

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 September 2006 (08.09.2006)

PCT

(10) International Publication Number
WO 2006/094149 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2006/007493

(22) International Filing Date: 1 March 2006 (01.03.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/657,841 1 March 2005 (01.03.2005) US

(71) Applicant (for all designated States except US): **EXACT SCIENCES CORPORATION** [US/US]; 100 Campus Drive, Marlborough, MA 01752 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **SHUBER, Anthony, P.** [US/US]; 6 Parker Road, Mendon, MA 01756 (US).

(74) Agent: **WALLER, Patrick, R., H.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR DETECTING ADENOMA

(57) Abstract: Aspects of the invention relate to methods and compositions for detecting adenomas in biological samples. Aspects of the invention provide panels of genetic markers that can be used to detect adenomas in biological samples with different levels of sensitivity. Embodiments of the invention may be used to screen stool samples for one or more indicia of colorectal adenoma(s).



WO 2006/094149 A2

- 1 -

METHODS AND COMPOSITIONS FOR DETECTING ADENOMA**Related Applications**

5 This application claims priority under 35 USC § 119(e) to US Provisional Application Serial No. 60/657,841 filed March 1, 2005, and entitled "Methods and Compositions for Detecting Adenoma."

Field of the Invention

10 The invention relates to methods and compositions for detecting indicia of cancer in a biological sample.

Background of the Invention

15 Methods for detecting indicia of cancer based on the detection of certain genetic abnormalities in biological samples are known. However, there remains a need in the art for improved methods for detecting certain forms of cancer and for screening patient samples to identify those with indicia of cancer.

Summary of the Invention

20 The invention provides methods and compositions for detecting early signs of cancer or precancer by detecting adenomas. In particular, aspects of the invention are useful for detecting indicia of precancer or early stage cancer (e.g., of the colon) in a subject by detecting one or more genetic abnormalities indicative of an adenoma. Aspects of the invention include interrogating a biological sample (e.g., a stool sample)
25 for the presence of one or more markers that are informative of adenoma. Panels or groups of informative markers may be used as sensitive detection assays to screen patient samples for the presence of adenoma. Aspects of the invention include detecting the presence of an adenoma with a sensitivity of greater than about 60%, greater than about 70%, greater than about 80%, greater than about 90%, or greater than about 95%. A
30 sensitivity of 95% means that 95% of adenomas are detected and 5% of adenomas are not detected. A 95% sensitivity results in a 5% false negatives. Accordingly, aspects of the invention include adenoma detection assays with a false negative percentage of less than about 40%, less than about 30%, less than about 20%, less than about 10%, or less

- 2 -

than about 5%. Accordingly, certain embodiments of the invention include diagnosing the presence or absence of an adenoma in a patient with high level of confidence.

Aspects of the invention include detecting adenoma by interrogating a biological sample for the presence of one or more markers belonging to an informative panel of markers. An informative panel is a panel that contains an informative combination of genetic markers. The biological sample may be a tissue biopsy sample. Alternatively, the biological sample may be a fluid, mucus, solid or other biological product that can contain biological material such as cells and cellular debris. In one embodiment, a colonic adenoma may be detected by interrogating a stool sample. In another embodiment, a colonic adenoma may be detected by interrogating a colon biopsy (e.g., a polyp or other tissue biopsy from the colon). Similarly, adenomas in other tissues (e.g., other gastro-intestinal tissues, pituitary tissue, lung, kidney, liver, or other epithelial, secretory, or glandular tissue, etc.) may be detected by interrogating appropriate biological samples from those tissues.

Aspects of the invention include panels of markers that are informative for adenoma. An informative panel may include a plurality of single base mutations (substitutions, insertions, or deletions) that are associated with adenoma. An informative panel may include one or more regions that contain a cluster of mutations that are associated with adenoma. An informative panel may include one or more CpG regions that are hypermethylated in adenomatous tissue. An informative panel of markers may include a combination of two or more of the above. Embodiments of the invention include panels of markers that are greater than about 60% informative, greater than about 70% informative, greater than about 80% informative, greater than about 90% informative, or greater than about 95% informative.

