaatttc ggggttgtt ttatagaagt aggtgggat	3240
cgatagtggt ttttcggttt agggaggaga gcgcggtcgc gggtttttt ttttagtttg	3300
gaggtttagt tcgttcagat cggttcgggt gggggcggg tgggggcggc gcggagggta	3360
cggagattac ggcggcgtta ttcgggatat ttagggttc gaggttttg gcggtttta	3420
cgcgagatcg taaattatga taataggtag ttattcagag ttaataaaa acggagtggg	3480
tttttcgccc gtcgctgttt ttcgctttt tggcgtttt tttcagggtt ttcgctgtt	3540

-continued

ttacgagttc gtagtagtgc gtggcgacgt cgttttcggt ttatTTTTTT gcgtaagtgc	3600
gaggttgc gtagcgcggc gtacgtttcg gtcgttttcg gttttcgcgt aaaatTTTTa	3660
ttttgtttac gtgaagttgt cgttgtttta gagaggggga aagagttgcg gaaaaagtgc	3720
gggagtgc attgcgcggc ttgggcgcgt ttttttattt tttttttttt tttttttttt	3780
tttgctgtag ttcggagttt tggttttttt tttttttttt tttttttcgg agtcggtttt	3840
ttttttcgtt tcgttttttt ttcgtttgtg tacgttattt gttgtggggg ggtcgaaggg	3900
gatgttttgt ttttattaga ggtatagcgc gaaggggaaa tttcgatatt ggaaggaacg	3960
agaataaata ttaattacg gacgtattga atcgcggttg gtagatata ttcgggaatt	4020
cgaagcggat cgggcgcgca ggtgagtgat tttttttttt aattttcggt ttagggtttt	4080
cgggggagtt tgagttgaga gaatTTTTaa atttttcggg aaagtgcgcg aggtttcgtc	4140
ggggacgtgc agcgttgggt attgagacg cgtagttgga cggtcgctgg gcgtttgcgt	4200
tttcgggggg cgtttggagg tcgggtgttt tacgtttgag ggttcggggtc gttcggatcg	4260
tagcgggtgt ttttgtttta gaagacgttt ttaagtttta agggtttttt tcgagtttgt	4320
ttgttttttt cggggtcggc gcgagtttg cgcgtaacgg agtttattta gtagtttagc	4380
gcgcggtttt tatttgtatt tcgtttttat ttggtagagg cgcgagtatc ggggtttttt	4440
ttatattttt tttatgacgt gtattatttt ttgatgattt tttagatggt ttaggcgcga	4500
ggatgttgat ttagagtttt tcggaggggt ataggcgttt gggttttttc ggtgtcgggt	4560
gcgtgtgat tttaaaggtt cgcgttttaa tttttaggtt ttgatcgggt tttttaattg	4620
cggcgatttt attttaatag tttttatgtg gcgtggattg aatgtttttt gtagtttgtt	4680
agggtcggtg aaattagagg cgttttggtt gagtagtgcg gtttattggt tcgagtagcg	4740
ggtgttatgg aaggtttata atttttttaa aggaagggat ttggttgggt agagtaggtt	4800
tttttttttt ttttaagttg ttgggtttgg ggaggtagtg gaatttgaaa tggttcggat	4860
tttttagcgtg gtgaagcagc gtttggaagt agacgtgtgt gtgtttgttt tattttgcgt	4920
cgtatagtaa ttcgaatttt cgtttggtag tatttgaaag agtttttttt ttttgtttgc	4980
gagatttga atagttcggc gcgattaggg aatttcggga tcgagttcgg tggtagagtt	5040
ggggcgaaag tagagagcgt aatttaattt ttgttatttt ttttttgttt gggaggatag	5100
tgtttttttt tattattatt tttttttttt ttttttatga agataacgga tttgcgtttg	5160
gggtgagagt gttgtcggga gagtgtgtg gagattgttt ttttttatcg cgttttttgc	5220
gttttttttc gttatttcca gcgggtttag agagttattt atgaatttta atttgagggt	5280
aggggaggaa ggtgtaggtt tttttgtttt tttgttaag gtgtagaata gcgttcgggc	5340
gtgtgttttg gttttagagt agttttacgt ggagtaattt cgtgtgtgtg tgtgtgtgtg	5400
tgtgtgtgtg ttgtgtattt gatttgtgag gagtaatag gattttgggt tttaaattta	5460
gtgggtcgtt ggttattagt ttgttttttt ggttgttatt atagatattt ttaaaatttg	5520
atatttaaga gaattaatag gttaggtttt atattaaggg tttatattta atagtttttt	5580
ttttttttta ttttttttgg tttggtattt tgggattaac gtaattgttg gagcgaata	5640
tatttttttg aatattggtt tttgtttttt tttttatttg tggttatttt gtgaattttc	5700
gggtgtttgt tagttttagt gcgggttttt gcgggattta ggtgggagtt ttaggagcgt	5760
ttaataatcg cgggtttttt attagtagtt tttgaagttt tatttatatc taggtgattg	5820

-continued

ataggattga aagttgacga tggTTTTTTT gtttGttGcg ttgTTTTTgt tttTaaAatt	5880
ttagggaagg gattgtattt gaattTTTTT ttcgggtacg gtttggTTTT aagtagaatt	5940
agtgtTTTTT tttttTTTT tgtttTTTaa aatattattg gtaagTTTaa atttgaagaa	6000
ttaaaatta gaggggtg ggagagaatt ttttTaaaa atatttgata cgtgatacgg	6060
agcgttttag gagattgat ttaaagatat ttgtgtattt ttaaaataa tattatttta	6120
agaaataaaa agtagtagta ataagagata gattgttttg tgtggagtaa gattgttaga	6180
atttgatttt tatgtaata atatcgaaag tagatataat tataattata tttattgtta	6240
taaatcgtaa aaatagtttg ttttatttga ttaaagtTg taagtattt aaagtTaaat	6300
tcgtatttag gaatgattg ataagaatgt taaaattttt gatagttata gttttgaaag	6360
taatataatt agatttgttg aaaatgtttt ttttttttga ttattagtga tatgtatttg	6420
ttttttataa taagtaaagg tttataatta atatggtttt taaaagtttg ttttttttg	6480
taaaattttt gaaaatagta tgttttaata tttaaagatt attgtaatta tcggttggtt	6540
tataattata attaaataa gatttattta gttgtttttt ttttattagt atttgttgaa	6600
tatttgaaat tgaatttttt atttttTaaa taaaacgttt tataatatag tatgtattta	6660
taattaatat ttttttTaaG tagtaggttt aagttttTaa atttttTaaa ttttaagatt	6720
tgtataaaaa ggagtattg tataaaaagg aataataatt attaattgtt aagaaaaata	6780
gatgttgaat ttagagggt ttttagaagt tggagaaaa aattgtatag aagtTTTTT	6840
tttgacgggt gattgttaga tgagtattt gttattttgt tataatttat ttttttttt	6900
taataagggt ttttggtta aaaaaaaaa aggaaaaata agtattttat ttttaagagt	6960
ttttttttta tttatgtttg attttgggt gtaaaataat tttatatagt tacgggaaat	7020
agagtattta attttatgtt ttttagtttt gttgtaagtt ttaaagatt gatgtcgtgt	7080
ttaaaatat attttatgg tttgataatt tattttTaa ttttatgtat tggaaataga	7140
tttttttagg tttttgaaa gatagaatga tagaaagacg atgattattg ttataaaaag	7200
ttagaatgtt aggaagtTta tattttTaaT ttaaataat taaatttatt ttgtttTata	7260
aaaattTata gtttttatat agattttTaaG aagagggtg ataatttttt ttattattta	7320
gttaagtgtt aagtaattta agaatgtaag tatgtattat gtttttttat tattattttt	7380
ttttgttagg taaatgtaag aatattggat attttgTaa aataggtttt tttttataag	7440
gagttattt atttatgttt tgtattgtag ttgttgatta tttttatgg agttgtTaa	7500
gtattTaaa tttaaatta ggtattgttt gttgtTaaat attttagatt atttattTaa	7560
gaaataaata gaaaagtgtg ttattTaaaG taggagacgt tttgaatttt tttattgaaG	7620
agttgaaata atttttTata aaatattttt tttttttttt agtgagaata taattttgat	7680
agtttttttt tttTaaatgga tattattttt atatgtatta aatgtTaaat tttTaaacg	7740
tgtttttttg tagagtattt taaattaatt atatttagaa atatagttgg gggattttat	7800
tataatgggt attaggtgaa gaaattaat attaggggaa tgggggtgta ttgtagtgtA	7860
tttgtttttt ttatgaattt tttttgtTaa attaagaaat gatatgttgt tttaggtatt	7920
tgttttggtg atgggtgaga tgtaatatgt gtaatgttgt attttatagt tagatgtgtt	7980
ttaatgaagg ggatattgta tagttattaa attatttttg gagttaaatt tggcgttatt	8040
tagtttgaat ttaagtggaa gaaaatgaat aagagttata tttttTaaag aagtataagt	8100

-continued

aatthttgatt gttthtttaaga ggtthttgaaga tthtttgtaaatt attatttgthta thttaatattg	8160
thtttgthgagth tghatattthaa thatatgthttt ggtghatathgt thttthttacgt thgaaattthg	8220
thttthttthtag thgtthttthttta thttthttthttt thttthtttagthaa ghagathattgha aaathghaatath	8280
thttthttthttt athttthttthttta thttthttthttat athttthtttagthaa aaaghagathgha aaththghattthaa	8340
thttthaaathag aaaththttthttt ghthghthttthttthaa aaththathatttht athathagthghat thghagathagth	8400
thttthagathagth ghthttthttthgt ghthttthttthttt aaththttthttt ghthttthttghaa ghthghthttthaa	8460
ghthttthtttaggh thattghatathata athagthathagth thghatathttthg aaaththttthttta ghgththghgthttt	8520
ghgthttthttthttt thttthttthgt ghthttthttthgt athghghatghghat ghgaaaghathata thttthttthttthaa	8580
athattghathaa thttthttthttt thttthttthttat ghaaathghthttt aathathagthath ghthttthatttht	8640
ththaaaghthttt thttthttthata thagthaaaghthaa athttthttthttta acghthattthttt ghghathagthaggh	8700
aghaaaghathttt athathghthttthttta thghthghthttthaa athttthttthtagth thttthttthttt ththghatathaggh	8760
aghthttthttthttta ghgththataathttt thttthagghthaa ghththghattthttt thattthattthata thghaatathttt	8820
thaaaghthghaa ththghghthttthaa ththaaathghthgha athagthttthttthgt thattthattthaa ghthathattthata	8880
athghathagthghat thaaathathghat ghthathattthttt thttghagthata thttthattthttt aaaghthghattth	8940
thttthaaathaa athghthghthattth ghthagthttthttth ghathaaaththata athttthghthghat thattthattthatt	9000
athaaathattthaa thttthattthattth thattthataath ththathagththtag thattthattthttta ghghaatthttthghat	9060
ghthaaathattth ghthttthaaaghthaa thttthathathgha ththagththaaathttt ththaaathttthttt thaaaththathghat	9120
ghthaaaththaggh thattthttthghttt athttthttthattth thathagathaa ghaghththgha thttthagththattgh	9180
thagthathagthghat ghagththagghat thttthaaathttthttt thagthathghaggh ththagthattthttt thattthttthaa gh	9240
ghghthghthghat ghthttthttthgtthttt thattthagththata thathattthattthata thttthttthagthghat ghthattthghaat	9300
athttthghthttt aathathattthttt athattthttthattth thaaaghghathaa thagathattthtag aathaaathttthghat	9360
thaaathttthghat aththaaathttthgt thaaathttthattth ghghthaaathattthgt thttthghagathaa ghaththaaathaa	9420
aghthattthghaagh ththaaaghathaa aathaaaththaaaghtht athttthghaaghthaa acghthattthaa ghthaaathghath	9480
thaaaghathathghat ghththaaathaa ththathagththatt ghaghghththgha thagthaaathgha thghaatthttthghat	9540
aghaaathghthghat thattthttthghthaa thghthttthttthaa ththattthaa ghththattghath ghathttthttthaa	9600
athattthttthaaath aathathghthgha athttthttthttt ghthttthttthaa ghthttthattthaa athghththaaath	9660
thghthattthaaath thttthghattthttt thttthghthttthttt ghthttthttthghat ghthttthaaathghat thghthttthghaa	9720
thttthaaathttt ghthathattthattth ghthghaaaththaa ththaaathaaaghtht thghaaaghaththt aaaghathghthata	9780
thghatghghattth ththghagghthag ghththaaaththata thaaathghghattth ghthttthattthttt thghghaaathata	9840
thttthaaathttthatt ththghghthattthghat thathghthghthaa ghthttthttthghthaa ghghthaaathattth ththghthghththata	9900
athaaathttthghat ththataathttthttt athagthattthttt thttthattthghaagh thagthghghthttt thattthttthttthata	9960
thttthattthata thattthagththata ghghathttthttt thattthaaathghat ghaaaththghatghat athathghthttthag	10020
thagathaaathghat thghthattthghthttt ththagghthghaagh thagghathghthag ghthagthttthghat aathghattthattth	10080
thaaathghattthttt thttthttthacghghat ghghthghthghat thttthghagthttt thttththaaaghthghat thttththaaathghat	10140
athghghaaaththghat ththghthaaathaaath thttthaaathgtthttt thttthghathaaathaa ghghghththaaathttt thagathattthttt	10200
ghthttthaaathath ghthttthattthghthghat thattthattthaaath aathathagaththghat ththagathacghghat thaaathattthttt	10260
thttthagthghthghat thathattthagththata ghaaaghghthghat thttthattthttthttt thttthttthataath ththghagthattthc	10320
ghthattthaaathata ghthttthttthaaath thghgathttthagthttt aththghagathaaath ththaaathattthgha ththaaathattthghat	10380

-continued

ttttgtgtgt agttataggt ttgaatgagg tataaaggat ttgtatttgt aaaggttgat	10440
tttattaatt agaataatattt ttttttttta aagattgtaa gaagaaatat tagtagtggt	10500
ttaatttgat atgatttttag atttttacgt aaattattgt tattattttt gttatttttt	10560
tttttttttt ttttaaaatg aaagggatat tttttgtgaa agattataat taaattataa	10620
aaatttatat ttatgtgtta ttaagtttaa ttttatttat aaaagtaata gtatagagtt	10680
tgaatttta tttttaatta agtagtgta ttatatatcg ggagggttta ttatgtataa	10740
ggttatata ataatttata gatttttgta tatatttaac ggagtgattt atttattata	10800
ttttattttt gattttgaat tttttaatgt taaaagattt gaaaagaatc gaatgttttg	10860
attaagagat tgaataatatt taatttaatg ttttttagtat gttgaaagt atgatgattt	10920
gggggaatta gtagattttt atattattta attttttttt tttatatttg aatgtaaatg	10980
tataattatg tgcggttatg atttaagtta tttttgttaa atttaatgac gttgtaggtg	11040
aattatattt agattttttt ttgtaggttt tttagtaatt taaaataatg tttttagtag	11100
gtaatttaag tatgtaaat ttaataaatt tgtaagaac ggtaatttta ttgttttatt	11160
ttgttttttt tttgttttg ttgtttttgt attaatttta gttgataaga tgatggtatt	11220
gttatttttt ttagtgtatt tatgaagaat ttttaatttag gtttagtttt tttttttaat	11280
ttgtgatttt agtttttaaa ggtttcgttt atgaaaatgg tttagtaaagt tgtgggtatt	11340
ttgtaaatgt ttgttaaatg tttttttttt attagtgttg ttgaagattt gtaaaattag	11400
agtgggatgg atagattttt tttttatttt gtatggtttt gaagattttg gagtttttat	11460
tgtaatttta tatttttata tattattttt aggaatattt agagagagaa ttattgtaat	11520
taaggttata ggttaaat ttaggtattt aaattagttt taggtggata attgttttat	11580
atataattgt tctgtattaat gttgttataa taaatatagt tataatattt tgatgttttt	11640
tttgggggtt tattttggat tgtgttgaag ttgtttttaa tttttttta cgtttttaag	11700
atatttaatt atattgttt atgattttta tgaggttttt tttggattta aaatttttta	11760
aaaaattgt ttgttgttta tagggtaaat tttaaaattt tttgtaggga agataatttt	11820
ttatattgag aattttgttt aggttgtttt tttatttttt tattttttat ttataataat	11880
tttttttttt agttgtatg aaaattgtag tttttaaata tattttggat atttttattt	11940
tatttttttg tgcgtggaat gttttttttt tggaaaatat cgtttttttt tttttgtaa	12000
tatttttttt attgtttttt taatatttaa ggtgtaaaat gttagttttt ttgataagtt	12060
tttttgattt ttagagtttg agttgaacgt tttttttttg ttatttttaa agtatttagt	12120
ttttatattt aatatagttt ttgttatatt gaaatgtaat agattatatt agtttaaat	12180
ttgttttaag ttgagaattt attggaagga ttagtgtat ttgttatttt tgtatcgtta	12240
atgtatagtg tattgttggat tatatagtta gtaggtagtt gttatatttt tattgaatga	12300
atgagtagag gtattgggaa tgagagtaga gattgtgata ttgaatattt ttttttgag	12360
attgggaata gtgagaaggg tagtgattat tataaggttg gattatttta ataataattt	12420
tagggagttt tggataaaat ttagtagaat ttagttgtg aatttaggtt taggttataat	12480
gtatgtatac gtgaatgtat aaatgtgttt attttttatt attgggaagt tagtgatttg	12540
gtggataagt taaggatttt gaaaattttt ggagatacgg taattttata gtatttgaaa	12600
tttgagagat ttatttttta tagtatagtg agtaagtaga aagaatatgg gggttgggcg	12660

-continued

cggtggttta	tatttgtaat	tttagtattt	tgggaggtcg	aggcggatgg	attatttga	12720
gtaggaggt	cgagattagt	ttgattaaaa	tggtaaaatt	tcgtttttat	taaaaatata	12780
aaaattagtt	ggcggtgtg	gtggcggttt	gtaatttttag	ttattcggga	ggttgaggta	12840
ggagaattat	ttgaatttag	aaggtagagg	ttgtagtgag	ttgagattgt	attattgtat	12900
tttagtttg	ataatagagt	gagattttat	tttaaaaata	aataaataa	taaaaaggaa	12960
gaagaatag	gagttaggta	tttgggttta	gatttttagt	ttaatattga	tgagtttga	13020
atattaaata	tattattgaa	atagattttg	ttttagtttt	ttatgtgtaa	tagaatatgt	13080
tttatgggt	ttgatgaga	atthaatgag	taaatatag	taaatatatt	tagagtagtg	13140
ttgatatga	gtattagtgt	ttattattat	gattgggtat	ttttatttag	attattttta	13200
aatgttttag	agatatttag	agttgaaggt	ttttttggg	gggttatgga	ttggtgatta	13260
taaatggagt	gttttagtag	ttaagtataa	attagatatt	gtgggttatt	ttggtaaatt	13320
tgaggttga	ttggagttt	taagagaaga	tattataggt	gtttaatagt	tggtattggt	13380
tatttacgtt	gtgtagtaat	atttttttta	gtaaatattt	tcgtatatag	taatttagat	13440
tgtttttag	ttttaatttg	tggtgttttt	taattatttt	aagtaattat	aagtacgttg	13500
tataatggta	gtattttata	gttaaatattt	aagttttttt	agttgttttt	tttttagtcg	13560
taaggtatta	attatgagaa	tttgtataat	ttttaatagt	tagtgggttt	gatgaattta	13620
atgtaaattt	tttaaagggt	agttttatat	attgttatga	ttttattttt	tagagagata	13680
gtaggaggta	attagtattt	ttatttatag	aatttattat	ttggatttag	atataatggt	13740
tatgtgtag	gatggggtg	aaggaaaaag	gaaagttaaa	atattaattg	ttaaggtttt	13800
atttttatag	ttgttttga	attgttatta	ggtgtatttg	tatagtttta	aaaatgataa	13860
cgtgtgattt	gaatttttta	gtatatttaa	ataagagttt	tttaagtttt	ggttattgat	13920
attttgggtt	ggataattta	tttaggggat	attgttttgt	gtttggtagg	atgttttagtg	13980
gtatcgttgg	atgttaggga	tttttttttt	tttgtaatg	atggtgtttt	tagatattgt	14040
taggtgattt	ttgggagtaa	aattgttttt	ttttgagaat	ttataattta	agtagatttt	14100
aatattattt	gatatagtta	aaaaattttt	agggttttgt	ataagaggat	tgttttattt	14160
ttagttaagt	gtggagaatt	ttaatatttt	ttttttaagt	ttaaataatga	agttaaataa	14220
tttttattag	tttgtgttga	tagattaatt	ttgtatagag	gggatttata	gatacgtttt	14280
atatagattg	attattttta	gtatattttt	tttttttgga	aataatata	gagtgggagt	14340
aatagaatat	ttgagagaga	gtatgtagga	agtaaatatt	tatttgaatt	atttatgtgt	14400
tttttttttc	ggttatttgt	tagtagatag	tttgattta	tttttaata	atgattattg	14460
tttttttatt	tttgtgggta	aatatttagt	aataagaaat	ttttatttta	tttagtttgt	14520
agtgttagaa	aaggtaaagt	tattgataat	tataattggt	gaagattaaa	atatttagtg	14580
agttaaatta	ttgttttga	atgagaaatt	tttaaagaa	ttatgtgtag	tttttagttg	14640
taattataag	tttaagattt	tgagttaata	atatttgtat	agttagaaga	aagtaaaaat	14700
aaatataatta	tatgtaaaaa	tttttttttg	tgaattttta	gatttaaatt	tttaagtttt	14760
tattttatgg	gtttgatggt	gattaatata	ttttatttat	gaaatgatat	ataattgatg	14820
taagtaaaatt	tgattgttgt	ttgggtatta	tttttatata	atthaagtat	atatgaaata	14880
tattttaaaa	attatttga	taaaatggga	ataataatgt	ttatataaag	tgagagagga	14940

-continued

taatTTTTT	tatatgagtt	ttgggTtgaa	TTTTAAAatt	TTTTGAAat	atgAAatatt	15000
aatattataa	atatatgatt	agattaattt	TTTTAagagt	TTTTAAtagt	agataatttt	15060
tattttatta	gaaattgtaa	agatatttga	aattttatat	ttatgtattg	TTTAatttag	15120
tggttgatt	TTAattgtat	taggtagata	attatggaag	tagttagatt	aatttggaga	15180
TTTTatatta	TTTattaatt	gatgtttaa	taggtagggt	agaggatttt	gaaaatata	15240
atatatata	atatatata	atatatata	atatatata	atttattttt	TTTTTTTTat	15300
TTTTTTgat	ttgtgttTta	attttgAAat	gtgtgtttag	gtggtaaagg	gTtgtaaata	15360
tagtatattg	Tgtttgatta	aaatattttt	ttgttttTta	ggTggagaag	gattttgtat	15420
agTTTTTgt	tagtagagag	ataaagtggg	Tgttgggtaa	TTTAatgata	ggtggaggTc	15480
gTTTTcggg	agtgttgcgg	agaatttgag	ataggggttt	aaattataga	tagtggagga	15540
gagTTTTgtt	TTatttatgg	ataggaagtt	taaagaaggt	atagcgaatt	ataggtttat	15600
Ttaggtattt	tatgttagat	attgaacggg	aagtttagta	gTTTTtagt	tattggagaa	15660
cggTTAAAA	TTtagagttt	Ttattttttt	aaatgagtat	TTTTTgggtt	Tataaattaa	15720
ggaattttat	agagtattta	agttatata	TTTTTTTTt	aggttgattt	TTTgtatagt	15780
gggtTTta	gtgtaagat	Ttatttatga	atattttatt	TTTTataatg	Tcgttagtta	15840
TTTTTTTTt	TTTTTTTggt	atatttttat	Tatttttatt	TTTTatgggt	TTAatttttt	15900
taatttattt	gTTtagtggt	TTTTgttgt	Ttatttggat	AAAataata	taataataat	15960
aatttattta	TTTTtaagg	TTtagtttaa	atgttatttt	TTTTgtgaa	TTTTTTTTaa	16020
ggTTtaggta	gaattagttt	TTTTTTTTgt	gtaatattat	agTatttttt	Ttggattggt	16080
aatataatag	tagattgtgt	Tggggtttat	Tgtatgtgtt	TgtTTTTtag	agagattatg	16140
aatttataga	ggggagagat	Tatatttttt	TTTTTcGta	atttAatat	Ttggagatgt	16200
taatgAAat	taagtattg	Tattatttta	Tgtatatttt	Taaagggtat	taggttttaa	16260
attagatatt	TgtAAaggat	gaattttgtt	aattttttgg	aaaattgggt	TTTTTTTgtt	16320
Tgtggttatg	aaagttgggt	tattggttat	TTTTatttt	aatttattta	aagttttatt	16380
Tgtattaaag	tattattgat	TTTTtagtgg	AAAAataggt	atgttcggcg	Tatattggta	16440
TTTTaataa	attgatattt	AAAagataga	TTTTcgttg	aggtttatgg	TTTTataat	16500
TTtagttttt	TTtagaaata	gattttgatt	TTTTTggttg	TTTTTgtta	agatcgattt	16560
tatgttattt	Tgaagaattt	Ttaattttta	gaattataga	TTTTattga	gagaaagtat	16620
gTTTTgtatt	Tgtttgtgtt	agTgttttaa	aatgtgagga	Ttaaaaattt	tagatgtatt	16680
TattAAAAat	atAcTggtt	TTTTgttgag	gGaaattata	TTTTTTTTt	taggtataat	16740
Tgaggaagta	AAAAaatttt	gTTtagtttt	tagTggtttg	TTTTtaggta	TTTTTatata	16800
aaagatgagt	aagatttttt	AAAAattatt	aatttgttta	TTtagggggg	aaagtttttt	16860
tatgtttaga	aataatatta	aattaagatt	attgttttta	atagttttaa	aaattgtggt	16920
TTTatttttt	Tgtatgagtt	atttttttagt	atagTtagt	atagattgt	attatattat	16980
gataaagtat	atttgttttg	Tgttttagtg	ggtattttta	TTTattagcg	Ttagatgatt	17040
agattagaag	agagatttag	TTtaattttt	Ttattttgta	gatgagatgt	Tgtattatgt	17100
gattttaaga	TTTTaagtta	TTTattttaga	ataaaataga	aaaagagaat	TTtatgttgg	17160
ggatatttta	Tgtttttaat	TTTatatttt	attattttat	TTTTatttgt	Ttgtttttga	17220