Accordingly, aspects of the invention include screening patient samples with panels of markers of different levels of sensitivity. In certain embodiments, a patient sample may be screened using a marker panel that is small and not as informative as a larger panel containing more markers. However, the advantage of using a smaller marker panel may be a reduced cost. Reduced cost may be desirable when an assay is being offered to a larger number of patients, particularly when the patients have no signs or risk factors for cancer. For example, a panel of relatively low informativeness (e.g. about 60%) may be used when screening a large number of young individuals who have no risk factors for cancer (e.g., no family or individual history of adenoma, colon cancer,

- 3 -

polyps, or any other tumor/cancer etc.). In contrast, if a patient has a polyp, genetic risk factors, or other indicia of cancer or precancer, it may be more appropriate to use a panel of markers of higher informativeness (e.g., about 80% or higher). For example, biopsy samples taken from a patient (e.g., taken from the colon during a colonoscopy) suspected of being cancerous or precancerous based on physical examination may be interrogated with a highly informative panel.

Aspects of the invention also may include screening subjects repeatedly for the presence of one or more markers of adenoma. For example, a panel of markers may be used as a regular assay for indicia of cancer (e.g., colon cancer). Such an assay may be part of a routine medical exam. For example, an assay (e.g., using one or more of the marker panels disclosed herein) for one or more indicia of adenoma may be performed approximately every six months, approximately once a year, approximately once every two or more years (e.g., every 3, 4, 5, 6, 7, 8, 9, or 10 years). In some embodiments, an assay may be performed as part of a general medical screen or checkup. It should be appreciated that a panel of markers that has a relatively low informativeness may actually become more informative when used repeatedly to test individuals in a population. As adenomas grow and develop, additional markers may appear. Accordingly, an adenoma that was not detected in a first assay may be detected in a subsequent assay if certain markers of the panel that were not initially present in the adenoma subsequently appear as the adenoma develops. In certain embodiments, regular screening may be initiated in individuals that are 40 years old or older (e.g., 50, 60, 70, or older). In some embodiments, the frequency of the assay may be increased for older individuals.

In another aspect, the invention provides groups of nucleic acid probes or primers that are useful for interrogating a biological sample for the presence of each genetic abnormality included in a genetic abnormality panel that is at least 60% informative for adenoma. The nucleic acid primers or probes (e.g., oligonucleotides) may be provided in a kit. The kit may include instructions for interrogating a biological sample for a plurality of genetic markers belonging to an informative genetic panel (e.g., a genetic panel that is at least 60% informative for adenoma).

Description of Sequences

SEQ ID NO:1 – Human APC sequence (GenBank reference NM_000038).

- 4 -

- SEQ ID NO:2 – Human Kras sequence (GenBank reference AF285779).
SEQ ID NO:3 – Human P53 sequence (GenBank reference U94788).
SEQ ID NO:4 – Human B-catenin sequence (GenBank reference AY463360).
SEQ ID NO:5 – DNA sequence encompassing human B-catenin codons 20-51.
5 SEQ ID NO:6 – Human Braf sequence (GenBank reference M95712).
SEQ ID NO:7 – Human PIK3CA sequence (GenBank reference NM_006218).
SEQ ID NO:8 – DNA sequence of exon 9 of human PIK3CA.
SEQ ID NO:9 – DNA sequence of exon 20 of human PIK3CA.
SEQ ID NO:10 – Human genomic Bat-26 sequence.
10 SEQ ID NO:11 – DNA sequence encompassing human APC codons 1286-1513.
SEQ ID NO:12 – DNA sequence of a portion of human Kras Cp2 (HUMRASK02 exon 1).
SEQ ID NO:13 – DNA sequence of a portion of human P53 (HSP53).
SEQ ID NO:14 – DNA sequence of exon 15 of human Braf.

15

Detailed Description of the Invention

The invention provides methods and compositions for detecting the presence of one or more adenomas in a biological sample. According to aspects of the invention, genetic assays involving a combination of different molecular markers that are
20 informative for adenoma can be used to detect adenomas with a high level of confidence.

It is difficult to develop sensitive and cost-effective cancer screens that can detect early stages of cancer in a population with a high level of confidence. Although mutations associated with cancer are known and methods for interrogating biological samples for the presence of one or more mutations are known, it is not economically
25 realistic to screen subjects for all known genetic abnormalities. However, if subjects are screened only for a handful of genetic abnormalities, many cancerous or precancerous lesions may not be detected. Accordingly, sensitive and cost-effective cancer screens require selecting those markers that are sufficiently informative for appropriate types of cancerous or precancerous lesions. Markers or combinations of markers that are
30 informative for early stage cancer or precancer are generally preferred, because patient prognosis is much better when cancer is detected early.

According to aspects of the invention, detecting the presence of an adenoma may be useful for detecting early signs of cancer or precancer. Adenomas are typically

- 5 -

glandular tumors or tumors of glandular origin. Adenomas may be early indicia of cancer, for example colon cancer. Not all adenomas become cancers. However, many cancers (e.g., carcinomas such as colorectal carcinomas) are thought to develop from adenomas. Indeed, a majority of colon cancers are thought to develop from adenomas.

5 Therefore, detecting adenomas is particularly useful for identifying early signs or risks of colorectal cancer (e.g., cancerous and precancerous lesions or growths in the colon).

Adenomas may be invasive adenocarcinomas, significant adenomas, or low potential polyps. Invasive adenocarcinomas may be, for example, adenocarcinomas at different TNM stages (e.g., TNM stages 1, 2, 3, or 4). Significant adenomas may be, for
10 example, carcinomas in-situ/high-grade dysplasias (CIS/HGD) having a diameter of greater than 1 cm, about 1 cm, less than 1 cm, or of unknown size; vilous adenomas having a diameter of greater than 1 cm, about 1 cm, less than 1 cm, or of unknown size; tubulovillous adenomas having a diameter of greater than 1 cm, about 1 cm, less than 1
15 cm, or of unknown size, and low-grade dysplasias (LGD) with a diameter of greater than or equal to 1 cm. Low potential polyps may be, for example, advanced polyps, and adenoma low-grade dysplasias (LGD) with an unknown diameter or a diameter of less than 1 cm. Aspects of the invention may be useful to detect any one or more of these different types of adenomas.

According to aspects of the invention, adenomas can be detected at different
20 positions in the colon and rectum (including the right and left colon and the transverse colon).

Aspects of the invention include panels of markers (e.g., genetic abnormalities such as mutations, including point mutations, deletions, and/or insertions) with different levels of informativeness for adenomas. According to the invention, a panel of markers
25 may include a plurality of different markers, any one of which may be indicative of disease (e.g., the presence of an adenoma) if it is detected. The sensitivity level of a marker panel is related to the percentage of diseased individuals (e.g., individuals with an adenoma) that are positive for at least one marker in the marker panel. Accordingly, only one positive marker out of all the markers tested in a panel may be sufficient to
30 detect an adenoma. However, an adenoma may be associated with the presence of two or more markers from the marker panel. It should be appreciated that different marker panels may have different levels of sensitivity. According to the invention, a marker may be a mutation (e.g., a point mutation, a deletion, an insertion, or other nucleic acid

- 6 -

alteration) relative to a normal sequence at a defined nucleic acid position (e.g., genomic position) or within a defined region (e.g., a defined genomic region). Accordingly, one or more of the markers described herein may be a mutation (e.g., a sequence difference) relative to one or more of the sequences provided in SEQ ID NOs. 1-14 at the specific positions or regions provided herein for the marker(s) in the marker panel(s).

In one embodiment, the following panel may be used to detect adenomas with greater than 60% sensitivity: assays are performed to detect one or more genetic abnormalities from a multiple mutation panel of genetic abnormalities at 22 loci including Kras mutations in codon 12 (K12 position 1, K12 position 2) and codon 13 (K13 position 2); mutations in APC codons 1309 (deletions), 1306 (mutations at position 1), 1312 (mutations at position 1), 1367 (mutations at position 1), 1378 (mutations at position 1), 1379 (mutations at position 1), 1450 (mutations at position 1), 1465 (deletions), 876 (mutations at position 1) and 1554 (insertions); mutations in P53 codons 175 position 2, 245 position 1, 245 position 2, 248 position 1, 248 position 2, 273 position 1, 273 position 2 and 282 position 1; and deletions at the BAT-26 locus. This panel is referred to herein as the V1 panel. Mutations at these loci can be detected using primer extension assays (including single base extension assays and assays designed to detect deletions or insertions in the polyA tract of the BAT-26 locus) or other assays that are useful to detect one or more of these genetic abnormalities (e.g., scanning or base tracking for identifying one or more mutations within a target region rather than assaying for a mutation at one specific position). In certain embodiments, mutant specific hybridization assays may be used. In some embodiments, sequencing assays may be used. In certain embodiments, mutant specific amplification assays may be used. Non-limiting examples of assays that may be used to detect one or more point mutations, deletions, or insertions, include one or more assays disclosed in issued U.S. patents 6,280,947; 6,482,595; 6,503,718; or in U.S. patent publication 20030203382, the disclosures of which are incorporated herein. One or more oligonucleotides may be used to capture, amplify, and/or assay for one or more of these markers (e.g., mutations). Different oligonucleotides may be designed based on the known nucleic acid sequences of Kras, APC, P53, and BAT-26 (for example, oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 2 or 12 may be used for Kras marker analysis; oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 1 or 11 may be used for APC marker analysis; oligonucleotides

- 7 -

having sequences identical to, or complementary to, portions of SEQ ID NO: 3 or 13 may be used for P53 analysis; and oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 10 may be used for Bat-26 analysis). The annealing position and size of the oligonucleotide(s) may be determined in part by the type of assay that is used.