-continued

tgaggagtg agaatggtg aattagatta tttttgtgtg tatttaggtt ttgttattta	17280
ttagcgatgt gattttggtt atgtaatfff ttgagtgggt gtaagttata gtttttagtat	17340
gtgtatatgg ggaataataa tagtattcgt tttatggagt tgtaaggaat aattgaattt	17400
atgtaggaaa gatatttggg ttagtgtttg gtttattata aaagtttagg ttatagtttg	17460
aattgatatt ttggaataaa atgttgagt tatttgatat ttagagaagg aaatttatgt	17520
tgttaaagta agttttttt ttttaattgt ttttaaaaa atataataa gttataattt	17580
gtataatata ttaaaataga gatgatagag gttatttttg agaggaatga tttaggggag	17640
taagagagag gttttggaat gttgattatg ttttttattt tggtttggtt gttggttatg	17700
tgaagtgtt ttttaataa aaatttatta gtgattaatt tttgattat gtattttatt	17760
ttattgtgta tatgttata ttttttaaaa agttttaaaa atacggatat atttaagta	17820
aaaagtggat ttattttatt ttattttttg taaattagga taagtgtta tatttagatg	17880
tgtatttttt taggtatttt tttttttata tatgtgtaaa tgtaataatt atgtgaggat	17940
gtaatttatt tttattttat atttttaatt ttaataggat tattgatttt tttaaaattt	18000
gtttttgatt tgttgaattt tgaatatttt tttatattag tataatagat tttgggtgtt	18060
tttatttaat gtatgtagat attttataat tttattgttt tttttttggt gtatgtttat	18120
attgttttta tatttttatt attattaagt atgttgattt gatgattttt atgttttttc	18180
gtatatttgt taaaaatatt ttagatataa attttagatg taatattaat agataaatat	18240
gtatgtgat tttttataga aaatattaaa atgtttttta aaaagattta tgaattttgt	18300
atttttaagg taatattttt atttttttta taaaaattag gtgtaaaaat ttttaattta	18360
ttgaattttg taattatttg tgttttggtt aagaaaaatg gttatagtag taaatttttt	18420
gaaatttata tagtaatcgg aatattttat tatgattagg ataaataatt ttattaacga	18480
tataatttta atatatatat ataatttata taattagttt tttatttaag aatatttagt	18540
atttattttt atgaatgat attatattta tttatggttc gtaagttta atgttaggaa	18600
gaaataaatt agtaatatta tattttattat attttttttt atttagtatt attaaatatt	18660
tattagaatt attaggagaa atggttttta tttttatgtg attggtgaga atttatgagg	18720
aatgatatat tgtattttta gtaagagttt attttgtgtt aggtatagtt ttaagtattt	18780
tatatatatt acgatagttt tttaatatag agatgaaagg tatagatatt gtataagtaa	18840
taaatgatat tattttatta taaatgtgtt gataaatatt attatatatg ttataatatt	18900
agttgtgtta tattaatgat taatttataa atttgttatt tattataatt tgtttaatgt	18960
ggaaataaaa gttaatgtat taaaatttag aaaattgtta tttaaatggt ggttaaagta	19020
atttgataga tatagtttta aaatgaagta tattaattag ttgattttgt tgttcgattt	19080
ttatatgtag aaaattttgt ttttagaata tttagttttt gtatgttttg gttaaagatt	19140
tgattaattt ttaatgttta tttagtgttt ttaaaaagta gtttaattta gaatagtaa	19200
tgtatttttt ttgtattgaa taaatatggt tattaatttt tgtggggttt ttttttttgt	19260
agaatttttt gttattttta gggatttaat aggatttata tatataaaat tgaataattat	19320
attattatgg ggaagtgtt ttgtgttttt tagaaggaat ttttttgta atttaagtat	19380
tgtatgtgtg tagaatagtt aatatttttt ttttaagaaa ttttatgtat gtagttaaaa	19440
tatttttaaa ttgattaagg atttgttata tttttttaga agtatttaag agtattttta	19500

-continued

tggttataag agtagagtgt attagaatgt tagaaattaa tgaatacgtt agaggttttt	19560
aaaattgtgt ttaatatattt tgattttatt atataaatat gtgggtagt ttggaaaaat	19620
aaaataattt tttttttttt aacgtagggtg aggtttgtga tgataatatt ttaattgggt	19680
tgaatattaa tttgatattt ttttgtaga agtaattttt attatttagt aaatgaaatg	19740
gagaggtaat aaatatgatt gatatttatt ttttgagtaa agtattattt tatattagat	19800
tcgattagtt atttttattt tttttgaaa atgtattata taaattatta tagtattttg	19860
tttattaatt agggaagtta aaaacgttt tataaacgtt gaattaaaat ttttattggt	19920
gtgaggaat taaattttta gtaatatatt tttttatag ggtgagaaat tgaagtatgg	19980
agaaattata aggttttttt aaagtattat attgagtat ggtagatata gaaatttata	20040
attttgttt tttgatttta tttgaaattt tatgaaacgt aatttattta ttttttgata	20100
tgtgttttag aagtatattt ttttatttgt ttatttatgt ttttattat ttatttatgt	20160
atgtatttat gtattattga ttatttatgt atttatttat ttatttatgt attttaaaag	20220
aatcgggtgt tttgtatgat atagttttga tttagggtt atgaatattt tgaatttata	20280
tatttttagat atttttttta agaaatttat tgtatttttt ttacgattat gaaaaagaaa	20340
ttaagatatt aagaaaatta agttggattg ttaagtgtg ggttatagta gtgtttatgg	20400
gtagtatttg aagttaattt aggagtgaat ttagtttggt taagtgtttg aataagaatt	20460
tggttgaagt tttggaggag atgtttgggg tatagagtgg ggacgaggaa aaattatagt	20520
attttttagag agtattgtaa ggattagtat ttatatattt attattgat gtttttata	20580
ttggttagt taagagagag ggggtaaatt attttgttt ttattttggt atttttttga	20640
ttttagttt tacgttttaa atttttttt attgcgattt tttttttatt agtttaatta	20700
gtattttttc gcgttttttt tattaaattc gttttttttt attagtttgc ggagttatat	20760
ttttgatcg aattgttggg gttataagta atttgaatg aggaaagtgt ttagtaattt	20820
tttttagttt agtattttgt aatagagatt tttaggtacg tagtgatgat atttgttttt	20880
ttttttttt atagagagtt tagtgggtaa tttttttggt ttgttttttt gttttttttt	20940
tgtagtttta agatttaga ttttagattg gtttttggtt gtatggtttt ttattttttt	21000
ttttggaaga gttttttttt ggggtttttt taggtttaag agtatttga tttttgggag	21060
atgaatttgt agattttata aagaatattt aagtattagg taaattatat tttttttag	21120
tttaaacgtt agtttttttt ttttttttaa ttttattgaa gaataattta gtttttaata	21180
aataagtaaa tagaattttt ggggtttttt taaaggtgta tttttgtgt tttgaagatt	21240
attattatta atttttattt tagttaattg gttttttttt atatttttat ttgggtataa	21300
atgtgatatt ttggtatata ttaagtgtga tttatattaa tacgttaatt tataattgag	21360
tgatttataa ataattgatg attatgattt agtttatatt tgttttttat ttaattgatt	21420
ttttaagtaa ttagtttttt gtggttgatt gaattttttt ataatttttt gttgggaaat	21480
aaagattttt ttttaattg aagaagtggg tttgttaatt aggggtatcg tatttaattt	21540
gagagaaaa gattaggaat gataggaatg atagtatatt tttaatattt aatagtttgt	21600
tatatatgtg aaagttagg atatagtagg gaagtataac gttgagtga tttaaaaat	21660
taaatgtatt tgatatatta ttttttgaat ttttataatt tttagttttt taaaggtttt	21720
tggattttta gaagttataa gggtagtttt tggtaaaggt ggttttattt ttaatttagt	21780

-continued

tatagagtta	gtggaatgat	ttttatattt	aaatgttata	ggggaattta	ttgagaggaa	21840
aaagggattt	taaatttttt	tggatgtttt	aattaaaaat	ttggattaat	aaaatttttt	21900
taagtgtttt	aaatttaagg	aaagtaaata	gtattttatt	taattatatt	tttaattagt	21960
atattatata	gatataaatt	aaattagtaa	tttggattta	gatgaattag	atattttggt	22020
atattatata	atttttatgt	ttattttggt	gtgtttggga	tttttttttt	tttataataa	22080
agaatatgat	ttataggtgt	tatatattga	tttttgagaa	attaatgtta	tagaaaagtt	22140
gtttggaag	atataattgt	gattttgtta	ttgtttttgt	tgttattatt	tttttttatg	22200
gggatgggta	tatatatttt	gagatttatt	agatttattt	tgtagtttat	aaattatttt	22260
tttagatgta	tttaatttgt	tgtttagttt	gttttttgta	aaaaaattta	gtgaatata	22320
ttttaaaaat	atttttatgg	ttttttttta	ggtttggtgt	tatatatttt	tcgtgtgtgg	22380
ttttttaaatt	ttatttttgt	ttttttgtg	tttgttttaa	gtatatttat	ttagtttttt	22440
ttatattggt	ttgtagtttt	atatttttag	atttttaatt	tgttggtgtg	gttatttttt	22500
attgtgtttt	gtttttttat	gtagtttata	attttttttt	tttttttgag	atggagtgtg	22560
gttttggtgt	ttaggttgga	gtgtagtggt	acgattttgg	tttatgttaa	tttttgtttt	22620
ttgggtttag	gcgatttttt	tgtttttagt	tttcgaaatg	ttgggattat	aggtgtttat	22680
tattatgttt	agttaaattt	tgtattttta	gtagaaaagg	ggtttcgtta	tgtagtttag	22740
gttggtttta	aatttttgat	tttaagtgat	ttgtttggtt	gagtttttta	atgtgttggg	22800
attataggta	tgagttatcg	tatttcgttt	gtagtttata	atttttggtg	gatttttttt	22860
ttagaatttt	ttgtattttg	tggagtaatt	ttgtattttt	ttggagttaa	gggattttta	22920
attttttagt	tagtttttat	tttattttat	ttttttggtt	tgaggttttt	atattataga	22980
tgtagtaaat	atttagaatt	tatatattgt	aagggtgaag	ttttgattta	ttttatttaa	23040
gaaattattt	ttttttattt	tttttttaaa	ggttttatta	taggataggt	ttttttgttg	23100
aatttgatgg	tgggctcagt	ttttttattt	tttttttaat	aaatgtgata	aatttttttag	23160
attttttttt	ttgtggattg	ttttttgttt	ttgttttttg	tttatcgtta	taaaagaaga	23220
taatgataat	aaaacgataa	taataataat	gtaaacgtat	tttttttttt	taattttttt	23280
atattattaa	ttttttatga	attaaggtgt	gatatttgaa	agttatattt	gttttattag	23340
ttatatttat	atagtttttt	tgttttattg	atgtattttt	tatattattt	tttatattta	23400
tttttttagg	ttttttttta	ttttatgtta	gttttagattt	ttttggtttt	ttatttgag	23460
cgttatagta	gtttttaagt	tgttatttaa	ttttattttt	tgttattttt	gtttgttaga	23520
aatttgaaat	agatgtgttg	ttgtaatgat	attatatgtg	ttattttttg	ttttgggggt	23580
tttaagggtt	ttaaaggtat	taaaatattg	taataattaa	ttgttttaat	tgtattttgt	23640
atattttttt	ttgtttttgt	taatttaaaa	tgttttaata	ttttgttaat	attgaatggt	23700
gatgtttttt	tatttttttt	ttttttatat	gttcgtttat	aattgtttat	tgattatatt	23760
agttttttta	tattaaaggt	attagttata	attttgatgt	tatatagaat	gataattaaa	23820
aggtgataaa	tgtaaatgat	ttaaatcgagg	agttgtttat	ttttataagt	ttatttgagt	23880
ttagagaatt	ggtaaaagat	tggtgatttt	tggttattat	ttaaatataa	ttgggttaatg	23940
tagtttttgg	agattagtgg	ttagtataat	attttttagtt	gagtaggggt	aaaaataatt	24000
atattttgtt	aagttgtaga	tattattttta	ttgtatagta	aagaatgtta	aaatgtttta	24060

-continued

atttagtata tgaagagtaa aatataaaaa taggtttttt tttattttta ttagtaagta	24120
aagtggtttg tatttttgaa atatattttt tgtttaatta taataaaata taatataaat	24180
gttattaanaa gttttatttt atggatgaag atagtgatat ttagagatat taattagttt	24240
ttttaagatt atatattttt ttagtagtaa aagtaagttt ttataagttg atttattttta	24300
ttttagtttt tatttggaat tattaaattt atattttttt tagtaaaatt atcgtttata	24360
tttttaggt ttttaagatt taatgtttaa tttgggagat ttatttttaa atttttttta	24420
agtgtaatat atttttttta ggatgattat aaagttatta gatgtgtgtt aagtaatata	24480
tatagttata tatttttgat taaaaaagat ttttaggta tatagaagag agatttatgg	24540
ttgagaaaaa ggtagattgt ttttaagaaa ggttttaaat atattttttt ttgatagggt	24600
ataggagagt tagttagatg aattattata gtgtgtataa aatgtttagt agaataattt	24660
tttagtattg gttaggaaaga gtaataggag taatataaga gaaggaaagt tggtttgtag	24720
agtttgtttt agtattcgtg ggggtgatga ggttgtttg gttagaatt tttttatag	24780
ttttttggg ataattttt tagaattatt ttgtgtttt ttaggttatt taggtaatta	24840
ttattaatta gtaaatatag ataaaaaga atattttgtt tggatttgtt tgttaaaaaat	24900
taaatgatta taatttataa ttttaggaag atttgttatt aaatgaattt atgataaatg	24960
atttaaatga tttttataa gtagagtttg ttgtaatttt tagtatgaga agtatgaatt	25020
ttattattat tttttaaata ttttgagta aatggatcgt tttgtaaagt ttatgtttta	25080
tttttataga tagtttaaaa atagtttagt atttagtttg taggttggtt ttatttaagg	25140
agtttgaagt taggtaaggt gtttattatt atttatttag atttgttttt tggatatattt	25200
agttgttttt tttttttacg gttaggagg agagttatgt gtatagttta agaattgtga	25260
tttttatggt ttggaatag tggttataga attggagta ggtgtttttt gtgttattgg	25320
ttttttttt tggtagttt ttcgggtggt tggagttagt tagtgtttga tttagttttt	25380
tagatttagt aatgaaattt ggagggtaga gtaggtaatt aggtatataa gaattgtaa	25440
gtaatgttta tttgtagatg tgattataat tataattttt taataatgta ataggaatat	25500
aaattgttag attttttata ttattgtaat ataaagaaga taattgaagg tctgaagat	25560
agtttatgaa atgtaaggaa gaattggata attaagtaat aggaagagta gagaaatag	25620
gagatttggt gagttataag atttgggtgt ttaaattttg ttaggatttt tttgttattt	25680
ttatttttgt tttttttttt cggttagttt tttttgttt ttttttatag attagttttt	25740
tttttatag atagaaaacg tttttaagtt ttatattttt ggggttttgt tattgaagga	25800
ggattagttt gatatttttt tagatttata ttttaaggaa gatgaattat tagatttaag	25860
taatttcgtg taattggaga agattttag atttttttt atgggttaag ttggatatgt	25920
atgtgtgga gggagatatg ttttggtggt ataatatata tgtaagtaaa ttttaatatt	25980
gtataagata cgtttttgaa aatttatttg tagtttgaat atttggaagt taaatagtag	26040
atattgtgtt gtgagtgta gaatttagtg taatgttatt ttttgtagt taggtaattt	26100
tgtagtagga gaaagatttt tgaagtttt atttagggtt gaaaggttt gtgttagagg	26160
ttatgagtg gatttttttg gtattttgtt tttttttta aggtttgtat atggattgga	26220
ataaaagttt tggtagttgt aatttttaaa tttttttttt aggtattttt gggatttaaa	26280
gtagtaggga atgtattttt ttattttttt ttgttttttt ttttattgta gtaagaggta	26340

-continued

agtatagttt ttagttttgt ttgtatnttt ataggagatt tatgtatagg tttcggggtt	26400
tggttgtagg atagtgtgag gttgttgata gggcgagaga atggcgagaa taatggaata	26460
ggattgaggg agaaggaaga ggaatcgttt ttagaattat tatttgggac ggtgggttta	26520
gggagtaaaa aatatttttg gagttttttt aagggtttt ttggtagagc ggtattatcg	26580
ttagatattg ttatggttgt aggtaggatt gattaaagaa tttttgattt tttttgagga	26640
tagggtaaag ttttgaattg tttttagaga ttgtatntag atttatattt cgaattatt	26700
atnttggttg tatttgaac ggtaacgttt agggagtagg attggaggga aggagattat	26760
tacgtggttg ttgttttttg ttaagtaggg aataatgatg gcggaacgtt tgtattggtg	26820
gagaggagag agaggaggga tttggcgaat ttttgagggt attggtagga tttggaata	26880
tattgaatgt ggagagtaaa tgggtaggga gttagtata aagtttaggt gtttggttt	26940
gatgattggt tatgtttgtg tccgttgtgt atagaattta gaggtttagc gtttttatgt	27000
gtagcgtcgg tttttatttc ccggtagttt ggtgtggagt tgagatgtgt ttttttatt	27060
ttatgtttac gtttttttag ttagagtaaa gtatgtttgg ataatttagg ttttttttt	27120
ttttttttt ttttttttag agagagaggg ttttcgtttt gttatntaga agaagtgtag	27180
tgtcgtgatt atagtttatt gtagttttga atntttggtt ttaagtatt agaggattag	27240
gtgtatnttt ttataaattt tttaaatttt ttttagtatt ttatntttta agggatttta	27300
tagttttttt tatttggacg ccgttatggt gttttatttt atntttgaga attgnttta	27360
aatnttatat agaaatttgt tattatgtat ttaaattgaga agaaatttg aatnttgaat	27420
ttaaattttt taagataaga gttttagttt cgttttttagt agagttaaag aaaaagatat	27480
gatgaatggt ttatatattg aaaattgtag aagaaataat ttaatatgga atgggtgagg	27540
ttgttatagg tgggttaaaa ataattcggg tttttgttgt ttttttagt ggtaaaaatc	27600
gtaattgntt ttgtataaat ttaataatat tgtttgaaat ttatatnttt tggtggaata	27660
gntttgagat agagcgggaa ttagattatt gntttataaa atagtatgta gttattgnta	27720
ttttttaatt ttatatnttt attatgaggt aatgnttttt ttatttgatt tggagaatgt	27780
ttttatnttt tttttatata tttgtgaaa ttaaattgnta aaggaattaa tataataatt	27840
attgttaaat tatgatgtaa aataatattg tgtttgtaat tagtttaggt agntttattt	27900
ggnttaataa tataaaatta ttaggnttat aatatttttag ttttttagtg tgattgnta	27960
agaaaaaat ttacgttatt aaaaaagtt aaattaagta ataattaatt taaattntaa	28020
ggaaatataa agtaaagtat tatgttgtaa tatggaatat taatgtaaat aatntnttaa	28080
tagnttatga aatagatntt gaattggtag ttgataaaga aagnttagnt ttggnttaat	28140
ttttttatgt tttaatattt tttttgaagt tttatntaat tttatttgcg ttagtaatag	28200
tattttgaga tgggtggtgta tttggaaaat ttttttttt tttttgataa ttagtatntt	28260
taattntaatt tttatatatg ttttttgaa atattatatt tatattttga tgacgtntta	28320
aaaataattt taatgtaata attatatatg tgtggtgagt aaattgtgnt tcgtaaaaga	28380
tattnttaaa ttttaattgt tgtatntggg aatgtgattt tttttgaaa taggnttttt	28440
gtaggntatg ttaagttaaa atgagatntt attgnttag gatggntttt aatataatga	28500
ttgttgntt taagaggaaa atntgtatag agatatatat atatataaaa gaatgntatg	28560
taacgttagg tacggagaat attatgtgat aatgaaggta gagattggag tgatgnttt	28620

-continued

```

atgagttaag gatggttagt aatttttaga agttggggag agataaagga tgttttttta 28680
ggatttttag agtaagtata aatttgtagt tattttgatt taggattttt tttttttaga 28740
attgtgagaa tagatttttag ttttaagtta ttttagttgt ggtattttgt tataggagtt 28800
ttaggaaatt aatgtaatat gttatgttgt aaataaatag tagttttaat tggataaaat 28860
aataatttta agtatatttt attattaggt taggtgagta ggaaattggt tggtagatta 28920
attattaatg ttgtattttt tgaaaagtta aaatattttt atttagtaaa taatatttgt 28980
ttaaagttt ttatttgat tgttttaaaa taaagtaaaa attaaatgg agttagtatt 29040
gtatttgaaa aatgtaaaa aggaagaaa aatttttttag gaaagatttt ttttaaatg 29100
gttatatatt tgagaagaa ttttaaaatt taaattattt ggtgtttat tttatttttt 29160
tgtttatatt attagatata ttatgtatga ttttatatat tatgtatggt atatttaatt 29220
aatatttatt aaatatttaa aatttgtagt gttttgagtt aggtattgta gataggtaaa 29280
taagataaat attttttttt tttaaagaga ttattttggt atgaaattta attttagaa 29340
gaagtgaag aggtattttg tagagtttat tatgaaaatt tatttttatt atttgttga 29400
gttttatttg attttgtaga attattttag tattattttg ttttattttg tgtttttgta 29460
ttatattgga gaaaataaaa tttatttatt gtatagataa tataaaatta attaaagtat 29520
gtgaaattaa aaatgtaatt gtttattaaa ttttttttag tgtttttat ttattagtat 29580
ttaaataat ttttagttgt gtatttatgt tagtataatt aaaatgtggt gtttttattt 29640
ttttatttaa tttataatta ataatgttta tttttttatt ttttttttt gagatagagt 29700
tttattttgt tatatttagg ttggagtgta gtggtagat tttggtagat tgaattttc 29760
gttttttggg ttttaagcgt tttttgttt tagttttttg agtaattggg attatagcgg 29820
ttcgttatta tgttcggtta attttttga ttttagtaga gatagggttt tattatggtg 29880
gttaggttgg tttcgaattt ttgattttag gtgattcgtt tatttcggtt ttttaagtg 29940
ttgggattat aggtatgaat tattacgttt ggtttttat tatttttata att 29993

```

<210> SEQ ID NO 3

<211> LENGTH: 29993

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

<400> SEQUENCE: 3

```

aattataaaa atgatgaggg gttaggcgtg gtggtttatg tttgtaattt tagtattttg 60
ggaggctcag gtggcggat tatttgaggt taggagttcg agattagttt gattaatatg 120
gtgaaatttt gtttttatta aaatataaaa aattagtcgg gtatggtggc gggcgtttgt 180
aattttaatt atttagaggg ttgaggtagg agaatcgttt gaatttagga ggcggaggtt 240
gtagtgtggt aagattatgt tattgtattt tagtttgggt gtgatagagt gagattttgt 300
tttaaaaaaa aaaagtggaa aagtgaatat tgttgattgt gaattagatg aaaaaataag 360
aatattatat tttagttagt ttaaatgga tgtataatta agaatgtatt tggatattga 420
tgaatgggaa atattaaaaa aaatttgata aataattata tttttggttt tatatatttt 480
aattaatttt atgttgtttg tgtaataagt aaattttatt ttttttaata taatatagaa 540
gtataaagta aaattaaatg atgttaaaat ggtttttatg gattaggtaa aattatagta 600

```

-continued

gatagtgaaa atgaatTTTT atggtagatt ttgtaaatgg tttttttat ttttttata	660
gattaaatTT tatgttaaaa taatTTTTTT aggggaaggg aatgTTTgtt ttatttattt	720
atttataata tttagtTTaa ggTTtagtaa atTTtaggta tttataaaat attgattgaa	780
tgtaatatat ataatgtatg aaattatata taatgtattt gataatgtaa ataaaaaaat	840
ggaataaata gttagatgat ttaagTTTT aaatTTTTTT ttaaatgtat agttatttta	900
ggaaaaatTT tttttaaaag atTTTTTTTT tttttttat atTTTTtagg tatataattg	960
atTTtatttt aatTTTTgtt ttgTTTTagg ataatataga tagaaatTTT tagataagta	1020
ttatttattg aatgggaata ttttaatTTT ttagaaaata taatattgat aattagtTTa	1080
ttaggtaatt ttttatttat ttgatttaat aataaaatat atTTaaaatt attatttata	1140
ttaattagaa ttattgtTTa tttataatat agtatgtTgt attagtTTTT taggattTTT	1200
ataataaagt attataaatt gggTggTTa gaattgaaat ttatttttat agTTTTggag	1260
gaaagaagt ttaaatTAag gtgtTgTtag atTTatgttt gTTTTggaga tTTtagggaa	1320
gtattTTTTg tttttttta gTTTTgggg gtgtTggtt atTTTTggtt tatagatgta	1380
ttattttaat ttttgtttt attgttatat gatgttttTc gtgtTTaacg ttatatggta	1440
ttttttTgt Tgtgtgatg tttttgtata aatTTTTTTT ttaagtataa tagttattgt	1500
attagattt atTTtaattt agtaaaatTT tattttaatt tgattatatt tgtaaagatt	1560
ttgTTTTaa atgaggTTat atTTTaaat atagatatta ggattTaaa gtattTTTa	1620
cggaatataa tttatttatt atatatgtat aattattata ttgaaattat tTTtaagacg	1680
ttattaaagt atgaatgtaa ttttttaga aagtatatgt aagaattaaa ttaaagatat	1740
tgattgttag gaaagaagt ataattTTTT agatgtatta ttattttaga gtattgtTgt	1800
taacgtaagt gaaattaaat aagattttag gaaagatatt ggagtataag agaattaagt	1860
taaaattagt tttttttgt taattgttag tttagaattt atTTtatgaa ttgtTgggga	1920
gtatttata ttaagtTTTT atgttataat ataataTTTT gttttatatt tttttgaaat	1980
ttagattaat TgtattTgg tttgattttt tatggtaacg tgaatTTTT tttTggtTaa	2040
ttatattaag gtattaaaa attatgagtt tagtgatttt atattattga attaagtaaa	2100
attatttagg ttaattgtag atatagtatt atTTtatatt ataatttagt aatggtTgtt	2160
atgttaattt ttttaattt tagtttttat aaatatatga gaaaagatg agaattttt	2220
ttaaatTAag taagagaagt attattttat aataaagata taagattaag gaataatagt	2280
aattgtatgt Tgtttatga gTtaataatt tagttttcgt tttattTaa gattgtTTTa	2340
ttagaaaata tgaatTTTaa atagtattat taggtTgtg taaaaataat tgcggtTTTT	2400
gttattgaaa gtaatgTaa aaatcgtaat tttttTgta ttattTgtag taattttatt	2460
tattttatat tagattattt ttttttagt ttttaatata tagattattt attatattt	2520
ttttttaat tttgttagaa cgcagttata aatTTTTgtt ttagaaaatt agaattataa	2580
atTTaatatt ttttttatt taaatatata atagtaagtt tttatataaa gttTtaagat	2640
aatTTTtaga aatggaatga ggtattatg cgtcgttttag gtgggagaga ttgtaaggtt	2700
ttTtaaggaa tgagatgtTg aaggaggTtt gagaagTttg tagagaagta tatttaattt	2760
tttgattatt tgaggTtagg agTtaaggT tgtagtgagt tatgattacg atattgtatt	2820
tttttaggt gatagacgG agattTTTT tttttagaga aaagaaGaa agaagaaag	2880

-continued

aaatntaggt tgttaggta tgtttgttt tggtaaaga aacgtgaata tggagtggag	2940
ggggtatatt ttaattttat attaggttgt cgcggaatga ggatcgacgt tgtatatgga	3000
aacgttaaat ttttgattt tgtatataat cggtataggt atagttagtt attaaagtta	3060
aatatttagg ttttgttatt gattttttgt ttatttgttt tttatattta atatattatt	3120
aagttttatt aatattttta aaagtctggt aaattttttt tttttttttt ttttgttagt	3180
atagcgtttt cgttattatt gttttttgtt tgggtaggag taatagttac gtaatggttt	3240
tttttttttt agttttgttt tttggacggt atcgtttata gtgtaattag agtgatgggt	3300
tcgaaatata gatttgaatg tagtttttaa aaataattta gagttttatt ttgtttttag	3360
aagaaattag aaatttttta attagtttta tttgtagtta tggtagtggt tagcggtagt	3420
atcgttttgt taaaaggttt tttgaagggg ttttaggggt gttttttgtt ttttgggatt	3480
atcgttttaa gtagtagttt tggggacgat tttttttttt ttttttttag ttttgtttta	3540
ttgttttctg tattttttct ttttgttagt agttttatat tgtttttag ttagagtctg	3600
ggattttgat atgagttttt tgtgagggta taagtagggg tgagggttat atttgttttt	3660
tgtttagtg taaaagaaaa taagggaaa tggggagta tattttttat tgttttaaat	3720
ttttagatat ttttgggaaa agatttaaaa attgtagttg ttaaaatttt tgttttagtt	3780
tatgtgtaga ttttgggatg gggaaatagg tgtttgaaga gttttattta tgatttttga	3840
tatagatttt ttttaatttg ggtggggttt ttaagaattt ttttgtttat gtagagtgt	3900
ttgggttata gaaagtgtta ttatattaaa ttttaatttt tataatagtt ggtttattgt	3960
ttggttttta gatttttaaa ttgtaagtaa attttttagaa cgcgtattttg tataatatta	4020
aggtttgttt gtagtggtat tatatttata aaatatattt tttttttata gatgtatatt	4080
tagtttaatt tattaggtag ggtttgtaag ttttttttag ttgtacgaaa ttatttggat	4140
ttaatgattt atttttttta ggatagagt ttgggaaaat attaggttag ttttttttta	4200
gtggtaaaat ttttaggata taaaatttgg aaacgtttta tatttatgtg gaaaaagtt	4260
ggttttagg gaaagataga agagaattga tcgagaaaga ggagtaaaaa tgaagatggt	4320
aagaaagttt tggtagaatt taagtattta gattttgtga tttttaaagt tttattgttt	4380
ttttgttttt tttgttgttt agttgtttta tttttttttg tattttatga gttattttta	4440
cggtttttaa ttattttttt tatattgtaa tggtaggaga agtttaatag tttatatttt	4500
tgttatatta ttgagaagt atagtgtga ttaattgttt agatgggtat tgttttataa	4560
tttttgtgta ttagttattt tattttgttt tttaaatttt attattgggt ttgtgaagtt	4620
ggattaggtta ttggttgggt ttaattattc gggaagtatg tttagagga aagttaatgg	4680
tataagaat atttagtttt aattttatgg ttaattgttt agagatatgg agatgtatat	4740
ttttggatta tgtatatgat tttttttttt ggtcgtagaa aagaaaaata gttgaggtgt	4800
ttaagaaata gatttggatg aatagtaatg agtattttgt ttgattttta gtttttttaa	4860
tggagttaat ttataaatta gatgttagt tatttttgag ttgtttataa agatagaata	4920
tggattttgt aaaacggttt atttaattta aaatgtttta aaaatgataa tggggtttat	4980
attttttatg ttaaaagtta taataagttt tattttgtga gaattatttg ggttatttgt	5040
tataagttta tttgatggta agtttttttg gaattataaa ttatgattat ttgatttttg	5100
ataagtaggt ttaaaatagg tgttttattt tgtttgtgtt tgttaattgg tgatggttgt	5160

-continued

ttagatagtt tagaaaatat agaagtaatt ttgagtgaat tgttttaggg agatatgtga	5220
aagaatTTTT agttaaagta attttattat ttttacggat attaaggtag attttgtaga	5280
ttaatTTTT tttttttgtg ttatttttat tgtttttttt tattagtatt gagaaattat	5340
tttgttgagt attttgata tattgtaatg gtttatttag ttggtttttt tgtagtttgt	5400
tagaagagaa tatatttgaa gtttttttta gaagtaattt gttttttttt tagttataaa	5460
tttttttttt gtgtaatttg gaaatTTTT ttgattaaga gtatatgatt atatatttg	5520
tttaatatat atttgatgat tttatgatta ttttgaggaa aatgtgtgtg attttgagag	5580
aatttaagaa tagatTTTT aagtttagta ttgaatTTTT gaaatTTtag aaatgaaac	5640
gatagttttg ttaaagggaa tgtaagttta gtgattttta ataaagattg gaataaaata	5700
gattagtttg tgaaaatttg tttttattat tggataaatg tatgattttg gaaaagttaa	5760
ttaatatttt taaatattat tgtttttatt tataaaatag ggtttttaat aatatttgta	5820
ttatattttg ttatgattaa ataaaaata ttttttaaaa atatagatta tttgtttgt	5880
tagtgagagt ggaagaaaat ttgtttttat attttgtttt ttatatattg aattaaata	5940
tttaaatatt tttgttgta tagtagagta gtatttgtaa tttaaataaa tgtaattatt	6000
tttaattttg ttaattaag agtattatat taattattga ttttaggag ttatattaat	6060
taattatatt tagatagtag ttaaaagtta ttagtttttt attaattttt tgaatttagg	6120
taagtttgta aaaatgaata atttttcgat taagttatta gtatttgta ttttttagtt	6180
gttattttgt ataatttaa agttgtgatt ggtgttttta gtgttgagag attgatataa	6240
ttaataaata attatagacg gatatgtgaa agaaaagga ataggaaagt attagtattt	6300
agtattaata ggatgttaag atattttaag ttggtaaaag taggagaata agtataaaat	6360
ataattaaga taattggtta ttataatatt ttgataattt taaaagtttt aaaatTTta	6420
aaataagaa tggatatatg gatattattg taataatata tttatttttag attttgata	6480
aataagatga ataagaagtg gagttaagtg gtaatttaaa ggtgtgtgtg acgttttaag	6540
tgagaagtta agagaattta gattgatatg ggggtgaaaa aggtttggag aagtggatat	6600
gagggataat atgagagata tattaataag ataaaggaat tatatgaata taattgataa	6660
ggtaaatgtg attttttagt attatatttt aatttatgag aaattgatga tatggaaaag	6720
ttaagaggaa aaggtacggt tatatttggt ttgttatcgt tttgttgta ttattttttt	6780
ttatggcgat gggtaggaag taaaagtaag aggtaattta tagaggaag agtttaaaaa	6840
gttgttata tttattggag ggggtaga aaaattcgggt ttattattag atttagtaaa	6900
gaagtttggt ttatgatgga gtttttggga agagaataaa agaaaataat tttttgagtg	6960
agataaatta aaatTTtagt tttgtaggat gtaaatttta aatgtttatt gtatttggtg	7020
tatggaaatt ttaagataaa aaaataaaat aaaataaaga ttgattttaga ggtttgaaat	7080
ttttagattt taggggaagt taaaattggt ttatagagta taaagggttt taagagagaa	7140
atTTtaaaa agttataaat tatagcgag gtgcggtggt ttatgtttgt aatTTtagta	7200
tattgggag tttaggtag tagattattt gaggttagga gtttgaggtt agtttggtta	7260
atagcgcaa attttttttt tattaaaaa ataaaagtta gttgggtatg gtgggtggta	7320
tttgtgattt tagttattcg ggaggttgag gtaggagaat cgtttgaatt taggaggtag	7380
agattgtagt aagttaagat cgtgttattg ttttttagtt tgggtaatag agtataaatt	7440

-continued

tattttaaaa agaaaaaaaa aattataagt tatataaaga aataaaatat agtagagagt	7500
agttataata ataaattagg aatttaagaa tgtgaaatta tagagtaatg tggaaaaaat	7560
tagataggta tattttaaaat agatataaaa gaaatagaaa taaaattgaa aagttatata	7620
cgaaaaaaaa atatatataa atttggaaga gaattataaa aatgttttta aaaatatatt	7680
tattgaattt ttttataaga aataaattaa atagtaaatt agatataattt aaagaaataa	7740
tttatgaatt ataaggtaaa tttggtaaat tttagaaggt gtgtgtttat ttttataaaa	7800
gaaaaaatg ataataaaaa taataataaa attataatta tattttttta aataattttt	7860
ttataatatt aattttttag gaattaaggt gtgatatttg tgagtatat tttttattat	7920
ggggaggag aagtttttaa tataataaag taaatatgag agttataaa tatggtaata	7980
aatttgattt atttaagtgt aagttattga tttgaattgt gtttatagtg aatattaatt	8040
aagaatgtaa ttaaataaga tgtgtttgt tttttttaa tttgggatat ttgaggaaat	8100
tttattgatt tagattttta attgaggtat ttaggaaagt ttgagtattt ttttttttt	8160
taataaattt tttttagta tttgaatata aaaattattt tattgatttt gtaattgaat	8220
tagaagtagg gttattttta ttaaaaattg tttttatagt ttttaaaaa ttaagattt	8280
ttaaaaggt tgaagtgtga agaatttaga aaataatgta ttaaatatat ttggtgttt	8340
gagttaattt aacgttatgt ttttttgta tgtgtttgt ttttatatgt atggtaggtt	8400
gttagatatt aggagtatgt tgttattttt gttattttta attttttttt ttttaagtaa	8460
gtacgggtgt tttggttaat aagattattt ttttaattta aaaaaaatt tttgttttt	8520
aataggaat tataaggaa tttagttagt tataaaagta taattgttta ggaatgtat	8580
taagtgaag gtagatataa attaggttat agtgtattat tattttttaa ttatttaatt	8640
ataaattgac gtattaatat agattaatat taatgtatat taaagtatta tatttatatt	8700
taaataaaaa tgaaggaag aattaattag ttgaaataga aattaatgat gataatttt	8760
aaagtagtaa gaatgtattt ttgagaaga ttttagggtt ttgtttgttt gtttgttga	8820
gattaaatta ttttttagta aaattgaagg gggaaaaaag attaacgttt aagttatag	8880
aagatgtgtt ttgtttgta tttggatgtt tttttagaa tttataagtt tattttttag	8940
gggtttaaat gtttttgat ttaaaaagtt ttcgggaaag aattttttta agggaagaga	9000
tggagagta tataagtaga aattaagttt gagttttaag ttttgggtt gtaagaaaag	9060
aataagagaa taagttaaa agttgttta ttaaattttt tatagaaggg gaggaggta	9120
ggtgttatta ttacgtattt tgaagttttt gttatagggt attggattga ggaagttgt	9180
tagatatttt ttttatttta aattatttgt aatttttagta gttcgattaa aaaatataat	9240
ttcgtaaagt gatgaaaga aacgaatttg gtgaagaaag cgcgggaaaa tattggttaa	9300
attagtaaga aaaagatcgt aataaaaaag aatttaaacg gtgagaattg aaattagga	9360
gatgtagaa tgagaagtag aataatttgt tttttttttt ttgattgaat taaatgtaa	9420
gatattaggt gtagagatg tggatattag tttttatagt attttttaag agtgttatgg	9480
ttttttttcg tttttatttt atgttttagg tttttttttt agggtttttag ttaagtttt	9540
gtttaagtat ttgaataaat tagatttatt tttgaattag ttttaggtgt tgtttataga	9600
tattgttatg gtttatattt ggataattta gtttggtttt ttttaagtgt taatttttt	9660
tttatgatcg tgaaaaaaat gtaatgagtt tttttgaaga aatatttaag atatataaat	9720

-continued

ttagaatatt tatagttttt aaattaagat tgtattatat aaaaatcgcg atttttttaa	9780
aatgtatgaa ttagtgaatg agtgataaaa tagttaatag tgtataaatg tatgtatgaa	9840
tgaatgaatg agagtataaa taagtaaata aggaaatata tttttaaagt atagtttagg	9900
aaatggataa attacgtttt atgaagtttt agattgaatt aggtaaatag aattgtagat	9960
ttttatgttt gttatgattt agtgatagat tttgggaaaa ttttgtaatt tttttatatt	10020
ttagtttttt attttataaa aggagtatat tgtttgaaat ttaatttttt tatagtagtg	10080
agagttttga ttttaacgttt ataaaacgtt ttttagtttt tttgattaat aggtaaaata	10140
ttgtgtagt tttgtataata ttttttttaa aaaggataaa ggtaattggg cgaatttgat	10200
gtaaaataat atttttatata agaaataaat attagttata tttattattt tttttttta	10260
tttgtttagt aataaaaatt atttttaata gaaaagtatt aaattaatat ttaagttagt	10320
tgaagtgtta ttattatagt tttttttac gttgggggaa ggagagttgt tttgtttttt	10380
taggttattt tatatgttta tatggtaaaag ttaggaatgt tgggtatagt tttgaagggt	10440
tttagcgtat ttattggttt ttgttatttt aatgtatttt atttttataa ttataaaagt	10500
atttttgagt gtttttgtaa atataataa aatttttaat taatttaaga gtgttttaat	10560
tgtatataatg aaattttttg ggaggaaaaat attagttatt ttatatatat atagtattta	10620
agtttagtaag agagtttttt ttgaaaggta taaagtagtt tttttatagt aatataattt	10680
ttagttttat atgtgtaagt tttattgaat tttttagaat aataggaaat tttgtaggaa	10740
aaaaaatttt ataaaaattg atagttatat ttatttaatg taagaggaat atattagttta	10800
ttttgagttg gattattttt taaaggtatt agataattat tgaagattaa ttaaaaatttt	10860
aattaaaaat tataagaatt gaatattttg gagatagaat tttttatata taaaaatcga	10920
ataataaaat taattgatta atataatttta ttttaaaatt gtatttatta gattatttta	10980
gttagttatt gagtagtaat tttttggatt ttgatgtatt aatttttatt tttatattaa	11040
atagattata ataagtaata gatttataaa ttagttattg atagatata attgatattg	11100
tgatatgtat aataatattt attaatatat ttataatgga gtaatgttat ttattattta	11160
tataatgttt atatttttta tttttatggt aaaaagttat cgtaatatat gtaagggttt	11220
tagaattgtg tttggtatag agtaggtttt tattagaata ataatatatt attttttata	11280
agtttttatt aattatatga gagttagaat tttttttttt aataattttg gtaaatattt	11340
gatggatta agtaaaaaaa aatgtagtga atataatatt attgatttat ttttttttaa	11400
tatttgggtt tacgaattat aagtaaatat aatattgatt tataaaaatg aatattaaat	11460
atttttaagt gagaaattaa ttatataaat tatatgtata tattaaatat atacgttaa	11520
tagaattatt ttttttaatt atagtgtagt atttcgatta ttatatgagt ttttaaggat	11580
ttattattat aattattttt ttttaattaga gtataaataa ttataagggt taattaagtt	11640
gaaatttttg tttttggttt ttgtgagaaa gatgagaata ttatttttaa gatgtaagat	11700
ttataaattt ttttaaaaag ttttttggtt ttttttatgg aaaatgtata tatatattta	11760
tttattaata ttgtatttag agtttttatt atagatattt ttggtagatg tacgaagaga	11820
tataaaaatt attagtgtaa tatatttagt aatagtaaaa ataggaagt aatgtaata	11880
tatattaaga ggggaatagt gaaattatga aatatttgta tatattaaat agaagtattt	11940
aaaattatat gtattgatgt ggaaaaatat ttaagattta gtatgataaa agtaagtttt	12000

-continued

aaggaaatta atgattttat tggagttaaa aatgtaaaat gaaaataaat tatattttta	12060
tatagttatt atatttata atatatagaa aaaaaaatgt ttgaaaagat gtatatttaa	12120
gtataataat tttttttgat ttgtagagaa taagaatgag taaatttatt ttttatttta	12180
aatatattcg tttttttaa attttttaag aaagtataat atatgtatag taaaataaaa	12240
tatataaatt aaaagttgat tattgatgaa tttttattga ataaatattt ttatatgatt	12300
agtaattaaa ttaagataaa agatataaatt agtattttta agtttttttt ttgttttttt	12360
tggttatttt ttttaaaagt aatttttatt atttttgttt taatatgta tataaagat	12420
gatttatttg ttttttttaa gtaatagatt aaaaaaaaa atttgtttta gtagtatgaa	12480
tttttttttt tgaatattaa atatatattag ttttttatt taagatatta atttaatta	12540
taatttgagt ttttatagtg ggttaggtat tgaattaaat gttttttttg tatgaattta	12600
attatttttt ataattttat gaggcgagta ttgttattat tttttatgta tatatgttg	12660
aattatggtt tatagttatt tagaaaatta tatgattaaa gttatattcg tgatgagtgg	12720
tagagtttgg atatatataa aggtagtttg attttagtat ttttaatttt ttattagaaa	12780
taaaataata aaaaataagt ggtaagatgt agagttggga gtatgaagtg ttttttagtat	12840
gaagtttttt tttttgttt tgttttgagt aaatgattta aggttttaag gttatataat	12900
ataatatttt atttgtaaa tgggaagggtt aggttagatt ttttttttga tttgattatt	12960
tgacgttgat aaaaagtg atttattgaa atatatagta aatatgtttt gttatgatgt	13020
agtgaatta tgtattgtat tatgttaaaa aatgatttat ataagaaat aaaattataa	13080
ttttgaaat tattggaat aataatttta atttaatttt atttttggat ataggaagat	13140
tttttttttt aaataagtaa attggtgggt tttgaaggt tttgtttatt tttgtatgga	13200
aaatgtttta aaaaataa ttaaaaatta aataaaattt ttttattttt ttaattgtgt	13260
ttgaaaaaa agatgtaatt ttttttagta aaaggttacg tatattttta atagatata	13320
ttgggatttt tggttttat attttgaat attgatataa ataaatatag aatatttttt	13380
tttttaaatg aaattataa ttttgaggt taaaaatttt ttaaaataat atgaaatcga	13440
tttttagtag aggtaattaa agagattaag gtttgttttt ggagaaaatt agaattatga	13500
agattataaa ttttaacgag aaatttgttt tttaaatggt agtttattga aaatattaat	13560
gtacgtcgg atattttgtt tttttattag aaaattagta atattttggt ataaatgagg	13620
ttttaaataa gttggagttag aaaataatta ataaattagt ttttatgatt ataagtaaaa	13680
gagaattagt ttttaaaaa gttaatagga tttatttttt gtagatgttt ggtttgggat	13740
ttaattgttt ttaaagatat gtatggaata gtgtagttat ttaattttta ttaatttttt	13800
taggtgttga ggttacggag ggaaaaagat atgatttttt ttttttgtga gtttatagtt	13860
tttttagaag atagatata ataatgagtt ttaatataat ttgttgttat attaatagtt	13920
taaaaaagt gttatggtat tatatagaga aaagaattaa ttttgtttga gttttaagga	13980
aggttttata gaggaggtga tttttgagtt aagttttgaa agatgagttag attattatta	14040
ttattattat tttgttttag taggtagtaa aagagtatta ggtaagtagg ttaagggagt	14100
tgggttatg gaggataaag gtggtgagaa tatgttaggg agaaaaagg aagtattga	14160
cggtattata agaaataaaa tgtttataag taagttttat tatattagga tttattatat	14220
aaagaattaa tttgagaaag gagatattgt atttgaatat tttgtggggt tttttgattt	14280

-continued

atgaatttag gaagtgttta tttggaaaa tagaattttt gggtttttgg tctgttttta	14340
gtaattaaag ggttattagg tttttcgttt aatatttggg ataaagtgtt tggatggatt	14400
tgtggttcgt tgtgtttttt ttggattttt tgtttatgaa tgaagtagaa ttttttttta	14460
ttatttga tttaaatttt tgttttaaat ttttcgtaat atttttcag aacggttttt	14520
atttgttatt gggttgttta gtatttattt ttttttttg ttaataagga gttatgtagg	14580
gtttttttt atttaggaaa tagaaaaata ttttggttaa atatagtatg ttgtatttgt	14640
agtttttgt ttttaagta tatattttaa aattaaagta taaattaaag gaagtgagag	14700
agagagagtg agtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtt	14760
tttaagtatt tttattttgt ttatttaggt attaattagt aaatgatgtg aaatttttag	14820
gtagtttga ttgtttttat gattatttat ttaataaat taaagtataa ttattaaatt	14880
agataatgta taagtgtgga attttaagta tttttgtaat ttttaataag ataaaaatta	14940
tttattgttt agaattttta agaaaattag ttttaattata tatttataat attaatgttt	15000
tataatttag aataattttg aaatttaatt taaaatttat atgggaaaag ttattttttt	15060
ttattttatg tgaatattat tgtttttatt ttgtttaagt gatttttgaa atgtgtttta	15120
tatatgttta aattatgtaa ggggtgtgtt taaataatag ttaaatttat ttatattaat	15180
tatatattat tttatagata aaatgtattg gttaatatta aatttataga atgagaattt	15240
aaaagtttg atttgagaat ttataaaaga gagtttttat atgtaataa tttattttta	15300
ttttttttt gttgtataaa tgttattaat ttaagatttt gaatttggg ttataattga	15360
aagttatata taattttttt gaaaattttt tattataaat aaataattta atttattaag	15420
tattttaatt ttttagtagt gtgattatta gtaattttat tttttttaat attatagatt	15480
aaataaaata aaagtttttt attattgaat atttatttat agaaatggaa aggtaatggg	15540
tattatttaa aataaatttt aaattattta ttgataaata atcggagaagaa aagatatata	15600
gatgatttaa ataaatattt attttttata ttttttttt taggtatttt gttattttta	15660
tttatatatt attttttagaa aaggaaaata tgttgagat ggtaatttg tgtgaagcgt	15720
atttgtaaat tttttttgtg taaagttaat ttattaatat agattaatga gaattgttg	15780
atttatgtt taaatttggg aggaaggtat taggattttt tatatttcat tggaaataga	15840
ataattttt tatgtagagt tttgggagtt ttttgattat attaatgat attaagattt	15900
gtttaggtta tgggttttta agaagaggta attttgttt tagagattat ttgataatgt	15960
ttggagatat tattattatt aggggagga gggtttttgg tatttagcga tgttattaag	16020
tattttatta ggtataggat agtatttttt gaataaatta tttagtttaa aatgttaata	16080
attagagttt gaaaaatttt tatttaata tattgaagaa tttaggata acgttattat	16140
ttttaaatt atatagatat atttgtagt agtttagat aagttgtaa aatgagattt	16200
tgataattaa tattttaatt ttttttttt ttttagttt attttgttat atagttatta	16260
tatttgaatt taagtataa attttgaag tagaggtatt aattattttt tattgttttt	16320
ttaaaggtaa gaattatagt aatagttaa gttatttttt aaaaagtta tattgagttt	16380
attaggttta ttgattattg aaaattatat aaatttttat aattggattt ttacggttga	16440
aaaaaagta attgagaaa tttaaatatt ggtataaag tattattatt atataacgtg	16500
tttataattg tttaaagtag ttaaggaatt atatagatta aagtttttaa ataatttaga	16560

-continued

ttattgtatg cggaaatatt tattaaga aatattattg tataacgtaa atggttaatg	16620
ttaattgttg agtatttata gtgttttttt ttaaggtttt aaatataatt ttagatttgt	16680
taagatgatt tatagtattt aatttatatt tggattattg gatattttat ttatggttat	16740
taatttatag ttttttaaaa aagattttta gtttttaata tttttggaat atttgggggt	16800
aatttaatga aagatattta attatagtaa taataattga tatttatatt aagtattatt	16860
ttaaatatat ttatatatat ttatttattg aatttttatt agaattttat gaggatggt	16920
ttgttatata tgagaaattg aagtaaaatt tgttttagta atgtgtttga tgtataaat	16980
ttattaatgt tagagttgag gtttgaattt aagtgtttgg ttttataatt ttttttttt	17040
ttattttatt atttttttt gagatggagt tttattttgt tgttttaggt ggagtgtaat	17100
ggtgtaattt tagtttattg taatttttgt tttttgggtt taagtgattt ttttgtttta	17160
gtttttcag tagttgggat tatagggcgtt tattattacg tttagttaat ttttgtattt	17220
ttagtagaga cggggttttg ttattttggt taggttggtt tcgaattttt gatttttaggt	17280
gattttattc tttcggtttt ttaaagtgtt gggattataa gtgtgagtta tcgcgtttag	17340
tttttatatt ttttttattt atttattgtg ttataagaat tgagtttttt agattttaag	17400
tattatgaaa ttatcgtatt ttttagagatt ttttaggatt ttggtttatt tattagatta	17460
ttgatttttt agtggaaga aataaatata tttttatatt tacgtatata tatatataat	17520
ttaagtttgg atttatagtt gtaattttat tgggtttata ttaaattttt ttgaggttgt	17580
tattgaggta gttttatttt gtaatagtta ttgttttttt tattgttttt agttttaagg	17640
aggggatggt tagtgttata atttttgttt ttatttttag tatttttatt tatttattta	17700
atgggaatat aataattatt tattgattat atgttttagta gtgtattgtg tattagcgat	17760
atagaaatga taaatatagt taattttttt agtaattttt tagtttaag taggatttag	17820
attaatgtaa tttattatat tttaatgtgg taagggttat gttaaatgta gaagttggat	17880
gttttgggag taataagaaa gaaacgttta atttagattt tggaggttag aggaatttat	17940
tagaggagtt aatattttgt attttaagta ttgaaaaagt agtagaaaag atattttag	18000
ggggagaggg gcgatgtttt ttagagaaaa agtattttac gtataaagaa atgaagtgga	18060
aatgtttaag gtgtatttga gaattgtagt tttttatata gttggaaaag agagttgttg	18120
tgagtgggga gtgagagatg agaagagtaa tttgagtaag gtttttagta tgaagggtta	18180
tttttttgt agagaatttt ggaatttgtt ttataggtag taggtaaatt ttttaaggaa	18240
ttttaagttt aggagggatt ttataaaaat tatgaataaa tgtggttgaa tattttaggg	18300
acgtgagaga gtgttagaga tagtttttaat atagttttaga atgggttttt aggagaggta	18360
ttaagatatt ataattatat ttgttatgat aatattaata cgaataagtg tgtataggat	18420
aattgtttat ttaaagttaa tttgaatatt tgaggattag atttgtgatt ttggttataa	18480
tgattttttt tttaaatatt tttgggaatg gtatgtaaaa atataagaat ataataaag	18540
ttttaaagtt tttaaagtta tgtagagtag aaaagaaatt tatttttttt attttaattt	18600
tgtaagtttt tagtaatatt gataaggaaa aagtattttaa taagtattta ttaagtattt	18660
ataattttgt taattatttt tatgaacgaa atttttataa attaaagtta ataattaagg	18720
gaaaaatta gatttaaat aaaatttttt atgagttaat taagaaaaat aatagtatta	18780
ttattttatt aattaaagtt aatgtaaaaa taataaata aaaaaaaaa taaaataaaa	18840

-continued

tagtagaatt gtcgTTTTTg ataggtttat tgaggtttgt atgtttaagt tttttattag	18900
aagtattggt ttagattatt ggaaaatttg taaaaggaaa tttgaatgtg attttattat	18960
aacgttatta aatttagtaa gaatgatttg agttataatc gtatatgaat atatatttgt	19020
atttaaatgt aagagaaaag aattaggtaa thtagagatt tgttgatttt ttttaagtta	19080
tattattttt aatatattga aaatattaag ttagaaatat ttaatttttt aattagaata	19140
ttcggttttt tttaaatttt ttaatattag ggagttaaa gtttagaggtg aaatgtaatg	19200
aatagattat ttogtgggt atatgtagag gtttgtgaat tgtatatatg attttgtgta	19260
taatgagttt tttcggtatg tggatatatt atttgattag ggatagaatt ttaaatTTTg	19320
tattgttgtt tttgtgagta gaattaaatt taatgtata tgaatgtaa tttttatgat	19380
ttaattgtag tttttataa gaaatgtttt tttttttta aaaagaaagg gggaaaggat	19440
aatagaaata gtagtaataa tttactgaa aatttaaagt tatgttaagt taggttattg	19500
ttggtgtttt tttttatagt ttttaaagga gaaagatgt ttttaattaat aaagttaatt	19560
tttgtaggta taagttttt gtgttttatt tagatttgta attatatata gggaaatttt	19620
tgattagttt taattttttt agtgttaaat cgaatttgaa gattggtaa tacggatatt	19680
taagttatag aaggagagt agatatattt tttgtttgat atgttatatt gggagatatt	19740
ttatcgtatt ttagatgttg ttttttagtg atgttatatg aatatgttag ggtaggtatt	19800
ttgaattaat tttttttgtt agaagatttg aatttttttg tagtagtttt tatatatga	19860
agatatttag agagaatttt agaggcggcg gttatcgtga gaggtattg ttggtgagtt	19920
atttttaggt tatttatatt ttgttttatt tagagtagtg gtatatTTTt ttattgggta	19980
tattattagt ttttttttaa atgaaaggat tcgtggttga gtatataagt aaataaaaa	20040
atgaaaatta ttgttttagt ggaaaaatat tataaaattg taaaataggt tattaatagt	20100
aaaagtattt atttatagaa aattttatat atgcgataat tagataaatt gaatgtattt	20160
ttaaaatag gtagtTTTta ttataattta aatttgTTTT tagaagtTTa ttatgtattt	20220
tttaattttt taagttttat tagttggttt aatataaatg thtagatttg agatttaaaa	20280
ttagtattaa aataataaaa aataaaaaa aagaaaatat aaagttgata atatttaatt	20340
aatttgata aggtttgaag agtaaaggaa gattttatat gttagttaa tttttaaag	20400
atttggttat aagtttttga taattgggaa atattataaa atatagtatt tttttaaatt	20460
tagttatttg ttgttttagt tttatgattg tggttattta ggttatattt ttgtttattg	20520
ttatttaata cgttttttta gatatttaat tttttttttt tgattttaat attttgtttg	20580
gttttgtttt aaataatatt gtttatgaag ttataggttt aatataaaat ttaataaatt	20640
tttttaatat ttagttattt ttatataaaa tataagtatt attaaaataa aatgtttagg	20700
tattatttaa aaatgggtgt atatgattga tggaaagtaag agtttagtat acgtttaagg	20760
ataaggtgt tggttttata ttggggagtt aggatttttg ttttattgtg ttgtatgatt	20820
gggttttagt tttttattt gtaaaatagg aataataaaa gtatttaatt tacggtatag	20880
ttaagaagat taaaattagt taattttatat aaatgtttta aatagtgttt tgtacgaaat	20940
ttttgagtaa atgttagtta ttagtattat atgggtatag tgattaatat tatagtaata	21000
atgatagtta aattgtaatt atttaaagat tatagtattt attttattt aaaaatatat	21060
tttaggatat aaatatattt aagaaaatat gtttttatat ttgatattta tattaatata	21120

-continued

tgttttggtg atataaaagt tttttatata tggtttaacg taattttatt ttgaaaatat	21180
ttatgtgata ataggaatta gattttgttt aaaaattata atttaataa attatttgtg	21240
tagaaaaaa aaaattaaaa aattttaata cgatagaata tataggtttt tttttattg	21300
tttaaaagta cgttaaaata aattttatta ttgtatgtaa ataaaagttt taaaaatatg	21360
aatatattgt attaaatat tttatattgg aaaaaatga gaatttttgg tattttagaa	21420
tggtgttttt tttttatttt tatgtattga aattttaggg taagaaaagg gtttagattaa	21480
ttttgaaag ttttaaatat taaattatat tattgtgta gtattttaag attaattagt	21540
attttttaag gataaaaaag gttatagtga gatttagagg tgtatttttt gaaattgttt	21600
taaattatta tataaatgtg tttttgata tataaaatgg tttttatttt aaattaatta	21660
gttttatttt tttttttgga ggtgtgaaag gggttagggg agtaagaata gtagtgttta	21720
tttttagtat ttttttatta ataaaaagat ggatggagaa atattaagaa tgataaattt	21780
taagcgtaaa atgatataatt attaaagtat atattaatgt atagttttat aagtaatat	21840
gagtgatagt aatgtttata aagtttttaa attttttaaa agtaattaa gttgtttgta	21900
tttttttga atgtgtaatt tttgtttatt tttttttatt taagttaag ttaaatgacg	21960
ttaagtttg ttttaaaaat gatttaatga ttatgtagtg tttttttat taaaatata	22020
ttaattataa gatgtagtat tatatatatt gtattttatt ttttttagg gtaaatgttt	22080
gaaatagtat gttatttttt agtttagtag gaaaaatga tgagagaagt aggtatattg	22140
taatgttatt ttattttttt ggtgttgatt ttttttatt agtaattatt gtaataaaat	22200
tttttaattg tatttttgaa tataaattgt ttaagatatt ttataggag gtacgttata	22260
agagtttagt atttagtata tgtaagaata atatttattt agaaaaagaa attgttaaga	22320
ttatatattt attgaaaaaa gaagaagggtg ttttgataga ggtttgttta gtttttaat	22380
gagaggattt aaaacgtttt ttatttttga tagtatattt ttttatttat tttttaaata	22440
agtgatttga aatatttata attagatagt atttaatttt aggttttaag tgttttaata	22500
gttttatgaa atagattag tagttgtaat atagaattat ggtgaataaa tttttataa	22560
agaggggttt gtttttatag aatgtttagt gtttttgtat ttatttaata gaaaaaatg	22620
ataatgagaa agtataatat atatttgtat ttttaaatga ttttaatttt aattaagtag	22680
tgaaaaagat tattatttttt ttttttgaaa tttgtataaa aattgtgagt tttgtgggg	22740
taaaaaagt ttgattattt gaaagttaga atgtgaattt ttttaatttt taatttttta	22800
taataatagt taccgttttt ttgttatttt atttttttta aaattttaaa aaatttaatt	22860
ttagtgata aatgtttaa atgaaatatt aagttatata aatatatttt tagatacgat	22920
attagttatt tagagtttgt aatagagatt gaaaatatga aattggatat tttatttttc	22980
gtggttatgt aaaatttatt tatattttta agtataatat aaataaaaag gaaattttta	23040
gaaatggagt atttgttttt tttttttttt ttttaattta aagagtttta ttaaaagaaa	23100
aaaataaatt atagtaaaat gataagtaat ttatttggta gttattcgtt aaaaaagaaa	23160
tttttatgta attttttttt ttagttttta gaaatttttt gaaatttagt atttattttt	23220
tttaataatt aataattatt attttttttt atataattat ttttttttat ataaatttta	23280
agattttaaa aatttttaaa ttttaagttta ttatttaaaa gaagtgttaa ttgtgagtat	23340
atatatatat gtgagacggt ttatttttga aatgagaaat ttaattttta atatttaata	23400

-continued

aatattaata	aaaagaaaat	aattgagtaa	atattatatt	agttataatt	gtgaattaat	23460
cggtagttat	aatggttttt	aaatattgaa	atatattatt	tttaggaaat	ttataaagga	23520
aaataggttt	ttgaaaatta	tattgattgt	agatttttat	ttattatgaa	agataagtat	23580
atattattaa	taattagaga	aaggagatat	ttttaataaa	tttaattatg	ttgtttttaa	23640
agttgtagtt	gttaaaggtt	ttaatatttt	tgtatagtta	tttttaaata	cgggtttaat	23700
tttgaatgat	ttataatatt	taattagggt	gaataaatta	tttttacggt	ttataataat	23760
aaatgtgggt	gtagttatgt	ttgttttcga	tattgttgtt	ataaaaatta	gattttggta	23820
gttttgtttt	atataaata	atgtgttttt	tgttattatt	gtttttgtt	ttttgggata	23880
atgttatttt	taaaaaata	taaatgtttt	tgaatgtagt	tttttaaacc	gtttcgtatt	23940
acgtattaaa	tatttttttg	ggggattttt	tttttatttt	tttttagttt	tggtttttta	24000
ggtttaagtt	tgtaataaat	attttaaaag	ataaaaaaaa	aaaaaaaggt	attgattttg	24060
tttaaaatta	aatcgtattc	gggaaaggaa	tttaagtata	attttttttt	taaggtttta	24120
aaaataaaaa	taacgtaata	ataaaaaaaa	ttatcgtaa	tttttaattt	tgtagttat	24180
ttacgtgtaa	gtgaaatatt	agaagttgtt	gatggggagt	tcgcggttgt	taaacgtttt	24240
taagattttt	atttaaattt	cgtaggagtt	cgtattgaaa	ttgataaata	ttcgaggggt	24300
tataaagtgg	ttatagatga	gaagaggaat	aaaatgttat	attttaggag	gtgtatttcg	24360
ttttaataat	tacgttgatt	ttaggtgtt	aagttaaaag	gggtaaggga	gaaagaaaat	24420
tgtaagtat	gagtttttaa	tgtgaaattt	aatgttttg	tttttttagg	tattagattt	24480
tggaagtgtt	tgtaatgata	gttaggaaag	taaatatgt	gttaacggtt	tattgggttt	24540
gaatattaga	gtttgttat	ttttttatag	attaagtata	taatatatat	atataatata	24600
atataatata	acgaagtgtt	tttactgga	attgttttg	aattaaata	tacgttcggg	24660
cgttattttg	tattttggt	gaaagggtag	aggattttgt	attttttttt	ttgtttttta	24720
ggtaagatt	tatgaatgat	tttttaggtt	cgttcgggat	ggcggggaga	ggcgtaggag	24780
acgcggtgag	agaggatagt	ttttatatta	tttttcgta	tataatttta	tttaggcgt	24840
aaatcgttg	ttttatag	aaggagaaa	aggggtgtg	gtgaaggga	atattatatt	24900
tttaagtagg	gaagagatga	taaaaattaa	attgcgtttt	ttgtttcgt	tttagttttg	24960
ttatcggatt	cgttcgtaa	atttttta	cgttcgagt	tgtttagaat	ttcgtaaata	25020
aaggagaaa	atttttttaa	gtgtattaa	acgaaagttc	gggtgttgt	gcggcgtaga	25080
ataaaataag	tatatatacg	ttattttta	aatcgtttt	tattacgta	aaaattcgag	25140
ttattttaag	ttttattgtt	tttttagatt	tagtaggttt	ggaaggagag	aaaaattgt	25200
tttatttagt	taggtttttt	tttttgaga	agttataagt	ttttatggt	atcgtttatt	25260
caggttaatg	aacgcgattg	ttttgataag	gcgtttttta	ttttatcgat	tttggtaaat	25320
tgtaggagat	attagttta	cgttatataa	aaattattaa	agtggtatcg	tcgtagttga	25380
aaagttcgat	tagtgttttg	agattagaac	gcgagttttt	aaagtata	cgtattcgggt	25440
atcgggaaag	tttaggcgtt	tgtgattttt	cgaaggattt	tggttagta	ttttcgcgtt	25500
tgattatatt	aggggttat	tagaaagtaa	tatacgttat	aagaaagatg	tgggggagat	25560
ttcgatgttc	gcgtttttgt	taggtggagg	cgggtgtgag	gtagaagtcg	cgcgttggt	25620
gtttggatga	atcgtttac	gcgtaggttt	cgcgtcgatt	tcggaaggga	taggtaggtt	25680

-continued

cggaaggat ttttggggtt tggggacgtt ttttaggta gagagtatcg ttgctggttcg	25740
agcggttcgg gtttttaggc gtgggttatt cggtttttaa cgcgttttcg gggacgtagc	25800
cgtttacgta tctgttagtt gcgcgttttt agtatttagc gttcggcgtt ttcggcggag	25860
tttcgcgtat ttttcggaa agtttggggg ttttttaaat ttagggtttt tcgggagttt	25920
tggggcgggg gttggaagaa ggggttattt atttcgtcgt tcggttcgtt tcgggttttc	25980
gaagtgtttg ttttagtcgc ggttttagtc gttcgttaatt aagtatttat tttcgttttt	26040
tttagtgcg aagttttttt ttcgcgttgt gtttttggtg aaaataggat atttttttcg	26100
gttattttat aataaatagc gtatataagc gggggagaag cggggcggag ggagaagtcg	26160
gtttcagagg ggaggaggaa aggagaggag ttaaaatctt ggattgcgat aggggggaaa	26220
ggagaagaaa agaaaatgag agagcgcgtt tagtcgtcgt agtcgttatt tttcggtttt	26280
tttcgtagtt tttttttttt ttttaaggta gcgataatth tacgtggata ggatggaagt	26340
tttgccgga agtcgggaac ggtcggagcg tgcgtcgcgt tgcgttagt ttcgtatttg	26400
cgtagggagg tggggcgggg cgcagctcgt tctcgttat tgcgggttcg tgaggctcgc	26460
gggggttttcg ggggaggtcg ttagggacgc gggggcggc ggcgcgcggg ggatttattt	26520
cgtttttatt tgatttcggg tgattgttta ttgttatggt ttgcgatttc gcgtggggat	26580
cgttttaggt ttcgggtttt tggatgtttc ggggtgctc gtcgtaatth tctgtttttt	26640
cgctcgttt tttttcgtt tttattcggg tctattcagc cggttgtagt ttttaggttg	26700
aggggaggga ttcgcgatcg cgtttttttt tttgggtcgg agagtattg tctattttta	26760
ttgtttttg tgggtagtt ttcggaattg tttggaatth tttcgttatt tagttttttt	26820
gtgtaggta ttgcgggttt tttgtcggat ttttttttat ttattttaag ggaggagata	26880
gaagggattt cgtttatttg taataacgtg aaataaaaat taatttagat tagattgggg	26940
cgtttttttt gacgggagga aatatttttt ttcgtggaga tatatgttat gaaggattaa	27000
gcggtttggg gataggttgt gttcgaagg gtttgggttt ttttagtgt tgttgggtgt	27060
aggtataggt attcgtttta gacgtttaaa ggttaggtag taacgattta tttttaagc	27120
ggcgttagag ttttttagg tatattgttg aaatgatatt tctgtttta gttttttgcg	27180
tatgtattac gcgattatat aggatttacg attttatatt tatacgttg tatgtataa	27240
ttaggggata tagtaagggt agacggaata aataatttat aaattttgta tttatttaat	27300
tatttttttt aaaaatttat tgtttatttt agcgtatttt tttttgttat attgataata	27360
aatgttttaa tttattgata aattattgag tgttaattat tctattggtt atttatggtta	27420
tatgagttta ttttaatttt atgataatth tacggtggtt attattaagt cggataaaaa	27480
gatgaggaaa ttaaagtta aagaattatg agtgattttt ttaaagataa atttagtgg	27540
ttttaattat aaagttatg atataataat ttttatatta ttttaggtgt ggtttaagag	27600
attatagatt aattcgtaga aattattttt atttaagaa agtttagttt taatatattt	27660
attaagtac gttttgaaaa aggtttagaa aaaagattht attttttatt ttaggttggt	27720
tttatgtaat tgatatgagt attgtttaat ttaaaagtag tattaatttt gaatatttat	27780
ggttataatt ttttttttaa tgagtggga attatttata atgtttaaga ggttttagaa	27840
aagggtgtgt tgtaaattha aaatgataaa ggtagtcgtt gttgttttta tttttattg	27900
ggtgtttttt tttgtagatt tttttggg ggtgtaggta ttgttaggta taatgttttt	27960

-continued

```

tttttgggtt ttaaggtag aaagggtttt gtagtgggg ttggtataaa gcggatttg 28020
agtatgggtg agagtatatt ttggtttaac gaggaatggt agttaatata tattggagag 28080
aaaaatatga atggatagat ttaattaatt ttgtagattt atttttttta ttttatatgt 28140
gagaaaatta agggtttaga attttttaga gtttgtaaat aaatggtaga tttgagtttt 28200
aagttagttt ttatttaagt ttatttttag ttatttgat gttattatta gtattaatgg 28260
tatttagaaa aatagtaatt tttttaaata tgtaatatata tataagttaa atatgtaatt 28320
taaatgtttt gtgaattata gtgtttaaga atgttttggt gttatagtaa atattatgta 28380
aggattaagt ataagatga atttaatatt gtttttaaaa ggtaataata ttttgaattt 28440
tattttggta taaaatatta tgtgttagtt ttatttataa ataaagggtg gagatgtatg 28500
atttgaataa aataaaaatt tatattttag tatattgta atataaatat aggggtgggtt 28560
tttttagtat gttttttgt atttttttaa agttagattt tttagttata ttgattttgt 28620
ttttaagggt attttgttt tgtatagttt tttgtatatt tagataagaa tatatgtata 28680
gaaattattt ttaaaaatta ggataatata ttagttattt gattttatat tattattttt 28740
taatagttgt tttaaatata tatatattaa tatggatatt attaggtgtt agataaaacg 28800
ttaattttta gtatatattt gaaataagtt taatatagga aatagaattt tagggttata 28860
aaggtaaaaa tgtattgttt gtttaaatga ggttttattt ttagtattta tattgaatat 28920
tatttttata attatagttat ttcgtttat attattgatt taatatgtat ttttaataa 28980
tttattttta ttaaaatagt ttatttgata aaggatattt tttatagttt ttataaatgg 29040
aaaattaata aaatcgtta tgaatagaag atgattttgt aaaggaattt atattgtaga 29100
aataaaagtt attttttatg tattatttaa aattatttta tatatagttt gagaaatatt 29160
ggtttaaaaga atttatattt atatgttttt taggtttatt aagagatatt ttttaagtttt 29220
ttttttttat aatcgaata aaatgttgtg ttttaattat tttgaggtat tattttttta 29280
attttgtggt tatttaatta ttgttttttt tgtttaaaat atatttgaag tttttgaagt 29340
tgaggcgttt ttttaaaatt tatgtattat ttattgtttt atttattatt ttatttgtgg 29400
taaaatgtaa ataatttata aatgtataat taaaaagtaa attatttttt tgaatattg 29460
tatcgtttat ttattatttt tttttagttt ttattagttt agtagttaa tatagtatag 29520
ataatgaaga agagaggaa agtagaatag tgggagaaac gtaaagggtt tagaaagata 29580
tatgggggaa attgaagtta gagatttgtt gtttagattt tgaagtgatt taaaaatatt 29640
agtgggtttt gtatgatatg agttagggtt ttagaatatt tgtatatggt taatttagtt 29700
tgaaaaggga tgtgttttgg gaggagtggg gagataaaaa atataagatt ttatagggtta 29760
aattagtgga aatataagat ttatttatta gtaatttttt taattaattt ttttggatg 29820
ggatgttttt gtagttatta tttattattt gtttattttt taaacgtgta tggatttgtt 29880
ttggtattag attttgata agttaaatta aataaagcgt tttaggagtt tataatttaa 29940
aatataattt ttgttttaatt tttattttaa agtaatttta agggaaaaat tag 29993

```

<210> SEQ ID NO 4

<211> LENGTH: 29993

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

-continued

<400> SEQUENCE: 4

```

ttggtttttt ttttaagggtt gtttttaaat gaaattaaaa taaaaattat gttttggatt    60
atgagttttt agagtgtttt gtttaatttg atttatataa aatttaagt taagatagtg    120
ttatgtatgt ttgaaaaata aataaatagt aaatggaat tgtagaggta ttttatttta    180
gaggagtgg ttaaggaat tgtaataag tgaattttat atttttattg atttatttta    240
taaagtttta tttttttat ttttttattt tttttaaaat attttttttt ttagattgag    300
ttgaatatgt ataaatattt tgagagtttg atttatgtta tgtagggttt attggtggtt    360
ttgaattatt ttaagaattt ggtagtaggt ttttgatttt aatttttttt atgtgttttt    420
ttaggttttt tgtgtttttt ttattgtttt attttttttt ttttttttat tgtttgtatt    480
gtattgtatt attgggttaa tagatgttag ggaagggtaa tgagtaattg gtataatatt    540
ttaggaaagt gatttgtttt ttaattgtgt atttatgaat tgtttgtatt ttattataaa    600
taagatagta agtaaaaaaa taaataatat atgaatttta aaaagatggt ttaattttag    660
agattttaag tatattttat ataagaaaag taataattaa ataagatag aatttaaaaa    720
gtaatatatt aaagtaattt ggtataata ttttgttttg gttatgaaaa ggggagggtt    780
gaaaaatatt tttggtgagt ttgagggata tatggaatgt aattttttta attaatggtt    840
tttaaattgt atgtgagata attttaggtg gtatatgaga aataattttt atttttataa    900
tataggtttt tttgtaagat ttttttttat ttatggtggg ttttattagt tttttattta    960
tggtgattat aaaaaatggt ttttattaga taaatttatt taataaaaaat agattattta 1020
aagatatata ttaagttaat aatatagggt atatggttat aattataaaa gtgatattta 1080
atgtaaagt tggaaataag gttttattta agtaaataat atatttttgt ttttataatt 1140
ttaaattttt gttttttatg ttaaatttat tttaaatggt tattgaaaaat tgatgtttta 1200
tttgatattt agtgggtggt atgttaatat atgtatattt aagataatta ttagaaggta 1260
gtaatatgaa gttaaataat taatatatta ttttgatttt taaaaataat ttttgatat 1320
gtatttttat ttagatgtat aaaaaattat gtaggagtaa agtgattttg aaagtaagggt 1380
tagtgtaatt gaaaaatttg attttagaag ggtataaaga aatatgttaa gaaaaattat 1440
tttatattta tgttaataat atgttaaagt atgaattttt atttatttta aattatata 1500
ttttattttt tattttaggt tgaaattagt atataatggt ttatgttaaa atgaaattta 1560
aaatggtggt gttttttgag gatagtatta aatttattat tgtatttaat ttttatatga 1620
tatttgttat agttatagaa tttttttaga tattatgatt tataaaatat ttaaattgta 1680
tatttgggtt gtaatatatt atatatttaa aagaattatt atttttttga atattattaa 1740
tattaataat agtattatag tgattaagag tggatttaaa taaagggttg tttgaaattt 1800
aggtttggtt tttattagta agtttttaaa aattttgagt ttttagtttt tttatatgtg 1860
aaatggagaa aataaatttg taaaataat taaattggtt ttttatattt ttttttttta 1920
gtgtatatta attggtattt tttgttaggt tagaatgtgt ttttaattat gttttaaatt 1980
tgttttgtgt taattttatt gttagaattt tttttatttt gagaattaga aaaggaata 2040
ttatgtttgg taatgtttat attttttaaa ataaatttgt aggaaagaat atttagtaag 2100
tgatgagagt agtaatgatt gtttttatta ttttaaaatt ataattattt tttttttaga 2160
gttttttaag tattgtagat aattttttat ttattaaaaa ataaattgta attataagta 2220

```

-continued

tttagggttg atattgtttt tgaattagat agtgtttata ttagttgtat aagattaatt	2280
taaagtagag gatgaaattt tttttttgaa ttttttttag aatgtaattt agtgaatata	2340
ttaaaattaa attttttttg aatgggagta atttttatgg attaatttgt aatttttttag	2400
attatattta aggtaagtga gaggttgttg tattataggt tttgtgatta gagattattg	2460
gatttgtttt tggaaaagt atttatgatt ttttgggttt tggttttttt attttttata	2520
ttggtttaat aatgattatt gtagagttgt tatggaagtt aaatgaattt atgtattata	2580
agtgattaat ataatgattg atatttagtg atttattaat aaattaaaa atttattatt	2640
aatatgatag agaaggtgtt gttaaaatag ataataggtt tttggaagag gtgattaaat	2700
ggatgtaaaa tttatggatt gtttattttg tttatttttg ttgtgttttt tggttgtggt	2760
atatatatgt gtgggtataa aattgtaaat tttatgtagt tgtgtagtgt atgtgtagaa	2820
ggtttagata tgaatgtta ttttagtaat gtgttttagag aagttttgat gttgttttg	2880
aagtaagttg ttgttgtttg atttttgggt gtttgggatg gatgtttata tttgtattta	2940
gtagtatttg aaggggttta ggttttttgt agtatagttt attttttagat tgttttagttt	3000
tttataatat atatttttat ggaaaaggtt attttttttt gttagaaaaa gtgttttagt	3060
ttggtttggg ttggttttta ttttatgttg ttgtaagtag gtgaagtttt ttttgttttt	3120
ttttttgggg taagtgaaa ggagtttgggt agggggtttg tagtggtttg tataggggaa	3180
ttgggtagtg agagagtttt aggtaatttt gggggttgtt ttatagaagt aggtggggat	3240
tgatagtggt tttttggttt agggaggaga gtgtggttgt gggttttttt ttttagtttg	3300
gaggtttagt ttgttttagt tggtttgggt gggggtggg tgggggtggt gtggagggtta	3360
tggagattat ggtggtgtta tttgggatat ttagggtttt gaggttttggt gtggttttta	3420
tgtgagattg taaattatga taataggtag ttatttgagg ttaaaaaaa atggagtggg	3480
ttttttgtgt gttgttgttt tttgtgtttt tgggtgtttt ttttgaggtt tttggtggtt	3540
ttatgagttt gtagtagttg gtggtgatgt tgtttttggt ttattttttt gtgtaagtgt	3600
gaggttgttg gtagtgtggt gtatgttttg gttgtttttg gtttttgtgt aaaattttta	3660
ttttgtttat gtgaagtttg tgttgtttta gagaggggga aagagttgtg ggaaggttg	3720
gggagtgatg attgtggttg ttgggtgtgt ttttttattt tttttttttt tttttttttt	3780
tttgtttagt tttggagttt tggttttttt tttttttttt ttttttttgg agttggtttt	3840
ttttttgtgt ttgttttttt tttgtttgtg tatgttattt gttgtggggt ggttgaaggg	3900
gatgttttgt ttttattaga ggtatagttg gaaggggaaa ttttgatatt ggaaggaatg	3960
agaataaata ttaattatg gatgtattga attgtggttg ggatagatat tttgggaatt	4020
tgaggtggat tgggtgatga ggtgagtgat tttttttttt aatttttgtt ttagggtttt	4080
tgggggagtt tgagttgaga gaatttttaa attttttggg aaagtgtgtg aggttttgtt	4140
gggagtgtg agtgttgggt attgaggatg tgtagttgga tgggtgtgtg gtgtttgtgt	4200
ttttgggggg tgtttggagg ttgggtgttt tatgttttag ggtttgggtt gtttggattg	4260
tagtgggttt ttttgtttta gaagatgttt ttaagtttta agggtttttt ttgagtttgt	4320
ttgttttttt tggggttgggt ttggagtttg tgtgtaattg agtttattta gtagtttagt	4380
gtgtggtttt tatttgtatt ttgtttttat ttggtagagg tgtgagtatt ggggtttttt	4440
ttatattttt tttatgatgt gtattatttt ttgatgattt ttttagatggt ttaggtgtga	4500

-continued

ggatgttgat ttagagtttt ttggagggtt ataggtgttt gggttttttt ggtggtgggt	4560
gtgtgtgtat tttaaagggt tgtgttttaa tttttaggta ttgattgggt tttttaattg	4620
tggtgatttt attttaatag tttttatgtg gtgtggattg aatgtttttt gtagtttgtt	4680
agggttggtg aaattagagg tgttttggtta gagtagttgt gtttattggt ttgagtagtg	4740
ggtgttatgg aaggtttata atttttttaa aggaagggat ttggttgggt agagtaggtt	4800
tttttttttt ttttaagttg ttgggtttgg ggaggtagtg gaatttgaag tggtttggt	4860
tttttagtgt gtgaagttag gtttggaagt agatgtgtgt gtgtttgttt tatttttgtt	4920
tgtatagtaa tttgaatttt tgtttgtag tatttgaaag agtttttttt ttttgtttgt	4980
gagatttga atagtttga gtgattaggg aatttggga ttgagtttgg tggtagagtt	5040
ggggtgaag tagagagtgt aatttaattt ttgttatttt tttttgtttt gggaggatag	5100
tgtttttttt tattattatt tttttttttt ttttttatga agataatgga tttgtgtttg	5160
gggtgagagt gtgtgtggga gagtgggtgt gagattgttt tttttattg tgttttttgt	5220
gttttttttt gttattttga gtgggtttag agagttattt atgaatttta atttgagggt	5280
aggggaggaa ggtgtaggtt tttttgtttt ttttgtaag gtgtagaata gtgtttgggt	5340
gtgtgttttg gttttagagt agttttatgt ggagtaattt tgtgtgtgtg tgtgtgtgtg	5400
tgtgtgtgtg ttgtgtattt gatttgtgag gaggtaatag gattttgggt tttaaattta	5460
gtgggttgtt ggttattagt ttgttttttt ggttgttatt atagatattt ttaaaatttg	5520
atatttaaga gaattaatag gtttaggtttt atattaagggt tttatattta atagtttttt	5580
ttttttttta ttttttttgg tttggatttt tgggattaat gtaattgttg gagtgaata	5640
tatttttttg aataggtat tttgtttttt tttttatttg tggttatttt gtgaattttt	5700
gggtgtttgt tagtttttag gtgggttttt gtgggattta ggtgggagtt ttaggaggtg	5760
ttaataattg tgggtttttt attagtagtt tttgaagttt tatttatatg taggtgattg	5820
ataggattga aagttagta tgggtttttt gtttgtttgt ttgtttttgt ttttaaaatt	5880
ttagggaagg gattgtattt gaattttttt tttgggtatg gtttggtttt aagtagaatt	5940
agtgtttttt tttttttttt tgttttttaa aatattattg gtaagtttaa atttgaagaa	6000
ttaaaaatta gagggggtgt ggagagaatt tttttaaaaa atatttgata tgtgatatgg	6060
agtgttttag gagattgtat ttaaagatat ttgtgtattt ttaaaaataa tattatttta	6120
agaaataaaa agtagtagta ataagagata gattgttttg tgtggagtaa gattgttaga	6180
atttgatttt tatgtaata atattgaaag tagatataat tatatttata tttattgtta	6240
taaattgtaa aaatagtttg ttttatttga ttaaaagttg taagtattt aaagttaaat	6300
ttgtatttag gaatgattgt ataagaatgt taaaattttt gatagttata gttttgaaag	6360
taatataatt agatttgttg aaaatgtttt ttttttttga ttattagtga tatgtatttg	6420
ttttttataa taagtaagggt tttataatta ataggttttt taaaagtttg tttttttttg	6480
taaatttttt gaaaatagta tgttttaata tttaaagatt attgtaatta ttggttggtt	6540
tataattata attaaaataa gatttattta gttgtttttt ttttattagt atttgttgaa	6600
tatttgaaat tgaatttttt atttttaaaa taaaatgttt tatatatatg tatgtattta	6660
taattaatat tttttttaag tagtaggttt aagttttaaa atttttaaaa ttttaagatt	6720
tgtataaaaa ggagtgattg tataaaaagg aataataatt attaattgtt aagaaaaata	6780

-continued

gatgttgaat tttagagggg ttttagaagt tggagaaaa aattgtatag aagttttttt	6840
tttgatgggt gattggtaga tgagttattt gttattttgt tataatttat tttttttttt	6900
taataagggt ttttgggta aaaaaaaaa aggaaaaata agtattttat ttttaagagt	6960
ttttttttta tttatgttgt attttggggg gtaaataaat tttatatagt tatgggaaat	7020
agagtattta attttatgtt tttagttttt gttgtaagtt ttaaatgatt gatgttgtgt	7080
ttaaaaaat attttatatgg tttgataatt tattttaata ttttatgtat tggaattaga	7140
tttttttaga tttttagaaa gatagaatga tagaaagatg atgattattg ttataaaaag	7200
ttagaatggt aggaagtta tattttaatt ttaataaat taaatttatt ttgttttata	7260
aaaatttata gtttttatat agattttaag aagaggggtg ataatttttt ttattattta	7320
gttaagtgtt aagtaattta agaatgtaag tatgtattat gtttttttat tattattttt	7380
ttttgttagg taaatgtaag aatattggat attttgtaaa aataggtttt tttttataag	7440
gagtttattt atttatagttt tgtattgtag ttgttgatta tattttatgg agttgttaa	7500
gtatttaaaa tttaaaatta ggtattgttt gttgttaaat attttagatt atttatttaa	7560
gaaataaata gaaaagtgtt ttattaaaag taggagatgt tttgaatttt tttattgaag	7620
agttgaataa atttttatta aaatattttt tttttttttt agtgagaata taattttgat	7680
agtttttttt tttaaatgga tattattttt atatgtatta aatgttaaat ttttataatg	7740
tgtttttttg tagagtattt taaattaatt atatttagaa atatagttgg gggattttat	7800
tataatgggt attaggtgaa gaaattaat attaggggaa tgggggtgta ttgtagtgta	7860
tttgtttttt ttatgaattt tttttgttaa attaagaaat gatatgttgt tttaggtatt	7920
tgttttggag atgggtgaga tgtaatatgt gtaatgttgt attttatagt tagatgtgtt	7980
ttaatgaagg ggaatattga tagttattaa attatttttg gagttaaatt tgggtttatt	8040
tagtttgaat ttaagtggaa gaaaatgaat aagagttata tatttaaaag aagtataagt	8100
aattttgatt gtttttaaga ggtttgaaga ttttgtaaat attattgtta tttaatattg	8160
tttggtgagt tgtatattaa tatatgtttt ggtgatatgt tattttatgt ttgaaatttg	8220
ttatttttag tgttttttta tttatttttt tattagtaaa gagatattga aaatgaatat	8280
tattattttt atttttttta ttttttttat attttttagaa aaagagatga aattgattaa	8340
tttaaaatag aaattatttt gtgttattaa aattatattt atatagtgat ttgagatagt	8400
tttagagagt gtatttttga gttttattgt aatttttttt gttattgaaa ggtgttaatt	8460
gatttttaggg tattgatata atagtatagt ttgatatttg aaatttttta gagttggttt	8520
ggtttttttt ttatttttag attttttagt atggatgatg aagaaagata ttattttaaa	8580
atattagaaa tttttatttt tttttaatat gaaatgtttt aatatagtat gtttatattt	8640
ttaaagtttt tatttatata tagtaagtaa atttatttta atgtattttt ggatagtagg	8700
agaagattt atatgtttta ttgtgttaga attttttagt tttttttttt ttgtataggt	8760
agtttattta ggttataatt tttaggtaaa gtttgatttt tattattata tgaatatttt	8820
taaagtgaaa ttgtgttaaa ttaatgtgga atagtttttg tattattaag gtatatatta	8880
atgtagatgt taaatagag agtatatttt tttagatata tttatatttt aaagtgtatt	8940
tttaaaaaa agtggttatt gtagttttta gataattata atttggttgt tattattatt	9000
ataatattaa ttattattat tattataata ttatagttag tattatttta ggaattttgt	9060

-continued

gtaaaatatt gttttaagta tttatatgga ttagttaatt ttaatTTTT taattatgtt	9120
gtaaattagg tatttttgtt atttttatTT tatagataag gaagttgaga tttagttatg	9180
tagtatagtg gagttaggat ttttaatTTTT tagtatgagg ttagtatttt tatttttaag	9240
tgtgtgttg gtttttgttt ttattagtta tatattatta tttttagtg gtatttgaat	9300
atTTTgtttt aataatattt atattttatg taaggataat tagatattag aaaaatttgt	9360
taaatttTgt attaaatttg taattttatg ggtaaatTtg tttgagataa gattaaataa	9420
agtattgaag ttaaagaaaa aaaattaagt atttgaagaa atgtattaag taatagtga	9480
taagaatag gtttaataa ttatagtTat gaagttgga tagtaaatga ttgaatttgg	9540
agaaatgtg tattttgtaa tgttttttaa ttattaagaa tttatgatta gattttttaa	9600
atatttaatt aatatgtgga atttttttt gttttttaag ttttatttaa attggttaa	9660
tgttattaat tttgtatttt ttttgtttt gtttttTgt ttttaatgt tggtttTga	9720
ttttaaattt gtatattTat gttgaattaa ttaataaggt ttgaagagtt aaagagtga	9780
tgatggattt ttggaggtag gtttaaatTaa taatggagtt gtttatattt tggaaatata	9840
tttaattTat ttggttattg tatgtgtaag gttttttgta ggtaaatatt tttgttTta	9900
ataatttTgt ttataatttt atagtatttt tttattgaa tagtggtttt tattttttta	9960
tttatttata tattttgtaa tggatttttt tatttaaatg gaaattgat atagtTtag	10020
tagagaaatg tgttattTgt ttaggtgaa taggatgtag gtagttTgag aatgattTat	10080
taatagtatt tttttatggt ggttTgtTgt tttggagttt ttttaagtTg ttttaaatgt	10140
atggaaatTg ttgtaaaaa atttaagttt tttgataaaa ggggttaatt tagagtattt	10200
gttttaatat gtttatgtgT tattattTaa aaatagattg ttagatatgT taaaatattt	10260
tttagTgTga tatattgTaa gaaaggtgTg tttatttttt tttttataat ttgagtattt	10320
gtattaatta gtttttaaat ttgatttagT attgagaaaa ttaaaattga ttaaaatgtt	10380
ttttgtgtgt agttataggt ttgaaTgag tataaaggat ttgtattTgt aaaggtTgat	10440
tttattaatt agaattatttt ttttttttta aagattgTaa gaagaaatTat tagtagTgt	10500
ttaattTgat atgattttTg atttttatgt aaattattgt tattattttt gttatttttt	10560
tttttttttt ttttaaatTg aaagggatTat tttttgtTaa agattataat taaattataa	10620
aaattTatat ttatgtTgTaa ttaagTttTaa ttttttttat aaaagTaaTat gtatagagtt	10680
tgaattttta ttttttaatta agtaggtgTaa ttatatattg ggaggtTttTaa ttatgtataa	10740
ggttatataT ataattTata gattttttgTaa tatattTaaT ggagTgattt atttattata	10800
ttttattttt gattttgaaT tttttaatgt taaaagattt gaaaagaatt gaatgtttTg	10860
attaagagat tgaattatttt taattTaatg ttttttagTat gttgaaagTg atgatgattt	10920
gggggaatta gtagattttt atattattTaa attttttttt tttatattTg aatgTaaatg	10980
tatatTtatg tTgtgTtatg attTaaTgTaa tttttgtTaa attTaatgat gttgtaggtg	11040
aattatattt agattttttt ttgtaggttt tttagTaaTt taaaataatg ttttttagTat	11100
gtaattTaaTg tatgTaaatt ttaataaatt TgtTaaTaaT ggTaaattTaa ttgtttTatt	11160
ttgttttttt ttttTgtTgt ttgtttttTgt attaaTttTaa gttgataaga TgatgTatt	11220
gttatttttt ttagttgatt tatgaaTaaT tttaaTttTg gtttagtttt tttttttTaaT	11280
tgttTgattt agttttTaaa ggtttTgttt atgaaatTgT ttagTaaagT TgtggTatt	11340

-continued

tggtaaatgt ttgttaaagtg tttttttttt attagtggtg ttgaagattt gtaaaattag	11400
agtgggatgg atagattttt tttttatttt gtatggtttt gaagattttg gagtttttat	11460
tgtattttta tatttttata tattattttt aggaatattt agagagagaa ttattgtaat	11520
taaggttata ggtttaattt ttaggtattt aaattagttt taggtggata attgttttat	11580
atatatttgt ttgtattaat gttgttataa taaatatagt tataatattt tgatgttttt	11640
tttggggggt tattttggat tgtgttgaag ttgtttttaa tttttttta tgtttttaag	11700
atatttaatt atatttgttt atgattttta tgaggttttt tttggattta aaatttttta	11760
aaaaatttgt ttgtgttata tagggtaaat tttaaaattt tttgtaggga agataatttt	11820
ttatattgag aattttgttt aggttgtttt ttttattttt tttttttat ttataataat	11880
ttttttttt agttgtatgg aaaattgtag tttttaaata tttttggat atttttattt	11940
tatttttttg tgtgtggaat gttttttttt tggaaaatat tgtttttttt tttttgtaa	12000
tatttttttt attgtttttt taatatttaa ggtgtaaaat gttagttttt ttgataagtt	12060
tttttgattt ttagagtttg agttgaatgt tttttttttg ttatttttaa agtatttagt	12120
ttttatattt aatatagttt ttgttatatt gaaatgtaat agattatatt agtttaaatt	12180
ttgttttaag ttgagaattt attggaagga ttagttgtat ttgttatttt tgtattgta	12240
atgtatagtg tattgttga tatatagtta gtaggtagtt gttatatttt tattgaatga	12300
atgagtagag gtattgggaa tgagagtaga gattgtgata ttgaatattt ttttttgag	12360
attgggaata gtgagaaggg tagtgattat tataagggtg gattatttta ataataattt	12420
tagggagttt tggataaaat ttagtagaat ttagttgtg aatttaggtt taggttata	12480
gtatgtatat gtgaatgtag aaatgtgttt attttttatt attgggaagt tagtgatttg	12540
gtggataagt taaggatttt gaaaattttt ggagatatgg taattttata gtatttga	12600
tttgagagat ttagttttta tagtatagtg agtaagtaga aagaatatgg gggttgggtg	12660
tgggtgttta tatttgaat tttagttatt tgggaggttg aggtggatgg attatttga	12720
gttaggagtt tgagattagt ttgattaaaa tggtaaaatt ttgtttttat taaaaatata	12780
aaaattagtt ggggtgtgtg gtgggtgttt gtaatttttag ttatttggga ggttgaggta	12840
ggagaattat ttgaatttag aaggtagagg ttgtagttag ttgagattgt attattgtat	12900
tttagtttg ataataagat gagattttat tttaaaaata aataataaa taaaaaggaa	12960
gaagaatag gagttaggta tttgggttta gatttttagt ttaatttga tgagtttga	13020
atattaaata tattattgaa atagattttg ttttagtttt ttatgtgtaa tagaatatgt	13080
tttatgggtt tttgatgaga atttaagtga taaatatatg taaatatatt tagagtagtg	13140
tttgatatga gtattagttg ttattattat gattgggtat tttttattag attattttta	13200
aatgttttag agatatttag agttgaaggt tttttttggg gggttatgga ttggtgatta	13260
taaatggagt gtttagtagt ttaagtataa attagatatt gtgggttatt ttggtaaaatt	13320
tgaggttgta tttggagttt taagagaaga tattataggt gtttaatagt tggatttgg	13380
tatttatggt tgttagtaat atttttttta gtaaatattt ttgtatatag taatttagat	13440
tgttttagag ttttaatttg tgtgggtttt taattatttt aagtaattat aagtatgttg	13500
tataatggta gtattttata gttaaatatt aagttttttt agttgttttt tttttagttg	13560
taaggatta attatgagaa tttgtataat ttttaatagt tagtgggttt gatgaattta	13620

-continued

atgtaaattt tttaaaaggt agttttatat attggtatga tttttatttt tagagagata	13680
gtaggaggta attagtattt ttattttatag aattttattat ttggatttag atataatggt	13740
tatgtgtag gatggggttg aaggaaaaag gaaagttaa atattaattg ttaaggtttt	13800
atttttatag ttgttttga attgttatta ggtgtatttg tatagtttta aaaatgataa	13860
tggtgtattt gaatttttta gtatatttaa ataagagttt ttttaagtttt ggttattgat	13920
attttgggtt ggataattta tttaggggat attgttttgt gtttggtagg atgttttagt	13980
gtattgttgg atgttaggga tttttttttt tttggtaatg atgggttttt tagatattgt	14040
taggtgattt ttgggagtaa aattgttttt ttttgagaat ttataattta agtagatttt	14100
aatattattt gatatagtta aaaaattttt agggttttgt ataagaggat tgttttattt	14160
ttagttaagt gtggagaatt ttaatatttt ttttttaagt ttaaatatga agttaataa	14220
tttttattag ttgtgttga tagattaatt ttgtatagag gggatttata gatattgttt	14280
atatagattg attattttta gtatattttt tttttttgga aataatata gagtgggagt	14340
aatagaatat ttgagagaga gtatgttaga agtaaatatt tatttgaatt atttatgtgt	14400
ttttttttt ggttatttgt tagtagatag ttggattta tttttaaata atgattattg	14460
tttttttatt ttgtgggta aatatttagt aataagaaat ttttatttta ttttagttgt	14520
agtgttagaa aaggtaaagt tattgataat tataattggt gaagattaaa atatttagt	14580
agttaaatta ttgttttga atgagaaatt tttaaagaa ttatgtgtag tttttagttg	14640
taattataag ttaagattt tgagttaata atatttgtat agttagaaga aagtaaaaat	14700
aaatataata tatgtaaaaa tttttttttg tgaattttta gatttaaatt ttttaagttt	14760
tattttatgg gtttgatgtt gattaatata ttttatttat gaaatgatat ataattgatg	14820
taagtaaat tgattgttg ttgggtatta tttttatata atttaagtat atatgaaata	14880
tattttaaaa attatttgaa taaaatggga ataataatgt ttatataaag tgagagagga	14940
taattttttt tatatgagtt ttgggttgaa ttttaaaatt attttgaaat atgaaatatt	15000
aatattataa atatatgatt agattaattt ttttaagagt tttaaatagt agataatttt	15060
tattttatta gaaattgtaa agatatttga aattttatat ttatgtattg ttttaatttag	15120
tggtgtatt ttaattgtat taggtagata attatggaag tagttagatt aatttggaga	15180
ttttatatta tttaataatt gatgtttaag taggtagggt agaggatttt gaaaatata	15240
atatatata atatatata atatatata atatatata atttatttt ttttttttat	15300
ttttttgat ttgtgtttta attttgaat gtgtgttttag gtggtaaagg gttgtaaata	15360
tagtatattg tgtttgatta aaatattttt ttgtttttta ggtggagaag gattttgtat	15420
agttttttgt tagtagagag ataaagtggg tgttgggtaa tttaatgata ggtggagggt	15480
gtttttggg agtgttgtgg agaatttgag ataggggttt aaattataga tagtggagga	15540
gagttttgtt ttatttatgg ataggaagtt taaagaaggt atagtgaatt ataggtttat	15600
ttaggatttt tatgttagat attgaaatgg aagtttagta gtttttagt tattggagaa	15660
tggttaaaaa ttttagagtt ttattttttt aaatgagtat tttttgggtt tataaattaa	15720
ggaattttat agagtattta agttatata tttttttttt aggttgattt tttgtatagt	15780
gggttttaat gtgtaagat ttatttatga atattttatt ttttataatg ttgttagtta	15840
ttttttttt tttttttggt atatttttat tatttttatt ttttatgggt ttaatttttt	15900

-continued

taatttattt gtttagtggt ttttgggtgt ttatttggat aaaataataa taataataat	15960
aatttattta tttttaagg ttttagttta atgttatttt ttttgtgaag ttttttttaa	16020
ggtttaggta gaattagttt tttttttgt gtaataattat agtatttttt ttggattggt	16080
aatataatag tagattgtgt tggggtttat tgtatgtgtt tgttttttag agagattatg	16140
aatttataga ggggagagat tatatttttt tttttttgta atttaatat ttggagatgt	16200
taatggaaat taagtgattg tattatttta tgtatatttt taaagggat taggttttaa	16260
attagatatt tgtaaaggat gaattttggt aattttttgg aaaattgggt tttttttggt	16320
tgtggttatg aaagttggt tattggttat tttttatttt aatttattta aagttttatt	16380
tgtattaaag tattattgat ttttagtggt aaaaataggt atgtttggtg tatattggta	16440
ttttaataa attgatattt aaaagataga tttttgttg aggtttatgg tttttataat	16500
tttagttttt tttagaata gattttgatt tttttggtg tttttgtta agattgattt	16560
tatgttattt tgaagaattt ttaattttta gaattataga ttttattga gagaaagtat	16620
gttttgtatt tgtttgtgtt agtgttttaa aatgtgagga ttaaaaattt tagatgtatt	16680
tattaaaaat atatgtggtt ttttgttgag gaaattata tttttttttt taggtataat	16740
tgaggaagta aaaaaatttt gtttagtttt tagtggtttg ttttaggta tttttatat	16800
aaagatgagt aagatttttt aaaaattatt aatttgttta tttagggggg aaagtttttt	16860
tatgtttaga aataatatta aattaagatt attgttttta atagttttta aaattgtggt	16920
tttatttttt tgtatgagtt attttttagt atagtgtagt atagattgt attatattat	16980
gataaagtat atttgttttg tgttttagtg gttattttta tttattagtg ttagatgatt	17040
agattagaag agagatttag ttttaatttt ttattttgta gatgagatgt tgtattatgt	17100
gattttaaga ttttaagtta tttattttaga ataaaataga aaaagagaat tttatgttg	17160
ggatatttta tgtttttaat tttatatttt attattttat ttttattgtt ttgtttttga	17220
tgaggagttg agaagtgttg aattagatta tttttgtgtg tatttaggtt ttgtattta	17280
ttagtgatgt gattttggtt atgtaatttt ttgagtggtt gtaagttata gtttttagt	17340
gtgtatatgg ggaataataa tagtatttgt tttatggagt tgaaggaat aattgaattt	17400
atgtaggaag gatatttggg ttaggttttg gttattata aaagtttagg ttatagtttg	17460
aattgatatt ttggaataaa atgttgagtg tatttgatat ttagagaagg aaatttatgt	17520
tgttaaagta agtttttttt ttttaattgt ttttaaaaa atataataa gttatatttt	17580
gtataatata ttaaaaataga gatgatagag gttatttttg agaggaatga ttagaggag	17640
taagagagag gttttggaat gttgattatg tttttatttt tggtttggtt gttggttatg	17700
tgaagtgtt tatttaataa aaatttatta gtgattaatt tttgatttat gtattttatt	17760
ttattgtgta tatgttatat ttttttaaaa agtttttaaaa atatggatat atttaagta	17820
aaaagtggat ttatttattt ttattttttg taaattagga taagtgtta tatttagatg	17880
tgtatttttt taggtatttt tttttttata tatgtgtaaa tgaataaatt atgtgaggat	17940
gtaatttatt tttattttat atttttaatt ttaataggat tattgatttt tttaaaattt	18000
gtttttgatt tgttgaattt tgaatatttt tttatattag tatatatagt tttgggtgtt	18060
tttatttaat gtatgtagat attttataat tttattgttt tttttttggt gtatgtttat	18120
attgttttta tattttttatt attattaagt atgttgattt gatgattttt atgttttttt	18180

-continued

gtatatttgt taaaaatatt tgtagataaa attttagatg taatattaat agataaatat	18240
gtatgtgtat tttttataga aaatattaaa atgtttttta aaaagattta tgaattttgt	18300
atttttaagg taatattttt atttttttta taaaaattag gtgtaaaaat ttttaattta	18360
ttgaattttg taattatttg tgttttggtt aagaaaaatg gttatagtag taaatttttt	18420
gaaatttata tagtaattgg aatattttat tatgattagg ataaataatt ttattaatga	18480
tatatattta atatatatat ataatttata taattagttt tttatttaag aatatttagt	18540
atttattttt atgaatgtat attatattta tttatggttt gtaagttta atgttaggaa	18600
gaaataaatt agtaatatta tattttattat attttttttt atttagtatt attaaaatatt	18660
tattagaatt attaggagaa atggttttaa tttttatgtg attggtgaga atttatgagg	18720
aatgatatat tgttatttta gtaagagttt attttgtgtt aggtatagtt ttaagtattt	18780
tatatatatt atgatagttt tttaatatag agatgaaagg tatagatatt gtataagtaa	18840
taaatgatat tattttatta taaatgtgtt gataaatatt attatataatg ttataatatt	18900
agttgtgtta tattaatgat taatttataa atttgttatt tattataatt tgtttaatgt	18960
ggaaataaaa gttaatgtat taaaatttag aaaattgtta tttaatgtt ggtaaagta	19020
atttgataga tatagtttta aatgaagta tattaattag ttgattttgt tgtttgattt	19080
ttatagttag aaaattttgt ttttagaata tttagttttt gtatgttttg gtttaagattt	19140
tgattaattt ttaatggta tttagtgtt ttaaaaagta gtttaattta gaatagttta	19200
tgtatttttt ttgtattgaa taaatatggt tattaatttt tgtggggtt tttttttgt	19260
agaatttttt gttattttta gggatttaat aggatttata tatataaat tgaataattat	19320
attattatgg ggaagtgtt ttgtgttttt tagaaggaat tttttgtta atttaagtat	19380
tgtatgtgtg tagaatagtt aatatttttt ttttaagaaa ttttatgtat gtagttaaaa	19440
tatttttaaa ttgattaagg atttgttata tatttttaga agtatttaag agtattttta	19500
tggttataag agtagagtgt attagaatgt tagaaattaa tgaatatgtt agaggttttt	19560
aaaattgtgt ttaatttttt tgattttatt atataaatat gtgggtagt ttggaaaaat	19620
aaaataattt tttttttttt aatgtagggt aggtttgtga tgataatatt ttaattgggt	19680
tgaatattaa tttgatattt ttttgttaga agtaattttt attatttagt aatgaaatg	19740
gagaggtaat aaatatgatt gatattttat ttttgagtaa agtattttt tatattagat	19800
ttgattagtt atttttattt tttttgaaa atgtattata taaattatta tagtattttg	19860
tttattaatt agggaagtta aaaaatgttt tataaatgtt gaattaaaa ttttattgtt	19920
gtgaggaat taaattttta gtaatatatt ttttttatag ggtgagaaat tgaagtatgg	19980
agaaattata aggttttttt aaagttatat attgagttat ggtagatata gaaatttata	20040
attttgtttg tttgatttta tttgaaattt tatgaaatgt aattatttta ttttttgata	20100
tgtgttttag aagtatattt ttttatttgt ttatttatgt ttttatttat ttatttatgt	20160
atgtatttat gtattattga ttatttatgt atttatttat ttatttatgt attttaaaag	20220
aattggtgtt tttgtatgat atagttttga tttagggtt atgaatattt tgaatttata	20280
tatttttagat atttttttta agaaatttat tgtatttttt ttatgattat gaaaaagaaa	20340
ttaagatatt aagaaaatta agttggattg ttttaagtgt ggttatagta gtgtttatgg	20400
gtagtatttg aagttaattt aggagtgaat ttagtgtgtt taagtgtttg aataagaatt	20460

-continued

tggttgaagt tttggaggag atgtttgggg tatagagtgg ggatgaggaa aaattatagt	20520
atTTTTtagag agtattgtaa ggattagat ttatTTTT attattgat gTTTTtata	20580
ttggttagt taagagagag ggggtaaatt atTTTgttt ttattttgtt atTTTTtga	20640
TTTTgtttt tatgTTTTaa atTTTTTTTT attgtgattt tTTTTtatt agTTtaatta	20700
gtatTTTTt gtgTTTTtTt tattaattt gTTTTTTTT attagttgt ggagtata	20760
TTTTgttg aattgtggg gttataagta attgaaatg aggaaagtgt ttagtaatt	20820
TTTTtagttt agtattttgt aatagagatt ttagagtatg tagtgatgat atttgtttt	20880
TTTTTTTT atagagagt tagtgggtaa tTTTTtggg ttgTTTTtTt gTTTTTTTT	20940
tgtagtTTta agatttaga tttagattg gTTTTtTtTt gtatggTTTT ttatTTTTt	21000
TTTTggaaga gTTTTTTTT gggggTTTT taggtttaag agtattgga tTTTTggag	21060
atgaattgt agatTTtata aagaatattt aagtattag taaattatat tTTTTttag	21120
TTTaatgtt agTTTTTTTT tTTTTTTaa tttattgaa gaataattt gTTTTtaata	21180
aataagtaa tagaatttta gggTTTTTTTT taaagggtta tTTTgtgtt tttgaagatt	21240
attattatta atTTTTatt tagttaattg gTTTTTTTT atattttat ttgggtata	21300
atgtgatatt ttgtatata ttaatgtga tttatattaa tatgttaatt tataattgag	21360
tgatTTaaaa ataatgatg attatgattt agtttatatt tgTTTTtat ttaatgtatt	21420
TTTTaagtaa ttatgtttt gtggttgatt gaatTTTTt ataatTTTT gttgggaaat	21480
aaagattttt tTTTaaattg aagaagtggg ttgttaatt aggggtattg tatttaatt	21540
gagagaaaa gattaggaat gataggaatg atagtatatt tTtaattt aatagttgt	21600
tatatatgtg aaagtaggat atatagtagg gaagtataat gttgagtga tTTaaaaat	21660
taaatgtatt tgatatatta tTTTTtgaat tTTTataatt ttagttttt taaaggttt	21720
tggatTTTta gaagtataa gggtagttt ttgtaaagg ggtttattt ttaatttagt	21780
tatagagtta gtggaatgat ttttatattt aaatgttata ggggaattt ttgagaggaa	21840
aaagggtatt taaatTTTT ttgatTTTT aattaaaaat ttggattaat aaaatTTTT	21900
taagtgtttt aaatttaagg aaagtaaata gtattttatt taattatatt tTtaattagt	21960
atTTattata gatataattt aaattagtaa tttgtatta gatgaattag atttattgtt	22020
atattatata atTTTTatgt ttattttgtt gtgtttggga tTTTTTTTT tttataata	22080
agaatgat ttataggtgt tatattttga tttttgagaa attaatgtta tagaaaagt	22140
gtttgaaag atataattgt gattttgtta ttgtttttgt tgttattatt tTTTTtatg	22200
gggatgggta tatatTTTT gagatttatt agatttattt ttagtttat aaattattt	22260
tttagatgta tTtaatttTt tgttagttt gTTTTtTga aaaaaattt gtgaatata	22320
TTTTaaaaat atTTTTatgg tTTTTTTTT ggtttgtgtg tatatTTTT ttgtgtgtgg	22380
TTTTTaaatt ttattttgtt tTTTTtTg ttgtttTaa gtatatttat ttagttttt	22440
ttatattgtt ttgtagttt atatttttag atTTTaaatt tgtgtgtgtg gttatTTTT	22500
attgtgtttt gTTTTttat gtatTTtata atTTTTTTTT tTTTTtTgag atggagtgt	22560
gttttTgtt ttaggttgga gtgtagtggt atgattttgg tttattgtaa tTTTgtttt	22620
ttgggttttag gtgattTTTT tgTTTTagtt tttgaaatg ttgggattat aggtgtttat	22680
tattatgttt agTtaatttt tgtatTTTta gtagaaaagg ggtttTgtta tgttagttag	22740

-continued

gttggtttta aatTTTTgat tttAAGtgat ttgtttgttt gagtttttta atgtgttggg	22800
attataggta tgagttattg tattttgttt gtagtttata atttttggtg gatttttttt	22860
ttagaatTTT ttgtattttg tggagtaatt ttgtattttt ttggagtta gggattttta	22920
atTTTtaggt tagtttttat tttattttat tttttgttt tgaggttttt atattataga	22980
tgtagtaaat atttagaatt tatattttgt aaggttgaag ttttgattta ttttatttaa	23040
gaaattatTT ttttttattt tttttttaa ggttttatta taggataggT tttttgttg	23100
aatttgatgg tgggttgagt ttttttattt ttttttaaat aaatgtgata aatTTTtag	23160
atTTTttttt ttgtggattg tttttgttt ttgttttttg tttattgtta taaaagaaga	23220
taatgataat aaaatgataa taatataaat gtaaatgtat tttttttttt taatTTTttt	23280
atattattaa ttttttatga attaaggTgt gatatttgaa agttatattt gttttattag	23340
ttatatttat atagtttttt tgttttattg atgtattttt tatattattt tttatattta	23400
tttttttagg ttttttttta ttttatgtta gtttagattt ttttggtttt ttatttgag	23460
tgttatagta gtttttaagt tgttatttaa tttatttttt tgtttatttt gtttgttaga	23520
aatttgaagt agatgtgttg ttgtaatgat attatatgtg ttattttttg ttttgggggt	23580
tttaaggTtt ttaaggTtat taaaatattg taataattaa ttgttttaaat tgtattttgt	23640
atTTtatttt ttgttttttg taatttaaaa tgttttaata ttttgTtaaat attgaatgtt	23700
gatgtttttt tatttttttt ttttttatat gtttgtttat aattgtttat tgattatatt	23760
agttttttta tattaaaggT attagttata attttgatgt tatatagaat gataattaa	23820
aggTgataaa tgttaatgat ttaattgagg agttgtttat tttataagt ttatttgagt	23880
ttagagaatt ggtaaaagat tggTgatttt tggttattat ttaaatataa ttggTtaatg	23940
tagtttttgg agattagtgg ttagtataat atttttagtt gagtaggggt aaaaataatt	24000
atatttgttt aagttgtaga tattatttta ttgtatagta aagaatgtta aaatgtttta	24060
atTTagtata tgaagagtaa aatataaaa taggtttttt tttattttta ttagtaagta	24120
aagTggtttg tatttttgaa atatattttt tgtttaatta taataaaata taatataaat	24180
gttattaaaa gttttatttt atggatgaag atagtgatat ttagagatat taattagttt	24240
ttttaagatt atatatttat tttagtagtaa aagtaagttt ttataagttg atttatttta	24300
ttttagtttt tatttggaat tattaaattt atattttttt tagtaaaatt attgTttata	24360
tttttagggT tttAagagtt taatgtttaa tttgggagat ttatttttaa atTTTttta	24420
agtGtaatat atTTTtttta ggatgattat aaagttatta gatgtgtgtt aagtaatata	24480
tatagttata tatttttgat taaaaaagat ttttaggtta tatagaagag agatttatgg	24540
ttgagaaaa ggtagattgt ttttaagaaa ggttttaaat atattttttt ttgataggtt	24600
ataggagagt tagttagatg aattattata gtgtgtataa aatgtttagt agaataattt	24660
tttagtattg gtagggaaga gtaataggag taatataaga gaaggaaagt tggTttgtag	24720
agTttgtttt agtatttgta ggggtgatga ggttGtttg gttAagaatt tttttatag	24780
ttttttggg ataattttat tagaattatt ttgtgtttt ttaggTtatt taggtaatta	24840
ttattaatta gtaaatatag ataaaaata atattttgtt tggattgtt tgttaaaaat	24900
tAAatgatta taatttataa ttttaggaag atttGttatt aaatgaattt atgataaatg	24960
atTTaaatga tttttataaa gtagagtttg ttgtaatttt tagtatgaga agtatgaatt	25020

-continued

ttattattat tttttaaata ttttgagtta aatggattgt tttgtaaagt ttatgtttta	25080
ttttataga tagtttaaaa atagtttagt atttagtttg taggttggtt ttatttaagg	25140
agtttgaagt taggtaaggt gtttattatt atttatttag atttgttttt tggatatttt	25200
agttgttttt ttttttatg gttagggagg agagttagt gtatagttaa agaagtgtga	25260
ttttatggt tttgaaatag tggttataga attggagtta ggtgtttttt gtgttattgg	25320
ttttttttt tggtagttt tttgggtggt tggagttagt tagtgttga tttagttta	25380
tagatttagt aatgaaattt ggaggtaga gtaggtaatt aggtatataa gaattgtaa	25440
gtaagttaa tttgtagatg tgattataat tataattttt taataatgta ataggaatat	25500
aaattgtag attttttata ttattgtaat ataaagaaga taattgaagg ttgtgaagat	25560
agtttatgaa atgtaaggaa gaattggata attaagtaat aggaagagta gagaaatag	25620
gagatttgg gagttataag atttgggtgt ttaaattttg ttaggatttt tttgttatt	25680
ttattttgt ttttttttt tggtagttt tttttgttt tttttatag attagttttt	25740
ttttatatg atagaaatg tttttaagtt ttatatttta ggggtttgt tattgaagga	25800
ggattagtt gatatttttt tagatttata ttttaaggaa gatgaattat tagatttaag	25860
taattttgt taattggaga agattttag attttattta atgggttaag ttggatatgt	25920
atttgggga gggagatag ttttgggtt ataatatata tgaagttaa ttttaatt	25980
gtataagata tgttttgaa aatttattg tagttgaaa atttggagt taaatagtag	26040
attaattggt gtgagtgtta gaatttagt taatggtatt ttttggatt taggtaattt	26100
tgtataggta gaaagatttt tgaagtttt atttaggtt gaaagggtt gtgttagagg	26160
ttatgagtgg gatttttttag gtatttgttt ttttatttta aggtttgtat atggattgga	26220
ataaaagtt tggtagttg aatttttaaa tttttttta aggtatttta gggatttaa	26280
gtagtaggga atgtattttt ttattttttt ttgtttttt ttttattgta gtaagggtta	26340
agtatagttt ttagttttgt ttgtattttt ataggagatt tatgtatagg ttttgggtt	26400
tggttgtagg atagtgtgag gttgttgata gggtagaga atggtagaa taatggaata	26460
ggattgaggg agaaggaaga ggaattgttt ttagaattat tatttgggat ggtggtttta	26520
gggagtaaaa aatatttttg gattttttt aagggtttt ttgtagagt ggtattattg	26580
ttagatattg ttatggttgt aggtaggatt gattaaagaa tttttgatt tttttgagga	26640
tagggtaaag ttttgaattg tttttagaga ttgtatttag atttatatt tggaaattat	26700
atttgggtt tattgtgaat ggtaattgtt agggagtagg attggaggga aggagattat	26760
tatgtggtt ttgtttttgt ttaagtagg aataatgatg gtggaatgtt tgtattggtta	26820
gagaggagag agaggaggga tttggtgaat ttttgagggt attgtagga tttgtaata	26880
tattgaatgt ggagagtaa tggtaggga gttagtata aagtttaggt gtttggttt	26940
gatgattggt tatgtttgtg ttggttgtg atagaattta gaggttagt gttttatgt	27000
gtagtgttg tttttattt gtgtagttt ggtgtggagt tgagatgtt ttttttatt	27060
ttatgtttat gtttttttag tttagttaa gtatgtttg ataatttag ttttttttt	27120
ttttttttt ttttttttag agagagagg tttttgttt gttatttaga agaagttag	27180
tgttgtgatt atagtttatt gtagtttga atttttgggt ttaagtgatt agaggattag	27240
gtgtattttt ttataaattt tttaaattt ttttagtatt ttatttttta agggattttta	27300

-continued

tagtttttt	tatttgatg	gtgttatggt	gttttatttt	atttttgaga	attgttttaa	27360
aattttatat	agaaatttgt	tattatgtat	ttaaagtaga	agaaatggtg	aatttgtaat	27420
tttaattttt	taagataaga	gtttgtagtt	tgtttttagt	agagttaaag	aaaaagatat	27480
gatgaatggt	ttatatattg	aaaattgtag	aagaaataat	ttaatatgga	atgggtgagg	27540
ttgttatagg	tggtgtaaaa	ataattgtgg	tttttgttgt	tatttttagt	ggtaaaaatt	27600
gtaattgttt	ttgtataaat	ttaataatat	tgtttgaat	ttatattttt	tggtggaata	27660
gttttgagat	agagtgggaa	ttagattatt	ggtttataaa	atagatgta	gttattgtta	27720
ttttttaatt	ttatattttt	attatgaggt	aatgtttttt	ttatttgatt	tgagagaatgt	27780
ttttattttt	ttttatata	tttgtaaaa	ttaaagtta	aaggaattaa	tataataatt	27840
attgttaaat	tatgatgtaa	aataatattg	tgtttgtaat	tagtttaggt	agttttattt	27900
ggtttaataa	tataaaatta	ttaggtttat	aatatttttag	tatttttagt	tgattgggta	27960
agaaaaaat	ttatgttatt	ataaaaagtt	aaattaagta	ataattaatt	taaattttta	28020
ggaaatataa	agtaaagtat	tatgttgtaa	tatggaatat	taatgtaaat	aattttttta	28080
tagtttatga	aatagatttt	gaattggtag	ttgataaaga	aaggttagtt	ttggtttaat	28140
ttttttatgt	tttaatattt	ttttggaagt	tttatttaat	tttatttggt	ttagtaatag	28200
tattttgaga	tggtgggtgta	tttgaaaaat	tattattttt	tttttgataa	ttagtatttt	28260
taatttaatt	tttatatatg	ttttttggaa	atattatatt	tatattttga	tgatgtttta	28320
aaaataattt	taatgtaata	attatatatg	tgtggtgagt	aaattgtggt	ttgtaaagaa	28380
tatttttaaa	ttttaatggt	tgtatttggg	aatgtgattt	tatttgaaa	tagggttttt	28440
gtaggatga	ttaagttaa	atgagatttt	attggattag	gatgggtttt	aatataatga	28500
ttgttgatt	taagaggaaa	atgtgtatag	agatatatat	atatataaaa	gaatgttatg	28560
taatgttagg	tatggagaat	attatgtgat	aatgaaggta	gagattggag	tgatgtattt	28620
atgagttaag	gatgggttag	aattttttaga	agttggggag	agataaagga	tgttttttta	28680
ggatttttag	agtaagtata	aatttgtag	tattttgatt	taggattttt	tttttttaga	28740
attgtgagaa	tagatttttag	tttaagttta	tttagtttgt	ggatttttgt	tataggagtt	28800
ttaggaaatt	aatgtaatat	gttatgttgt	aaataaatag	tagttttta	tggtataaat	28860
aataatttta	agtatatttt	attattaggt	taggtgagta	ggaaattggt	tggtagatta	28920
attattaatg	ttgtattttt	tgaaaagtta	aaatattttt	attagttaa	taatatttgt	28980
ttaaaagt	ttatttgtat	tgttttaaaa	taaagtaaaa	attaaaatgg	agtttagttat	29040
gtatttgaaa	aatgtgaaaa	aggaaagaaa	aatttttttag	gaaagatttt	ttttaaaatg	29100
gttatatatt	tgagaaagaa	ttttaaaatt	taaattattt	ggtgttttat	tttatttttt	29160
tgtttatatt	attagatata	ttatgtatga	ttttatata	tatgtatggt	atatttaatt	29220
aatatttatt	aaatatttaa	aatttgtag	gttttgagtt	aggtattgta	gataggtaa	29280
taagataaat	attttttttt	tttaagaga	ttattttggt	atgaaattta	atgttagaa	29340
gaagtgaag	aggttatttg	tagagtttat	tatgaaaatt	tatttttatt	atgttgtga	29400
gttttatttg	attttgtag	attatttttag	tattatttgg	ttttatttgg	tgtttttga	29460
ttatattgga	gaaaataaaa	tttatttatt	gtatagataa	tataaaatta	attaaagtat	29520
gtgaaattaa	aaatgtaatt	gtttattaaa	tttttttttag	tgttttttat	ttattagtat	29580

-continued

```

ttaaataat ttttagttgt gtatttatgt tagtataatt aaaatgtggt gtttttattt 29640
ttttatttaa tttataatta ataatgttta tttttttatt tttttttttt gagatagagt 29700
tttattttgt tatatttagg ttggagtgta gtggatgat tttggatat tgtaattttt 29760
gttttttggg ttttaagtgt ttttttgttt tagttttttg agtaattggg attataggtg 29820
tttgttatta tgtttggtta attttttgta ttttagtaga gatagggttt tattatggtg 29880
gttaggttgg ttttgaattt ttgattttag gtgatttggt tattttgggt ttttaaagtg 29940
ttgggattat aggtatgaat tattatgttt ggttttttat tatttttata att 29993

```

<210> SEQ ID NO 5

<211> LENGTH: 29993

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

<400> SEQUENCE: 5

```

aattataaaa atgatgaggg gttagggtg gtggtttatg tttgtaattt tagtattttg 60
ggagggtgag gtgggtggat tatttgaggt taggagttg agattagttt gattaatatg 120
gtgaaatttt gtttttatta aaatataaaa aattagtttg gtatgggtgg ggtggtttgt 180
aattttaatt attttaggag ttgaggtagg agaattggtt gaatttagga ggtggagggt 240
gtagtgtggt aagattatgt tattgtattt tagtttgggt gtgatagagt gagattttgt 300
tttaaaaaaa aaaagtggaa aagtgaatat tgttgattgt gaattagatg aaaaaataag 360
aatattatat tttagttatg ttaaatgga tgtataatta agaatgtatt tggatattga 420
tgaatgggaa atattaaaa aaatttgata aataattata tttttggttt tatatatttt 480
aattaatttt atgttgtttg tgtaataagt aaattttatt ttttttaata taatatagaa 540
gtataaagta aaattaaatg atgttaaaat ggttttatag gattaggtaa aattatagta 600
gatagtgaat atgaattttt atggttagatt ttgtaaatgg ttttttttat tttttttata 660
gattaatttt tatgttaaaa taattttttt aggggaaggg aatgtttggt ttatttattt 720
atttataata tttagtttaa ggtttagtaa attttaggta ttttaataat attgattgaa 780
tgtaatatat ataatgtatg aaattatata taatgtattt gataatgtaa ataaaaaat 840
ggaataaata gttagatgat ttaagtttta aaattttttt ttaaatgtat agttatttta 900
ggaaaaattt tttttaaaag attttttttt ttttttttat attttttagg tatataattg 960
attttatttt aatttttggt ttgtttttagg ataatataga tagaaatttt tagataagta 1020
ttattttattg aatgggaata ttttaatttt ttagaaaata taatattgat aattagttta 1080
ttaggtaatt ttttatttat ttgatttaat aataaaatat atttaaaatt attatttata 1140
ttaattagaa ttattgttta tttataatat agtatgttgt attagttttt taggattttt 1200
ataataaagt attataaatt ggggtggtta gaattgaaat ttatttttat agttttggag 1260
gaaagaagtt ttaaatgaag gtgttggtag atttatgttt gttttggaga ttttagggaa 1320
gtattttttg ttttttttta gtttttggg gttgttggtt atttttggtt tatagatgta 1380
ttatttttaat ttttgttttt atgtttatat gatgtttttt gtgtttaatg ttatatggta 1440
tttttttggt tgtgtgtatg tttttgtata aatttttttt ttaagtataa tagttattgt 1500
attagatttt attttaattt agtaaaattt tatttttaatt tgattatatt tghtaaagatt 1560

```

-continued

ttgtttttaa atgaggttat atttttaaat atagatatta ggatttaaaa gtatttttta	1620
tggaatataa tttattttatt atatatgtat aattattata ttgaaattat ttttaagatg	1680
ttattaaagt atgaatgtaa tattttttaga aagtatatgt aagaattaaa ttaaagatat	1740
tgattgttag gaaagaagtg ataatttttt agatgtatta ttatttttaga gtattgtttgt	1800
taatgtaagt gaaattaaat aagatttttag gaaagatatt ggagtataag agaattaagt	1860
taaaattagt ttttttttgt taattgttag tttagaattt attttatgaa ttgttgggga	1920
gttatttata ttaatgtttt atgttataat ataataattt gttttatatt tttttgaaat	1980
ttagattaat tgttatttgg tttgattttt tatggtaatg tgaatttttt ttttggttaa	2040
ttatattaag gtattaaaa attatgagtt tagtgatttt atattattga attaagtaaa	2100
attatttagg ttaattgtag atatagtatt attttatatt ataatttagt aatggttgtt	2160
atgttaattt ttttaatatt tagtttttat aaatatatga gaaaaagatg agaataattt	2220
ttaaattaag taagagaagt attattttat aataaagata taagattaag gaataatagt	2280
aattgtatgt tgttttatga gtttaataat tagtttttgt tttattttaa gattgtttta	2340
ttagaaaaata tgaattttta atagatttat taggtttgtg taaaaataat tgtggttttt	2400
gttattgaaa gtaatagtaa aaattgtaat ttttttga ttattttag taattttatt	2460
tattttatat tagattattt tttttgtagt ttttaatata tagattattt attatattt	2520
tttttttaat tttgttagaa gtgagttata aatttttgtt ttagaaaatt agaattataa	2580
atttaattt tttttttatt taaatatata atagtaagtt tttatataaa gttttaagat	2640
aattttttaga aatggaatga ggtattatgg tgttgttttag gtgggagaga ttgtaagggt	2700
ttttaaggaa tgagatgttg aaggaggttt gagaagtttg tagagaagta tatttaattt	2760
tttgattatt tgaggttagg agtttaaggt tgtagttagt tatgattatg atattgtatt	2820
tttttttagg gatagagtgg agattttttt ttttttagaga aaagaaagaa agaaagaaag	2880
aaatttaggt tgttttagta tgtttttgtt tggttaaaga aatgtgaata tggagtggag	2940
gggttatatt ttaattttat attaggttgt tgtggaatga ggattgatgt tgtatatgga	3000
aatgttaaat ttttgattt tgtatataat tggtagatt atagtttagt attaaagttt	3060
aatatttagg ttttgtatt gattttttgt ttatttgtt tttatattta atattattt	3120
aagttttatt aatattttta aaagtttgtt aaatttttt ttttttttt ttttgttagt	3180
ataggtgttt tgtattattt gttttttgtt tgggtaggag taatagttat gtaatggttt	3240
tttttttttt agttttgttt tttggatggt attgtttata gtgtaattag agtgatgggt	3300
ttgaaatata gatttgaatg tagtttttaa aaataattta gagttttatt ttgtttttag	3360
aagaaattag aaatttttta attagtttta tttgtagtta tggtagtgtt tagtggtagt	3420
attgttttgt taaaaggttt tttgaagggg ttttaggggt gttttttgtt ttttgggatt	3480
attgttttaa gtagtagttt tggggatgat ttttttttt ttttttttag ttttgtttta	3540
ttgtttttgt tatttttttg ttttgttagt agttttatat tgtttttag tttaggtttg	3600
ggattttgat atgagttttt tgtgagggta taagtagggt tgagggttat atttgttttt	3660
tgtttagtg aaaaggaaaa taagggaag tggggagta tttttttat tgttttaaat	3720
ttttagatat ttttgggaaa agatttaaaa attgtagttg ttaaaatttt tgttttagtt	3780
tatgtgtaga ttttgggatg gggaaatagg tgtttgaaga gttttattta tgatttttga	3840

-continued

tatagatttt ttttaatttg ggtggggttt ttaagaattt ttttgtttat gtagagtgt	3900
ttgggttata gaaagtgtta ttatattaaa ttttaatttt tataatagtt ggtttattgt	3960
ttggttttta gattttttaa ttgtaagtaa attttttagaa gtgtattttg tataatatta	4020
aggtttgttt gtatgtgtat tatatttata aaatatattt tttttttata gatgtatatt	4080
tagtttaatt tattaggtag ggtttgtaag ttttttttag ttgtatgaaa ttatttggat	4140
ttaatgattt atttttttta ggatattgagt ttgggaaaat attaggttag tttttttta	4200
gtggtaaaat ttttaggata taaaatttg aaatgtttta tatttatgtg gaaaaagtt	4260
ggtttgtagg gaaagataga agagaattga ttgagaaaga ggagtaaaaa tgaagatggt	4320
aagaaagttt tggtagaatt taagtattta gattttgtga tttttaaagt tttattgttt	4380
ttttgttttt tttgttgttt agttgtttta tttttttttg tttttatga gttattttta	4440
tggtttttaa ttattttttt tatattgtaa tggtagatga agtttaatag tttatatttt	4500
tgttatatta ttgagaagtt atagttgtga ttaattgttt agatgggtat tgttttataa	4560
tttttgtgta tttagtattt tattttgttt tttaaatttt attattgggt ttgtgaagtt	4620
ggattaggta ttggttgggt ttaattattt gggaaagtat tttagagga aagttaatgg	4680
tataaagaat atttagtttt aattttatgg ttaattgttt agagatatgg agatgtatat	4740
ttttggatta tgtatatgat tttttttttt ggtttagtaa aagaaaaata gttgaggtgt	4800
ttaagaaata gatttggatg aatagtaatg agtattttgt ttgattttta gttttttaa	4860
tggagttaat ttataaatta gatgttgagt tatttttgag ttgtttataa agatagaata	4920
tggattttgt aaaatggttt atttaattta aaatgtttta aaaatgataa tggggtttat	4980
attttttatg ttaaaagtta taataagttt tattttgtga gaattatttg ggttattttg	5040
tataagttta tttgatggta agtttttttg gaattataaa ttatgattat ttgatttttg	5100
ataagtaggt ttaataaagg tgttttattt tgtttgtgtt tgttaattgg tgatggtgt	5160
ttagatagtt tagaaaatat agaagtaatt ttgagtgaat tgttttaggg agatatgtga	5220
aagaattttt agtttaagta attttattat ttttatggat attaaggtag atttttaga	5280
ttaatttttt tttttttgtg ttatttttat tgtttttttt tattagtatt gagaaattat	5340
tttgttagt attttgata tattgtaatg gtttattttg ttggtttttt tgtagttttg	5400
tagaagagaa tatatttgaa gttttttttt gaagtaattt gttttttttt tagttataaa	5460
tttttttttt gtgtaatttg gaaatttttt ttgattaaga gtatatgatt atatatttg	5520
tttaatatat atttgatgat tttatgatta ttttgaggaa aatgtgtgtt attttgagag	5580
aatttaagaa tagatttttt aagtttagta ttgaattttt gaaatttttag aaatgtaa	5640
gatagttttg ttaaagggaa tgaagtttta gtgattttta ataaagattg gaataaaata	5700
gattagtttg tgaaaatttg tttttattat tggataaatg tatgattttg gaaaagttta	5760
ttaatatttt taaatattat tgtttttatt tataaaatag ggtttttaat aatatttga	5820
ttatattttg ttatgattaa ataaaaata ttttttaaaa atatagatta ttttgtttgt	5880
tagtgagagt ggaagaaaat ttgtttttat attttgtttt ttatatattg aattaaaaata	5940
ttttaatttt ttttgttga tagtagagta gtatttgtaa ttttaataaa tgaattatt	6000
tttaattttg ttttaattaag agtattatat taattattga tttttaggag ttatattaat	6060
taattatatt tagatagtag ttaaaagtta ttagtttttt attaaatttt tgaatttagg	6120

-continued

taagtttga	aaatgaata	atTTTTgat	taagttatta	gtatttgta	tttttagtt	6180
gtattttgt	ataatattaa	agttgtgatt	ggtgtttta	gtgttgagag	attgatataa	6240
ttaataaata	attatagatg	gatatgtgaa	agaaaagga	ataggaaagt	attagtattt	6300
agtattaata	ggatgttaag	atattttaag	ttggtaaaag	taggagaata	agtataaaat	6360
ataattaaga	taattggta	ttataatatt	ttgataaatt	taaaagtttt	aaaattttta	6420
aaataaagaa	tggtatatgt	gatattattg	taataatata	tttattttag	atttttgata	6480
aataagatga	ataagaagtg	gagttaagtg	gtaatttaaa	ggtgtgtgtg	atgttttaag	6540
tgagaagtta	agagaattta	gattgatatg	gggtggaaaa	aggtttgag	aagtggatat	6600
gagggataat	atgagagata	tattaataag	ataaaggaat	tatatgaata	taattgataa	6660
ggtaaatgtg	atttttaggt	attatatttt	aatttatgag	aaattgatga	tatggaaaag	6720
ttaagaggaa	aaggtatgtt	tatatttggt	ttgttattgt	tttgttgta	ttattttttt	6780
ttatggtgat	gggtaggaag	taaaagtaag	aggtaattta	tagaggaagg	agtttaaaa	6840
gttgttata	ttattggag	gggggataga	aaaattgggt	ttattattag	atntagtaa	6900
gaagtttgtt	ttatgatgga	gtttttggga	agagaataaa	agaaaataat	tttttgagtg	6960
agataaatta	aaattttagt	tttgtaggat	gtaaatttta	aatgtttatt	gtatttggg	7020
tatggaaatt	ttaagataaa	aaaataaaat	aaaataaaga	ttgatttaga	ggtttgaaat	7080
ttttagattt	taggggaatg	taaaattggt	ttatagagta	taaagggttt	taagagagaa	7140
atttatataa	agttataaat	tataggtgag	gtgtggtggt	ttatgtttgt	aattttagta	7200
tattgggagg	tttaggtagg	tagattattt	gaggttagga	gtttgagggt	agtttggta	7260
ataggtgaa	atttttttt	tattaaaaat	ataaaagta	gttgggtatg	gtggtgggta	7320
tttgtgattt	tagttatttg	ggaggttgag	gtaggagaat	tgtttgaatt	taggaggtag	7380
agattgtagt	aagttaagat	tgtgttattg	tattttagtt	tgggtaatag	agtataaatt	7440
tattttaaaa	agaaaaaaa	aattataagt	tatataaaga	aataaaatat	agtagagagt	7500
agttataata	ataaattagg	aatttaagaa	tgtgaaatta	tagagtaatg	tggaaaaaat	7560
tagataggta	tatttaaaat	agatataaaa	gaaatagaaa	taaaattgaa	aagttatata	7620
tgaaaaaat	atataataa	atttggaaga	gaattataaa	aatgttttta	aaaatatatt	7680
tattgaattt	ttttataaga	aataaattaa	atagtaaatt	agatatattt	aaagaaataa	7740
tttatgaatt	ataaggtaaa	tttggtaaat	tttaagaagt	gtgtgtttat	ttttataaaa	7800
gaaaaaatg	ataataaaaa	taataataaa	attataatta	tattttttta	aataattttt	7860
ttataatatt	aattttttag	gaattaaggt	gtgatatttg	tgagtatat	tttttattat	7920
ggggagggag	aagtttttaa	tataataaag	taaatatgag	agttatataa	tatggtaata	7980
aatttgattt	atttaatggt	aagttattga	tttgaattgt	gtttatagtg	aatattaatt	8040
aagaatgtaa	ttaataaaga	tgtgtttgt	tttttttaa	tttgggatat	ttgaggaaat	8100
tttattgatt	tagattttta	attgaggtat	ttaggaaagt	ttgagtattt	tttttttttt	8160
taataaattt	ttttgtagta	tttgaatata	aaaattattt	tattgatttt	gtaattgaat	8220
tagaagtagg	gttattttta	ttaaaaattg	tttttatagt	ttttaaaaat	ttaagatttt	8280
ttaaaaggtt	tgaagttgta	agaatttaga	aaataatgta	ttaaatatat	ttggtgtttt	8340
gagttaaatt	aatgttatgt	ttttttgta	tgtgttttgt	ttttatatgt	atggtaggtt	8400

-continued

gtagatatt	aggagtatgt	tggtatTTTT	ggtatTTTTa	atTTTTTTTT	tttaagttaa	8460
gtatggtggt	tttggttaat	aagattatTT	ttttaattta	aaaaaaaaatt	tttgTTTTTT	8520
aataggaat	tataagggaa	tttagttagt	tataaaagta	taattgttta	ggaatgtat	8580
taagtgaag	gtagatataa	attaggttat	agtgtattat	tatTTTTaaa	ttatTTaatt	8640
ataaattgat	gtattaatat	agattaatat	taatgtatat	taaagtatta	tatTTatatt	8700
taaataaaaa	tgtaaaggaa	aattaattag	ttgaaataga	aattaatgat	gataatTTTT	8760
aaagtagtaa	gaatgtatTT	ttgagaaga	tttaggggt	ttgTTTTgtt	gTTgttgga	8820
gattaaatta	TTTTttagta	aaattgaagg	gggaaaaaag	attaatgttt	aagttatagg	8880
aagatgtggt	ttgTTTTgga	tttggatggt	ttttgtagaa	tttataagtt	tatTTTTtag	8940
gggtTTaaat	gTTTTtgat	ttaaaaagtt	ttgggaaag	aatTTTTtta	aggaagaga	9000
tggagagtta	tataagtaga	aattaagttt	gagTTTTtaag	ttttgggggt	gtaagaaaag	9060
aataagagaa	taagTTaaaa	aagttgttta	ttaaatTTTT	tatagaaggg	gaggagggtta	9120
ggtgttatta	ttatgtatTT	tgaagTTTTt	ggtataggg	attggattga	gggaagttgt	9180
tagatTTTT	TTTTatTTta	aattatTTgt	aatTTTTtagta	gTTtgattaa	aaaatataat	9240
tttgtaagtt	gatggaaaaga	aatgaatTTg	gtgaagaaa	tgtgggaaaa	tattggttaa	9300
attagtaaga	aaaagattgt	aataaaaaag	aatTTaaagt	gtgagaattg	aaattaggaa	9360
gatggtagaa	tgagaagtag	aataatTTgt	TTTTTTTTtt	ttgattgaat	taaatgtaa	9420
gatattaggt	gtagatagat	tggatattag	TTTTtatagt	atTTTTtaag	agtgttatgg	9480
TTTTTTTTtg	TTTTatTTtt	atgTTTTtag	tatTTTTTTTT	agggtTTtag	ttaaagTTTT	9540
gtttaaagtat	ttgaataaat	tagatTTtatt	tttgaattag	ttttaggtgt	tgTTtataga	9600
tattgttatg	gTTtatatTT	ggataatTTta	gTTtggtTTt	tttaatgttt	taatTTTTTT	9660
tttatgattg	tgaaaaaat	gtaatgagtt	TTTTtgaaga	aatTTtaag	atatataaat	9720
ttagaatatt	tatagTTTTt	aaattaagat	tgtattatat	aaaaatattg	atTTTTTTaa	9780
aatgtatgaa	tgagtgaatg	agtgtataaa	tagTTaatag	tgtataaatg	tatgtatgaa	9840
tgaatgaatg	agagtataaa	taagtaaata	aggaaatata	TTTTtaaagt	atatgttagg	9900
aaatggataa	attatgtTTt	atgaagTTTT	agattgaatt	aggtaaatag	aattgtagat	9960
TTTTatgTTt	gTTatgattt	agtgtatagt	tttgggaaaa	TTTTgtaatt	TTTTtatatt	10020
ttagTTTTTT	atTTTTataa	aggagtatat	tgTTTTgaaat	ttaatTTTTt	tatagtagtg	10080
agagTTTTga	tttaatgTTt	ataaaatgTT	ttttagTTTT	tttgattaat	aggtaaaata	10140
ttgtggtagt	ttgtataata	tatTTTTTTaa	aaaggataaa	ggtaattgggt	tgaatTTgat	10200
gtaaaataat	atTTTTatTTa	agaaataaat	attagTTtata	tttattatTT	TTTTatTTta	10260
tttgTTaagt	aataaaaaatt	atTTTTtaata	gaaaagtatt	aaattaatat	ttaaagttagt	10320
tgaagtgtta	ttattatag	TTTTatTTat	gttgggggaa	ggagagttgt	tttgTTTTTT	10380
taggttattt	tatatgTTta	tatggtaaag	ttaggaatgt	tgggtatagt	tttgaagggt	10440
tttagtgtat	ttattggTTt	ttgtatTTTT	aatgtatTTt	atTTTTtata	ttataaaagt	10500
atTTTTgagt	gTTTTtgaa	atatataata	aatTTTTtaat	taatTTaaga	gtgTTTTaat	10560
tgtatataatg	aaatTTTTtg	ggaggaaaaat	attagTTtatt	ttatataat	atagatTTta	10620
agttagtaag	agagTTTTTT	ttgaaaggta	taaagtagtt	TTTTtatagt	aatataatTT	10680

-continued

ttagttttat atgtgtaagt tttattgaat tttttagaat aataggaat tttgtaggaa	10740
aaaaaatttt ataaaaattg atagttatat ttatttaatg taagaggaat atattagtta	10800
ttttgagttg gattatTTTT taaaggtatt agataattat tgaagattaa ttaaaatttt	10860
aattaaaaat tataagaatt gaatattttg gagatagaat tttttatata taaaaattga	10920
ataataaaat taattgatta atatatTTTA ttttaaaatt gtatttatta gattatTTTA	10980
gtagtattt gagtagtaat tttttggatt ttgatgtatt aatttttatt tttatattaa	11040
atagattata ataagtaata gatttataaa ttagttattg atatgatata attgatattg	11100
tgatattgat aataatattt attaatatat ttataatgga gtaatgtat ttattattta	11160
tataatgttt atatttttta tttttatggt aaaaagtta tgtaatatat gtaagggtt	11220
tagaattgtg tttggtatag agtaggtttt tattagaata ataatatatt attttttata	11280
agtttttatt aattatata gagttagaat ttttttttt aataattttg gtaaatattt	11340
gatggtatta agtaaaaaa aatgtagtga atataatatt attgatttat ttttttttaa	11400
tatttgggtt tatgaattat aagtaaatat aatattgtatt tataaaaatg aatattaaat	11460
atttttaagt gagaaattaa ttatataaat tatattgata tattaaatat atattgttaa	11520
tagaattatt tttttaatt atagtgaggt attttgatta ttatatgagt ttttaaggaat	11580
ttattattat aattatTTTT ttttaattaga gtataaataa ttataagggt taattaaggt	11640
gaaatttttg ttttgggttt ttgtgagaaa gatgagaata ttattttaaa gatgtaagat	11700
ttataaattt ttttaaaaag ttttttggtt ttttttatgg aaaatgtata tatatattta	11760
tttattaata ttgtatttag agttttattt atagatattt ttggtagatg tatgaagaga	11820
tataaaaatt attagtgtaa tatatttagt aatagtaaaa atatggaagt aatgtaata	11880
tatatataaga ggggaatagt gaaattatga aatatttga tatattaaat agaagtattt	11940
aaaattatat gtattgatgt ggaaaaatat ttaagattta gtagattaaa agtaagtttt	12000
aaggaaatta atgattttat tggagttaaa aatgtaaaat gaaaaataat tatattttta	12060
tatagttatt atatttatat atatatagaa aaaaaaatgt ttgaaaagat gtatatttaa	12120
gtataataat ttattttgat ttgtagagaa taagaatgag taaatttatt ttttatttta	12180
aatatatttg tttttttaa attttttaag aaagtataat atatgtatag taaaataaaa	12240
tataaaaatt aaaagttgat tattgatgaa tttttattga ataaatattt ttatatgatt	12300
agtaattaaa ttaagataaa agatataatt agtattttta agtttttttt ttgttttttt	12360
tggttatttt ttttaaaaag aatttttatt atttttggtt taatatgtta tataaagtat	12420
gatttatttg ttttttttaa gtaatagatt aaaaaaaaa atttgtttta gtagtatgaa	12480
tttttttttt tgaatattaa atatatttag tttttattt taagatatta atttaatta	12540
taatttgagt ttttatagtg ggttaggtat tgaattaaat gttttttttg tatgaattta	12600
attatttttt ataattttat gaggtgagta ttgttattat tttttatgta tatattgttg	12660
aattatgggt tatagttatt tagaaaatta tatgattaaa gttatattgt tgatgagtg	12720
tagagtttgg atatatataa aggtagtttg attttagtat ttttaatttt ttattagaaa	12780
taaataaata aaaaataaag ggtaagatgt agagttggga gtatgaagtg ttttttagtat	12840
gaagtttttt ttttttggtt tgttttgagt aaatgattta aggttttaag gttatataat	12900
ataatatttt atttgtaaaa tggagggtt aggttagatt ttttttttga tttgattatt	12960

-continued

tgatggtgat	aaataaagtg	atattattgaa	atatagagta	aatatgtttt	gttatgatgt	13020
agtgtaatta	tgtattgtat	tatgttaaaa	aatgattttat	ataagaaaat	aaaattataa	13080
tttttgaaat	tattggaaat	aataatttta	atttaaatggt	atattttgat	ataggaagat	13140
tttttttttt	aaataagtaa	attgggtggtt	tttgaagggt	tttgtttatt	tttgtatgga	13200
aaatgtttaa	aaataaatta	ttaaaaatta	aataaaattt	ttttattttt	ttaattgtgt	13260
ttgaaaaaaa	agatgtaatt	tttttagta	aaaggttatg	tatatatttt	atagatatata	13320
ttgggatttt	tggtttttat	atattgaaat	attgatataa	ataaatatag	aatatgtttt	13380
tttttaaatg	aaatttataa	ttttgagggt	taaaaatttt	ttaaaataat	atgaaattga	13440
tttttagtag	aggtaattaa	agagattaag	gtttgttttt	ggagaaaatt	agaattatga	13500
agattataaa	ttttaatgag	aaatttgttt	tttaaatggt	agtttattga	aaatattaat	13560
gtatggtgga	tatatttggt	tttttattag	aaaattagta	atattttggt	ataaatgagg	13620
ttttaaataa	gttgagtag	aaaataatta	ataaattagt	ttttatgatt	ataagtaaaa	13680
gagaattagt	tttttaaaaa	gttaatagga	tttatttttt	gtagatgttt	ggtttgggat	13740
ttaatgtttt	ttaaagatat	gtatggaata	gtgtagttat	ttaattttta	ttaatatattt	13800
taggtgttga	ggttatggag	ggaaaaagat	atgatttttt	ttttttgtga	gtttatagtt	13860
tttttagaag	atagatatata	ataatgagtt	ttaatataat	ttgttgttat	attaatagtt	13920
taaaaaaagt	gttatggtat	tatatagaga	aaagaattaa	ttttgtttga	gttttaagga	13980
aggttttata	gaggaggtga	tatttgagtt	aagttttgaa	agatgagtag	attattatta	14040
ttattattat	tttgtttagg	taggtagtaa	aagagtatta	ggtaagtagg	ttaagggagt	14100
tggggttatg	gaggataaag	gtggtgagaa	tatgttaggg	agaaaaaagg	aagtgattga	14160
tggatttata	agaaataaaa	tgtttataag	taagttttat	tatattagga	tttattatat	14220
aaagaattaa	tttgagaaa	gagatatgtg	atttgaatat	tttggtgggt	tttttgattt	14280
atgaatttag	gaagtgttta	tttgaaaaa	tagaattttt	gggtttttgg	ttgtttttta	14340
gtaattaaag	ggttattagg	ttttttgttt	aatatttggt	ataaagtgtt	tggatggatt	14400
tgtggtttgt	tgtgtttttt	tggatttttt	tgtttatgaa	tgaagtagaa	ttttttttta	14460
ttatttgtag	tttaaatatt	tgttttaa	tttttgta	atatttttag	aatggttttt	14520
atattgtatt	gggttgttta	gtatttattt	tatttttttg	ttaataagga	gttatgtagg	14580
gttttttttt	atattagaaa	tagaaaaata	ttttggttaa	atatagtatg	ttgtattttgt	14640
agttttttgt	tatttaagta	tatatattta	aattaaagta	taaattaaag	gaagtgagag	14700
agagagagtg	agtggtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtt	14760
tttaagtatt	tttattttgt	ttatttaggt	attaattagt	aatgatgtg	aaatttttag	14820
gttagtttga	ttgtttttat	gattattttat	ttaatataat	taaagtataa	ttattaaatt	14880
agataatgta	taagtgtgga	attttaagta	tttttgta	ttttaataag	ataaaaaatta	14940
tttattgttt	agaattttta	agaaaattag	tttaattata	tatttataat	attaatgttt	15000
tatatatttag	aataattttg	aaatttaatt	taaaaattat	atgggaaaag	ttattttttt	15060
ttattttatg	tgaatattat	tgtttttatt	ttgtttaagt	gattttttgaa	atgtgtttta	15120
tatatgttta	aattatgtaa	gggtggtgtt	ttaataatag	ttaaatattat	ttatattaat	15180
tatatattat	tttatagata	aatgtattg	gttaatatta	aatttataga	atgagaattt	15240

-continued

aaaagtttg atttgagaat ttataaaaga gagtttttat atgtaataata tttattttta	15300
ttttttttta gttgtataaa tgttattaat ttaagatfff gaatttggtg ttataattga	15360
aagttatata taatfttttt gaaaatfttt tattataaat aaataattta atttattaag	15420
tattttaatt tttagtagtt gtgattatta gtaatftttat tttttttaat attatagatt	15480
aaataaaata aaagttfttt attattgaat atttatttat agaaatggaa aggtaatggt	15540
tattatttaa aaataaattt aaattattta ttgataaata attggaagaa aagatatata	15600
gatgatttaa ataaatattt atftttttata tftttftttt taggtatttt gttatftttta	15660
tttatatatt atfttttagaa aaggaaaata tgttgagat ggftaatttg tgtgaagtgt	15720
atttgtaaat tftttttgtg taaagftaat ttattaatat agattaatga gaattgfttg	15780
atftttatgft taaatttggt aggaaggtat taggtatttt tatatttgat tggaaataga	15840
ataatftttt tatgtagagt tttgggagtt ttttgattat attaatgat attaagattt	15900
gfttaggtta tgggftttta agaagaggta atftttgttt tagagattat ttgataatgt	15960
ttggagatat tattattatt aggggaggtta gggftttttg tatttagtga tgttattaag	16020
tattttatta ggtataggt agtatftttt gaataaatta tttagtttaa aatgftaata	16080
attagagttt gaaaaatftt tatttaaata tattgaagaa tttaggtata atgttattat	16140
ttttaaaatt atatagatat atfttgtagt agfttttagat aagttgtaaa aatgagattt	16200
tgataattaa tattttaatt tftttftttt ttttagtttt atftttgttat atagttatta	16260
tatttgaatt taagtataaa atftttgtaag tagaggtatt aattatfttt tattgtfttt	16320
ttaaaggtaa gaattatagt aatatgtaaa gttatftttt aaaaagtta tattgagttt	16380
attaggttta ttgattattg aaaattatat aaatftttat aattggtatt ttatggttga	16440
aaaaaaagta attgaggaaa tttaaatatt ggttataaag tattattatt atataatgtg	16500
tttataattg tttaaagtag ttaaggaatt atatagatta aagftttaaa ataattaga	16560
ttattgtatg tggaaatatt tattaaaaga aatattattg tataatgtaa atggftaatg	16620
ttaattgttg agtatfttata gtgtfttttt ttaaggtttt aaataaatt ttagatttgt	16680
taagatgatt tatagtattt aatttatatt tggattattg gatattttat ttatggttat	16740
taatttatag tfttttaaaa aagatftttta gttttaaata tftttggaat atttgggggt	16800
aatttaatga aagatattta attatagtaa taataattga tatttatatt aagtattatt	16860
ttaaataat ttatatata tttattattg aatftttatt agaatfttat gaggtatgft	16920
ttgttatata tgagaaattg aagtaaaatt tgttttagta atgtgfttga tgttataaat	16980
ttattaatgt tagagttgag gtttgaattt aagtgtttgg tttataattt tftttftttt	17040
ttatttatftt atftttfttt gagatggagt tttatftttgt tgtttaggtt ggagtgtaat	17100
ggtgtaattt tagtttattg taatftttgt tftttgggtt taagtattt tfttgtttta	17160
gftttttgag tagttgggat tataggtgftt tattattatg tttagftaat tfttgtattt	17220
ttagtagaga tggggttttg ttatftttgt taggttggft ttgaatfttt gatftttaggt	17280
gatttattttg tftttgfttt ttaaagtgtt gggattataa gtgtgagftta ttgtgtttag	17340
tftttatatt tfttttatftt atftttattgt ttataagaat tgagfttttt agatftttaag	17400
tattatgaaa ttattgtatt tttagagatt tttagattt ttggftttatt tattagatta	17460
ttgattftttt agtggtaaga aataaatata tfttttatatt tatgtatata tatatataat	17520

-continued

ttaagtttgg atttatagtt gtaattttat tgggtttata ttaaattttt ttgaggttgt	17580
tattgaggta gttttatttt gtaatagtta ttgttttttt tattgttttt agttttaagg	17640
aggggatggt tagtggtata atttttgttt ttattttttag tttttttatt tttttattta	17700
atgggaatat aataattatt tattgattat atgttttagta gtgtattgtg tattagtgat	17760
atagaaatga taaatatagt taattttttt agtaaatttt tagtttaag taggatttag	17820
attaatgtaa tttattatat tttaatgtgg taagggttat gttaatgta gaagttggat	17880
gttttgggag taataagaaa gaaatgttta atttagattt tggaggttag aggaatttat	17940
tagaggagtt aatattttgt attttaagta ttgaaaaagt agtagaaaag atattttag	18000
ggggagaggg gtgatgtttt tttagagaaa agtattttat gtataaagaa atgaagtga	18060
aatgtttaag gtgtatttga gaattgtagt tttttatata gttggaaaag agagttgttg	18120
tgagtgggga gtgagagatg agaagagtaa tttgagtaag gtttttagta tgaagggtta	18180
tttttttgt agagaatttt ggaatttgtt ttataggtag taggtaaatt ttttaaggaa	18240
ttttaagttt aggagggttt ttataaaaaat tatgaataaa tegtgtttaa ttttttaggg	18300
atgtgagaga gtgttagaga tagttttaat atagttttaga atgggttttt aggagaggta	18360
ttaagatatt ataattatat ttgttatgat aatattaata tgaataagtg tgtataggat	18420
aattgtttat ttaaagttaa tttgaatatt tgaggattag atttgtgatt ttggttataa	18480
tgattttttt tttaaatatt tttgggaatg gtatgtaaaa atataagaat ataataaag	18540
ttttaagtt tttaaagtta tgtagagtag aaaagaaatt ttttttttt attttaattt	18600
tgtaagtttt tagtaatatt gataaggaaa aagtatttaa taagtattta ttaagtattt	18660
ataattttgt taatttttt tatgaatgaa atttttaaaa attaaagtta ataattaagg	18720
gaaaaaatta gatttaaaat aaaatttttt atgagttaat taagaaaaat aatagtatta	18780
ttattttatt aattaaagtt aatgtaaaaa taataaata aaaaaaaaa taaaataaaa	18840
tagtagaatt gttgtttttg ataggtttat tgaggtttgt atgtttaagt tttttattag	18900
aagtattggt ttagattatt ggaaaaattg taaaaggaaa tttgaatgtg atttatttat	18960
aatgttatta aatttagtaa gaatgatttg agttataatt gtatatgaat atatatttgt	19020
atttaaatgt aagagaaaag aattaggtaa tgtagagatt tgttgatttt ttttaagttat	19080
tatttttttt aatataattga aaatattaag ttagaatat ttaatttttt aattagaata	19140
tttggttttt tttaaatttt ttaatattag ggagttaaaa gtttagagggtg aaatgtaatg	19200
aatagattat tttgttgggt atatgtagag gtttgtgaat tgtatatatg attttgtgta	19260
taatgagttt ttttggtagt tgggtatatt atttgattag ggatagaatt ttaaattttg	19320
tattgttgtt tttgtgagta gaattaaatt taatgttata tgaatgtaaa tttttatgat	19380
ttaattgtag ttttttataa gaaatgtttt ttttttttta aaaagaaagg gggaaaggat	19440
aatagaaata gtagtaataa tttatgtgaa aatttaaagt tatgttaagt taggttattg	19500
ttgtgttttt tttttatagt ttttaagga ggaagatgt ttttaattaat aaagttaatt	19560
ttttaggta taagtttttt gtgttttatt tagattttag attatatata gggaaatattt	19620
tgattagttt taattttttt agtgtttaa tgaattttaa gattggttaa tatggatatt	19680
taagttatag aaggagagat agatatattt ttttgttgat atgttatatt gggagatatt	19740
ttattgtatt tgatagtttg ttttttagtg atgttatatg aatatgttag ggtaggtatt	19800

-continued

ttgaattaat tttttttggtt agaagatttg aatttttttg tagtagtttt tatataattga	19860
agatatttag agagaatfff agaggtgggtg gttattgtga gagggatttg ttggtgagtt	19920
atfttttaggt tatttatatt ttgttttatt tagagtagtg gtatattttt ttattgggta	19980
tattattagt ttttatttaa atgaaaggat ttgtggttga gtatataagt aaataaaaaa	20040
atgaaaatta ttgttttagt ggaaaaatat tataaaattg taaaataggt tattaatagt	20100
aaaagtattt atttatagaa aatftttatat atgtgataat tagataaatt gaatgtattt	20160
ttaaaatatg ggtagtttta ttataaatta aatftgtttt tagaagtta ttatgtattt	20220
tttaattttt taagtftttat tagttggttt aatataaatg tgtagatttg agatttaaaa	20280
ttagtattaa aataataaaa aataaaaaa aagaaaaat aaagttgata atatttaatt	20340
aatftggata aggtttgaag agtaaaggaa gattfttatat gttagttaaa tatttaaaag	20400
atftggttat aagfttttga taattgggaa atattataaa atatagtatt tttftaaatt	20460
tagttatttg ttgttttagt tttatgattg tggttattta ggttatattt ttgtttattg	20520
ttatttaata tgtfttttta gatatttaat tttftttttt tgattttaat atfttgtttg	20580
gftttgtttt aaataatatt gftttatgaag ttataggftt aatataaaat ttaataaatt	20640
ttfttaatat ttagtattt ttatataaaa tataagtatt attaaaataa aatgfttagg	20700
tattatttaa aaatgggtgt atagtattga tggaaagtaag agfttagtat atgtttaagg	20760
ataaggtgt tggftttata ttggggaggt aggattfttg ttttattgtg ttgtatgatt	20820
gggtfttagt tttfttattt gtaaaatagg aataataaaa gtatttaatt tatggtatag	20880
ttaagaagat taaaattagt taattfttat aaatgfttaa aatagtgttt tgtatgaaat	20940
ttfttagtaa atgttagtta tgatgttatg atgggtatag tgattaatat tatagtaata	21000
atgatagtta aattgtaatt atftaaagat tatagtattt atftttattt aaaaatatat	21060
tttaggatat aaatatattt aagaaaaat gftftftatat ttgatattta tattaatata	21120
tgtfttgggtg atataaaagt tattfttat tggfttaatg taattfttatt ttgaaaatat	21180
ttatgtgata ataggaatta gattftgttt aaaaattata atftaaataa attatttgtg	21240
tagaaaaaa aaaattaaa aatfttaata tgatagaata tataggfttt tttfttattg	21300
tttaaaagta tgttaaaaa aatfttatta ttgtatgtaa ataaaagftt taaaaatag	21360
aatatattgt attaaaatat tttatattgg aaaaaatga gaatftttgg tattfttagaa	21420
tgggtftttt ttttattatt tatgtattga aatfttaggg taagaaaagg gttagattaa	21480
ttftgaaagg ttttaaatat taaattatat tattgtgta gtattftaag attaattagt	21540
atftfttaat gataaaaaag gttatagtga gattfttaggg tgtattfttt gaaattgttt	21600
taaattatta tataaatgtg gftttgataa tataaaatgg tttftatttt aaattaatta	21660
gftftatttt tttftttgga ggtgtgaaag gggfttaggg agtaagaata gtagtgttta	21720
ttfttagtat tttfttatta ataaaaagat ggatggagaa atattaagaa tgataaattt	21780
taagttaaa atgatattt attaaagtat atattaatgt atagftttat aagtaatatt	21840
gagtgatagt aatgfttata aagftfttaa atftfttaaa agtaattaa gttgfttga	21900
ttfttttga atgtgtaatt tttgtttatt tttftttatt taagfttaag ttaaatgatg	21960
ttaagfttg ttttaaaaat gatttaatga ttatgtagt tttfttttat taaaatatat	22020
ttaattataa gatgtagtat tatatatatt gtattftatt tattfttagg gtaaatgttt	22080

-continued

gaaatagtat gttatTTTTT agtttagtag gaaaaattta tgagagaagt aggtatattg	22140
taatgttatt ttatTTTTT ggtgttgatt tttttattt agtaattatt gtaataaaat	22200
tttttaattg tatttttgaa tataaattagt ttaagatatt ttatagggag gtatgttata	22260
agagtttagt atttagtata tgtaagaata atatttattt agaaaaagaa attgttaaga	22320
ttatattttt attgaaaaa gaagaagggtg ttttgataga ggtttgttta gttttttaat	22380
gagaggattt aaaatgtttt ttatTTTTg tagtatattt ttttatttat tttttaaata	22440
agtgatttga aatatttata attagatagt atttaatttt aggttttaag tgttttaata	22500
gttttatgaa atatgattag tagttgtaat atagaattat ggtgaataaa tttttataa	22560
agaggggtt gtttttatag aatgtttagt gtttttgat ttatttaata gaaaaaatg	22620
ataatgagaa agtataatat atatttgtat ttttaaatTA ttttaatttt aattaagtag	22680
tgaaaaagat tattatTTTT ttttttgaaa tttgtataaa aattgtgagt ttttgtggg	22740
taaaaaagt ttgattattt gaaagttaga atgtgaattt ttttaatttt taatttttta	22800
taataatagt tattgttttt ttgttatttt atttttttaa aaattttaaa aaatttaatt	22860
ttagtgtata aaatgttaaa atgaaatatt aagttatata aatatatttt tagatatgat	22920
attagttatt tagagtttgt aatagagatt gaaaatatga aattggatat tttatTTTT	22980
gtggttatgt aaaatttatt tatattttaa agtataatat aaataaaaag gaaattttta	23040
gaaatggagt atttgttttt ttttttttt ttttaattta aagagtttta ttaaaagaaa	23100
aaaataaatt atagtaaaat gataagtaat ttatttgta gttatttgtt aaaaaagaaa	23160
tttttatgta attttttttt ttagttttta gaaatttttt gaaatttagt atttatTTTT	23220
tttaataatt aataattatt attttttttt atataattat ttttttttat ataaatttta	23280
agattttaaa aaattttaaa ttttaagttta ttatttaaaa gaagtgttaa ttgtgagtat	23340
atatataat gtgagatggt ttatTTTgga aatgagaaat ttaattttaa atatttaata	23400
aatattaata aaaagaaaat aattgagtaa attttatttt agttataatt gtgaattaat	23460
tggtagtat aatggttttt aaatattgaa atatattatt tttaggaaat ttataaagga	23520
aaatagggtt ttgaaaatta tattgattgt agatttttat ttattatgaa agataagtat	23580
atattattaa taattagaga aaggagatat ttttaataaa ttttaattatg ttgtttttaa	23640
agttgtagtt gttaaagggt ttaatatttt tgtatagtta tttttaaata tgggtttaat	23700
tttgaatgat ttataatttt taattagggt gaataaatta tttttatggt ttataataat	23760
aaatgtgggt gtagtattgt ttgtttttga tattgttgtt ataaaaatta gattttggtta	23820
gttttgtttt atataaaata atttgttttt tgttattatt gttttttgtt ttttgggata	23880
atgttatttt taaaaata taaatgtttt tgaatgtagt tttttaaaat gttttgtatt	23940
atgtattaaa tattttttgg ggggattttt tttttatttt ttttagtttt tggtttttta	24000
ggtttaagtt tgtaataaat attttaaaag ataaaaaaa aaaaaagggt attgattttg	24060
tttaaaatta aattgtattt gggaaaggaa ttttaagtata attttttttt taaggtttta	24120
aaaataaaaa taatgtaata aataaaaaaa ttattgttaa tttttaattt tgttagttat	24180
ttatgtgtaa gtgaaatttt agaagttggt gatggggagt ttgtggttgt taaatgtttt	24240
taagattttt atttaaattt tgtaggaggt tgtattgaaa ttgataaata tttgaggggt	24300
tataaagtgg ttatagatga gaagaggaat aaatgttat attttaggag gtgtattttg	24360

-continued

ttttaataat	tatggtgatt	ttaggggtgt	aagttaaaag	gggtaaggga	gaaagaaaat	24420
tgттааgтat	gagtttttaa	tgtgaaattt	aatttggttg	tttttttagg	tattagattt	24480
tggaagtgtt	tgtaagtata	gttaggaaag	taaattaatg	gttaatgggt	tattgggttt	24540
gaatattaga	gttttgttat	ttttttatag	attaagtata	taatatatat	atatatatat	24600
atatatatat	atgaagttgt	tttatgtgga	attgtttttg	aattaaata	tatgtttggg	24660
tgttattttg	tattttggta	gaaagggtag	agggatttgt	attttttttt	tttgttttta	24720
ggttaagatt	tatgaatgat	tttttagggt	tgtttgggat	gggggggaga	ggtgtaggag	24780
atgtggtgag	agaggatag	ttttatatta	tttttttgta	tatattttta	ttttagggtg	24840
aaatttggtg	ttttatag	aaggagaaag	aggggtggtg	gtgaagggaa	atattatttt	24900
tttaagtagg	gaagagatga	taaaaattaa	attgtgtttt	ttgtttttgt	tttagttttg	24960
ttattggatt	tggtttgtaa	attttttaat	tgttttgagt	tgtttagaat	tttgtaaata	25020
aagggagaaa	atttttttaa	gtgttattaa	atgaaagttt	gggttggtgt	gtggtgtaga	25080
ataaaataag	tatatatatg	tttattttta	aattttgttt	tattatgtta	aaaatttgag	25140
ttattttaag	ttttattgtt	tttttagatt	tagtaggttt	ggaaggagag	aaaaatttgt	25200
tttatttagt	taggtttttt	ttttggaga	agttataagt	tttttatggt	attgtttatt	25260
tgagttaatg	aatgtgattg	ttttgataag	gtgtttttaa	ttttattgat	tttggtaaat	25320
tgtaggagat	atttagttta	tgttatataa	aaattattaa	agtgggattg	ttgtagttga	25380
aaagtttgat	tagtgttttg	agattagaat	gtgagttttt	aaagtatata	tgtatttggg	25440
attgggaaag	tttaggtgtt	tgtgattttt	tgaaggattt	tgggttagta	tttttgtgtt	25500
tggattattt	aggggtttat	tagaaaagtaa	tatatgttat	aagaaagatg	tgggggagat	25560
tttgatgttt	gtgtttttgt	taggtggagg	tgggtgtag	gtagaagtgg	tgtgttggat	25620
tgttggatga	attttgttat	gtgtaggttt	tgtgttgatt	ttggaaggga	taggtaggtt	25680
tggaagggat	ttttggggtt	tggggatggt	ttttagggta	gagagtattg	ttgtggtttg	25740
agtggtttgg	gtttttaggt	gtgggttatt	tggtttttaa	gtgttttttg	gggatgtagg	25800
tgtttatgta	ttgtttagtt	gtgtgttttt	agtatttagt	gtttgggtgt	tttgggtggag	25860
ttttgtgtat	ttttttggaa	agtttggggg	tttttttaat	ttaggttttt	ttgggagttt	25920
tgggtgggg	gttggagaaa	gggttatttt	attttgttgt	ttggtttgtt	ttgggttttt	25980
gaagtgtttg	ttttagttgt	ggttttaggt	gtttgtaatt	aagtatttat	ttttgttttt	26040
tttagtgttg	aagttttttt	tttgtgtgtt	gtttttggtg	aaaataggat	attttttttg	26100
gttattttat	aataaatagt	gtatataagt	gggggagaag	tgggggtggag	ggagaagttg	26160
gttttgaggg	ggaggaggaa	aggagaggag	ttaaaatttt	ggattgtgat	aggggggaaa	26220
ggagaagaaa	agaaaatgag	agagtgtgtt	tagttgttgt	agttgttatt	ttttgttttt	26280
ttttgtagtt	tttttttttt	ttttaaggta	gtgataattt	tatgtggata	ggatggaagt	26340
tttggtgga	agttgggaat	ggttgaggtg	tgtgttgtgt	tgttggtagt	tttgatattg	26400
tgtagggagg	tgggtggggg	gtgatgttgt	tattggttat	tgtgggtttg	tgaggttgtt	26460
gggggttttg	ggggaggttg	ttagggatgt	gggggtggtg	ggtgtgtggg	ggatttattt	26520
tgtttttatt	tgattttggg	tgattgttta	ttgttatggt	ttgtgatttt	gtgtggggat	26580
tgtttagggt	tttgggtttt	tggatgtttt	gggtggtgtt	gttgaatttt	ttgtgttttt	26640

-continued

tgtgttgttt ttattttggt tttatttggg ttgatttgag tggttgtagt ttttaggtt	26700
aggggagggg tttgtgattg tgtttttttt tttgggttgg agagtattg ttgattttta	26760
tttgtttttg tggggtagtt tttggaattg tttggaattt ttttgttatt tagttttttt	26820
gtgtagggta tttgtgggtt tttgttgat ttttttttat ttattttaag ggaggagata	26880
gaagggattt tgtttatttg taataatgtg aaataaaaat taatttagat tagattgggg	26940
tgtttttttt gatgggagga aatatttttt tttgtggaga tataatgtat gaaggattaa	27000
gtggtttggg gataggttgt gttgtgaagg gtttgggtt ttttagtgt tgttgggtgt	27060
aggtataggt atttgtttta gatgtttaaa ggttaggtag taatgattta tttttaaggt	27120
ggtgttagag ttttttagg tataatgttg aaatgatatt ttgtgttta gttttttgtg	27180
tatgtattat gtgattatat aggatttatg attttatatt tataatgttg tatgttataa	27240
ttaggggata tagtaaaggt agatggaata aataatttat aaattttgta tttattta	27300
tatttttttt aaaaatttat tgtttatttt agtggatttt ttttggtat attgataata	27360
aatgttttaa tttattgata aattattgag tgttaattat tgtattggtt atttatggta	27420
tatgagttaa ttttaatttt atgataattt tatggtggtt attattaagt tggataaaaa	27480
gatgagaaa ttaaagttta aagaattatg agtgattttt ttaaagataa atttagtggt	27540
ttttaattat aaagtttatg atataataat ttttatatta ttttaggtgt ggtttaagag	27600
attatagatt aatttgtaga aattattttt atttaagaa agtttagttt taatatattt	27660
attaagttat gttttgaaa aggtttagaa aaaagatttt attttttatt ttaggttgg	27720
tttatgtaat tgatatgagt attgtttaat ttaaaagtag tattaatttt gaatatttat	27780
ggttataatt tattttttta tgagtggga attatttata atgtttaaga ggttttagaa	27840
aaggtggtgt tgtaaatata aaatgataaa ggtagtgtt gttgttttta ttatttattg	27900
ggtgtttttt tttgtagatt tattttggag ggtgtaggta ttgttaggta taatgtttt	27960
ttttttgggt ttttaaggtg aaagggtttt ggtagtggg ttggtataaa gtggatttg	28020
agtatggtg agagtatatt ttggtttaat gaggaatgtt agttaatata tattggagag	28080
aaaaatga atggatagat ttaattaatt ttgtagattt atttttttta ttttatatgt	28140
gagaaaatta aggttttaga attttttaga gttgtttaat aaatggtaga tttgagttt	28200
aagttagttt ttatttaagt ttatttttag ttattgtgat gttattatta gtattaatgg	28260
tatttagaaa aatagtaatt tttttaaata tgtaatatata tataagttaa atagtaatt	28320
taaatgtttt gtgaattata gtgtttaaga atgttttggt gttatagtaa atattatgta	28380
aggattaagt ataagtaga atttaaatatt gtttttaaaa ggtaataata ttttgaattt	28440
tattttggta taaaatata tgtgttagtt ttatttataa ataaaggtgg gagatgtatg	28500
atttgaataa aataaaaatt tatatttttag tatattgta atataaatat aggggtgggt	28560
tttttagtat gttttttgt atttttttaa agttagattt tttagttata ttgattttgt	28620
ttttaaggtt atttgtttt tgtatagttt tttgtatatt tagataagaa tatatgtata	28680
gaaattattt ttaaaaatta ggataatata ttagttattt gattttatat tattttttt	28740
taatagttgt tttaaatata tatataatata tatggatatt attaggtgtt agataaaatg	28800
ttaattttta gtatatattt gaaataagtt taatatagga aatagaattt tagggttata	28860
aaggtaaaaa tgtattgttt gtttaaatga gtttttattt ttagtattta tattgaatat	28920

-continued

```

tatttttata attatagtta tattgtttat attattgatt taatatgtat ttttaaataa 28980
tttattttta ttaaataagt ttatttgata aaggatattt tttatagtta ttataaatgg 29040
aaaattaata aaatttgta tgaatagaag atgattttgt aaaggaattt atattgtaga 29100
aataaaagtt attttttatg tattatttaa aattatttta tatatagttt gagaaatatt 29160
ggtttaaaga attatatttt atatgttttt taggtttatt aagagatatt ttttaagttt 29220
ttttttttat aattggaata aaatgttggt ttttaattat tttgaggtat tattttttta 29280
attttgtggt tatttaatta ttgttttttt tgtttaaaat atatttgaag tttttgaagt 29340
tgagggtgtt ttttaaaatt tatgtattat ttattgtttt atttattatt ttatttgtgg 29400
taaaatgtaa ataatttata aatgtataat taaaaagtaa attatttttt tgaatattg 29460
tattggttat ttattatttt tttttagtat ttattagttt agtagtgtaa tatagtatag 29520
ataatgaaga agagagggaa agtagaatag tgggagaaat gtaaagggtt tagaaagata 29580
tatgggggaa attgaagtta gagatttgtt gttagatttt tgaagtgtt taaaaatatt 29640
agtgggtttt gtatgatatg agttaggttt ttagaatatt tgtatatgtt taatttagtt 29700
tgaaaagggg tgtgttttgg gaggagtggg gagataaaaa atataagatt ttatagggta 29760
aattagtgga aatataagat ttatttatta gtaatttttt taattaattt ttttggatg 29820
ggatgttttt gtatgtatta tttattattt gttttttttt taaatgtgta tggatttgtt 29880
ttggtattag attttgtata agttaaatta aataaagtgt tttaggagtt tataatttaa 29940
aatataattt ttgttttaat tttattttaa agtaatttta agggaaaaat tag 29993

```

```

<210> SEQ ID NO 6
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PRIMER

```

```

<400> SEQUENCE: 6

```

```

cggaggggtac ggagattacg 20

```

```

<210> SEQ ID NO 7
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PRIMER

```

```

<400> SEQUENCE: 7

```

```

cgacgacgcg cgaaa 15

```

```

<210> SEQ ID NO 8
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PRIMER

```

```

<400> SEQUENCE: 8

```

```

tggatgatgga ggaggtttag taagt 25

```

```

<210> SEQ ID NO 9
<211> LENGTH: 25
<212> TYPE: DNA

```

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PRIMER

<400> SEQUENCE: 9

ggtgattggt tattgttatg gtttg 25

<210> SEQ ID NO 10
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PRIMER

<400> SEQUENCE: 10

ccccctcaacc taaaaactac aac 23

<210> SEQ ID NO 11
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PROBE

<400> SEQUENCE: 11

cgaaacccta aatatcccga ataacgccg 29

<210> SEQ ID NO 12
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PROBE

<400> SEQUENCE: 12

accaccaccc aacacacaat acaaacaca 30

<210> SEQ ID NO 13
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PROBE

<400> SEQUENCE: 13

aaaattacga cgacgccacc cgaaa 25

<210> SEQ ID NO 14
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: BLOCKER

<400> SEQUENCE: 14

gttatggttt gtgattttgt gtggg 25

<210> SEQ ID NO 15
<211> LENGTH: 26

-continued

<212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: BLOCKER

<400> SEQUENCE: 15

aaactacaac cactcaaatc aaccaca

26

1. A method of diagnosing a colon cell proliferative disorder in a subject, comprising the steps of:

- (a) obtaining one or more test samples from colon tissue or serum or both of said subject; and
- (b) detecting a decrease in the amount or expression of a polypeptide expressed from the EYA4 gene

2. The method of claim 1, wherein said colon cell proliferative disorders are taken from the group comprising adenocarcinomas, squamous cell cancers, carcinoid tumours, sarcomas, and lymphomas.

3. The method of claim 1 or 2, wherein said detection is by immunoassay, in particular by an ELISA.

4. The method of claim 3, wherein said immunoassay is a radioimmunoassay.

5. A method of diagnosing a colon cell proliferative disorder in a subject, comprising the steps of:

- a) obtaining one or more test samples from colon tissue or serum or both of said subject; and
- b) contacting said sample with an antibody immunoreactive with the EYA4 polypeptide to form an immunocomplex;
- c) detecting said immunocomplex;
- d) comparing the quantity of said immunocomplex to the quantity of immunocomplex formed under identical conditions with the same antibody and a control sample from one or more subjects known not to have colon cancer; wherein a decrease in quantity of said immunocomplex in the sample from said subject relative to said control sample is indicative of colon cancer.

6. The method of claim 5, wherein said immunocomplex is detected in a Western blot assay.

7. The method of claim 5, wherein said immunocomplex is detected in an ELISA.

8. The method of claim 1, wherein said detection is by expression analysis.

9. The method of claim 8, comprising detecting the presence or absence of mRNA encoding a EYA4 polypeptide in a sample from a patient, wherein a decreased concentration of said mRNA above the concentration for a healthy individual indicates the presence of colon cell proliferative disorder cells.

10. The method of claim 8, comprising the steps of:

- (a) providing a polynucleotide probe which specifically hybridises or is identical to a polynucleotide consisting of SEQ ID NO: 1,
- (b) incubating said sample with said polynucleotide probe under high stringency conditions to form a specific hybridisation complex between an mRNA and said probe;
- (c) detecting said hybridisation complex.

11. The method according to claim 10 wherein the detecting step further comprises the steps of:

- a) producing a cDNA from mRNA in the sample;
- b) providing two oligonucleotides which specifically hybridise to regions flanking a segment of the cDNA;
- c) performing a polymerase chain reaction on the cDNA of step a) using the oligonucleotides of step b) as primers to amplify the cDNA segment; and
- d) detecting the amplified cDNA segment.

12. A method for repressing transformation in a colon cell, the method comprising contacting said cell with a EYA4 polypeptide in an amount effective to inhibit a transformed phenotype.

13. The method of claim 12, wherein said EYA4 polypeptide is introduced into said cell by the direct introduction of said EYA4 polypeptide.

14. The method of claim 12, wherein said EYA4 polypeptide is introduced into the cell through the introduction of a EYA4-encoding polynucleotide.

15. The method of claim 14, wherein said EYA4-encoding polynucleotide further comprises control sequences operatively linked to said EYA4 encoding polynucleotide.

16. The method of claim 14, wherein said EYA4-encoding polynucleotide is present in a vector.

17. Use of the vector according to claim 16 for therapy of colon cell proliferative disorders.

18. A method for preventing or treating a colon cell proliferative disorder in a subject which comprises administering to the subject a therapeutically effective amount of a compound that agonises EYA4.

19. Use of a polypeptide expressed from the EYA4 gene for detecting, differentiating or distinguishing between colon cell proliferative disorders.

20. Use of a polypeptide expressed from the EYA4 gene for therapy of colon cell proliferative disorders.

21. The method of claim 8, wherein said detection comprises methylation analysis of the gene EYA 4, its promoter and/or regulatory elements, in particular through the methylation analysis of a genomic DNA sequence according to SEQ ID NO: 1.

22. 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 NO: 2 to SEQ ID NO: 5 and sequences complementary thereto.

23. An oligomer, in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer, said oligomer comprising in each case at least one base sequence having a length of at least 9 nucleotides which is complementary to, or hybridises under moderately stringent or stringent con-

ditions to a pretreated genomic DNA according to one of the SEQ ID NO: 2 to SEQ ID NO: 5 and sequences complementary thereto.

24. The oligomer as recited in claim 23; wherein the base sequence includes at least one CpG, TpG or CpA dinucleotide.

25. The oligomer as recited in claim 24; characterised in that the cytosine of the CpG dinucleotide is located approximately in the middle third of the oligomer.

26. A set of oligomers, comprising at least two oligomers according to any of claims 23 to 25.

27. A set of at least two oligonucleotides as recited in claims 23 to 26, which can be used as primer oligonucleotides for the amplification of DNA sequences of one of SEQ ID NO: 2 to SEQ ID NO: 5 and sequences complementary thereto.

28. A set of oligonucleotides as recited in claim 23, characterised in that at least one oligonucleotide is bound to a solid phase.

29. Use of a set of oligomer probes comprising at least ten of the oligomers according to any of claims 24 to 28 for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) within one of the sequences according to SEQ ID NO: 1, and sequences complementary thereto.

30. A method for manufacturing an arrangement of different oligomers (array) fixed to a carrier material for analysing diseases associated with the methylation state of the CpG dinucleotides of one of SEQ ID NO: 1, and sequences complementary thereto wherein at least one oligomer according to any of the claims 23 through 26 and 28 is coupled to a solid phase.

31. An arrangement of different oligomers (array) obtainable according to claim 30.

32. An array of different oligonucleotide- and/or PNA-oligomer sequences as recited in claim 31, characterised in that these are arranged on a plane solid phase in the form of a rectangular or hexagonal lattice.

33. The array as recited in any of the claims 31 or 32, characterised in that the solid phase surface is composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold.

34. A composition of matter comprising the following:

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 NO: 1 to SEQ ID NO: 5 and sequences complementary thereto, and

A buffer comprising at least one of the following substances: 1 to 5 mM Magnesium Chloride, 100-500 μ M dNTP, 0.5-5 units of taq polymerase, an oligomer, in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer, said oligomer comprising in each case 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 NO: 2 to SEQ ID NO: 5 and sequences complementary thereto.

35. Use of the gene EYA 4, its promoter and/or regulatory elements for detecting, differentiating or distinguishing between colon cell proliferative disorders.

36. A method for detecting, differentiating or distinguishing between colon cell proliferative disorders according to claim 21, 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 18 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: 2 to 5, and 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, or an average, or a value reflecting an average methylation state of a plurality of CpG dinucleotide sequences of SEQ ID NO: 1.

37. A method for detecting, differentiating or distinguishing between colon cell proliferative disorders according to claim 21, 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 hybridisation properties;

c) amplifying one or more fragments of the treated DNA such that only DNA originating from colon or colon cell proliferative disorder cells are amplified d) detecting the amplicates or characteristics thereof and thereby deducing on the presence or absence of a colon cell proliferative disorder.

38. The method of one of claims 36 or 37, 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 fluids, serum, plasma, stool, urine, blood, and combinations thereof.

39. The method of one of claims 36 to 38, wherein in step b) treating the genomic DNA, or the fragment thereof, comprises use of a solution selected from the group consisting of bisulfite, hydrogen sulfite, disulfite, and combinations thereof.

40. The method of one of claims 36 to 39, wherein treating in b) is subsequent to embedding the DNA in agarose.

41. The method of one of claims 36 to 40, where treating in b) comprises treating in the presence of at least one of a DNA denaturing reagent or a radical trap reagent.

42. The method of one of claims 36 to 41, 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 amplificate nucleic acid molecule carrying a detectable label; and combinations thereof.

43. The method of claim 42, 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.

44. The method of claim 43, comprising in step d) use of mass spectrometry for detecting the amplificate, or detachable amplificate fragment mass labels.

45. The method of claim 44, wherein the mass spectrometry technique is selected from the group consisting of: matrix assisted laser desorption/ionisation mass spectrometry (MALDI), electron spray mass spectrometry (ESI), and combinations thereof.

46. The method according to one of claims 36 to 45 wherein in step c) of the method 6 or more different fragments are amplified.

47. The method according to one of claims 36 to 46 wherein said amplicates are between 100 and 200 base pairs in length.

48. The method according to one of claims 36 to 47 wherein said amplicates are between 100 and 350 base pairs in length.

49. The method according to one of claims 36 to 48 wherein one or more of said primers comprise sequences taken from the group according to SEQ ID NO: 11 to SEQ ID NO: 15.

50. The method according to one of claims 36 to 48 wherein one or more of said primers comprise one or more CpG, TpG or CpA dinucleotides.

51. The method of claim 50 wherein said primers comprise between two to four CpG, TpG or CpA dinucleotides.

52. The method according to one of claims 50 or 51 wherein said one or more CpG, TpG or CpA dinucleotides are located within the 3' half of the primer.

53. The method according to one of claims 50 to 52 wherein said primers comprise one or more bases which hybridise to positions that were converted in the treatment of step b).

54. The method of claim 53 wherein said bases are located within the 3' half of the primer.

55. The method according to one of claims 50 to 54 wherein said primers do not comprise more than 2 cytosine or guanine bases within the first 5 bases at the 3' end.

56. The method according to one of claims 36 to 55 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.

57. The method according to one of claims 36 to 56, 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: 2 to 5, and complements thereof, and wherein said nucleic acid mol-

ecule or peptide nucleic acid molecule suppresses amplification of the nucleic acid to which it is hybridised.

58. The method according to claim 57 wherein the sequence of said nucleic acid(s) or peptide nucleic acid(s) is selected from the group consisting SEQ ID NO: 6 to SEQ ID NO: 10, and sequences complementary thereto.

59. The method of claim 57, wherein amplification of DNA that was unmethylated prior to treatment of step b) is suppressed.

60. The method of one of claims 57 to 59, 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.

61. The method of one of claims 57 to 60, wherein said nucleic acid molecule or peptide nucleic acid molecule in each case lack a 3' hydroxyl group.

62. The method of one of claims 57 to 61, wherein the amplification enzyme is a polymerase lacking 5'-3' exonuclease activity.

63. The method of one of claims 57 to 62, 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.

64. The method of one of claims 57 to 63, 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.

65. The method of claim 64, 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.

66. The method of one of claims 57 to 64, wherein said oligonucleotide or peptide nucleic acid oligomer hybridises between the binding sites of the forward and reverse primers.

67. The method of one of claims 36 or 37, wherein determining in step d), comprises hybridisation of at least one nucleic acid molecule or peptide nucleic acid (PNA) 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: 2 to 5, and complements thereof.

68. The method of claim 67, wherein at least one such hybridising nucleic acid molecule or peptide nucleic acid molecule is bound to a solid phase.

69. The method of claim 68, 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.

70. The method of one of claims 36 or 37, wherein determining in step d), comprises sequencing of the amplificate.

71. The method of one of claims 36 or 37, 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: 2 to 5, and complements thereof; and extending at least one such hybridised nucleic acid molecule by at least one nucleotide base.

72. The method according to claim 71 wherein the sequence of said nucleic acid(s) or peptide nucleic acid(s) is selected from the group consisting SEQ ID NO: 11 to SEQ ID NO: 15, and sequences complementary thereto.

73. The method according to claim 67 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.

74. The method according to claim 73 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.

75. The method of one of claims 36 or 37, wherein the background DNA concentration is at between 100 to 1000 fold excess of the concentration of the DNA to be investigated.

76. A method for detecting a colon cell proliferative disorder according to claim 21, 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 NO:1 or a sequence that hybridises under stringent conditions to SEQ ID NO:1, 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 NO: 1, or an average, or a value reflecting an average methylation state of a plurality of CpG dinucleotide sequences of SEQ ID NO: 1, 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.

77. The method of claim 76, further comprising, prior to determining in step d), amplifying of the digested or undigested genomic DNA.

78. The method of claim 77, 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.

79. The method of claim 78, 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.

80. The method of claim 79, comprising use of mass spectrometry for detecting amplificate, or detachable amplificate fragment mass labels.

81. The method of claim 80, 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.

82. The method of claim 76, wherein the biological sample obtained from the subject is selected from the group consisting of histological slides, biopsies, paraffin-embedded tissue, bodily fluids, urine, serum, plasma, stool, blood, and combinations thereof.

83. A kit useful for detecting, differentiating or distinguishing between colon cell proliferative disorders, comprising:

- a) a bisulfite reagent; and
- b) 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 5, and complements thereof.

84. The kit of claim 83, 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.

85. The use of a nucleic acid according to claim 22, of an oligonucleotide or PNA-oligomer according to one of the claims 23 to 28, of a kit according to claims 83 or 84, of an array according to one of the claims 32 to 33, of a set of oligonucleotides according to one of claims 26 to 28, or a method according to claims 36 to 82, for the classification, differentiation and/or diagnosis of colon cell proliferative disorders or the predisposition to colon cell proliferative disorders.

86. The use of a nucleic acid according to claim 22, of an oligonucleotide or PNA-oligomer according to one of the claims 23 to 28, of a kit according to claims 83 or 84, of an array according to one of the claims 32 to 33, of a set of oligonucleotides according to one of claims 26 to 28, or a method according to claims 36 to 82, for the therapy of colon cell proliferative disorders.

* * * * *

专利名称(译)	用于分析结肠直肠细胞增殖性疾病的方法和核酸		
公开(公告)号	US20050003463A1	公开(公告)日	2005-01-06
申请号	US10/482823	申请日	2003-02-13
[标]申请(专利权)人(译)	ADORJAN PETER BURGER MATTHIAS MAIER SABINE LESCHE RALF COTTRELL SUSAN DE VOS THEO		
申请(专利权)人(译)	ADORJAN PETER BURGER MATTHIAS MAIER SABINE LESCHE RALF COTTRELL SUSAN DE VOS THEO		
当前申请(专利权)人(译)	ADORJAN PETER BURGER MATTHIAS MAIER SABINE LESCHE RALF COTTRELL SUSAN DE VOS THEO		
[标]发明人	ADORJAN PETER BURGER MATTHIAS MAIER SABINE LESCHE RALF COTTRELL SUSAN DE VOS THEO		
发明人	ADORJAN, PETER BURGER, MATTHIAS MAIER, SABINE LESCHE, RALF COTTRELL, SUSAN DE VOS, THEO		
IPC分类号	G01N27/62 A61K48/00 A61P35/00 A61P43/00 C12M1/00 C12N15/09 C12Q1/02 G01N21/78 G01N33/53 G01N33/543 G01N33/574 G01N33/58 G01N37/00 C12Q1/68		
CPC分类号	C12Q1/6886 Y10T436/143333 C12Q2600/154 A61P35/00 A61P43/00		
优先权	2002004551 2002-02-27 EP		
其他公开文献	US8691503		
外部链接	Espacenet USPTO		

摘要(译)

本发明提供了通过分析基因EYA4及其启动子和调节序列来检测，区分或区分结肠细胞增殖性疾病及其治疗的方法和核酸。本发明进一步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。

FIGURE 1