In another embodiment, the following panel may be used to detect adenomas with greater than 60% sensitivity: assays may be performed to detect hypermethylation at one or both of the HMTF locus and the V29 locus (a Vimentin locus). Hypermethylation at these loci can be detected using methylation specific primer analysis (e.g., MSP amplification) or other assays that are useful to detect hypermethylation at one or more of these genetic loci. Non-limiting examples of assays that may be used to detect hypermethylation at one or more loci include one or more assays disclosed in U.S. patent publications 20040053304; 20040242510; 20050106593; and in issued U.S. patents 5,786,146; 6,017,704; 6,200,756; 6,265,171; 6,818,404; and 6,960,436, the disclosures of which are incorporated herein.

In one embodiment, scanning for one or more mutations at the APC-MCR (the APC mutation cluster region, see for example Miyoshi et al., 1992, Hum. Mol. Genet. 1(4):229-33) may detect adenomas with greater than 74% sensitivity. Scanning may be performed using a base scanning technique described herein, or any other suitable detection assay, to scan for the presence of one or more mutations within the APC sequence, the APC-MCR, for example within the sequence of SEQ ID NO: 1 or 11 or any portion thereof.

In one embodiment, the following panel may be used to detect adenomas with greater than 90% sensitivity: scanning for one or more mutations in the APC-MCR locus, exon 9 of the PIK3CA locus, exon 20 of the PIK3CA locus, B-catenin (e.g., exon 5), or a mutation in BRAF that results in a V599E amino acid change. Scanning as described herein can be used to detect one or more mutations in the APC-MCR locus, exon 9 of the PIK3CA locus, or exon 20 of the PIK3CA locus. Mutations at the BRAF locus can be detected via primer extension or other appropriate methodology (including scanning). One or more oligonucleotides may be used to capture, amplify, and/or assay for one or more of these markers (e.g., mutations). Different oligonucleotides may be designed based on the known nucleic acid sequences of PIK3CA (see, for example, Samuels et al., 2004, Science 304(5670):554), B-catenin (see, for example, Sparks et al.,

- 8 -

1998, Cancer Res. 58(6), 1130-4), and BRAF. For example, oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 4 or 5 may be used for B-catenin marker analysis; oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 6 or 14 may be used for BRAF marker analysis; and oligonucleotides having sequences identical to, or complementary to, portions of SEQ ID NO: 7, 8, or 9 may be used for PIK3CA analysis. The annealing position and size of the oligonucleotide(s) may be determined in part by the type of assay that is used.

In one embodiment, a combination of all of the above loci may be used to detect adenomas with a greater than 95% sensitivity (e.g., greater than 98% sensitivity).

Data obtained for 50 colonic adenomas that were analyzed using these different panels of markers is described in more detail in the examples.

It should be appreciated that other combinations of these markers can be used to obtain different levels of sensitivity for detecting adenomas (e.g., colonic adenomas).

Assays can be performed on stool samples (e.g., representative stool samples) using methods that can detect small amounts of mutant genetic material in a heterogeneous sample containing a majority of normal genetic material (e.g., where mutant genetic material accounts for less than 10%, less than 5%, less than 2.5%, and even less than 1% of the total genetic material in the sample). Accordingly, a genetic marker belonging to an informative panel of markers may be detected using one or more methods that can detect a low frequency event in a heterogeneous biological sample. Such methods may include PCR amplification, primer extension, and/or mutant scanning methods. The specificity and sensitivity of a primer extension or scanning reactions that include acyclo terminators and deoxynucleotides can be improved by using a combination of acyclo polymerase and TAQ polymerase (or other combination of a polymerase that preferentially incorporates acyclo terminators and a polymerase that preferentially incorporates deoxynucleotides). In some embodiments, one or more assays may be performed in a digital format (e.g., diluted so that on average 1, 1 to 5, or a few more different molecules are analyzed in each assay – it should be appreciated that the diluted sample may be amplified to increase the signal in each assay). In one embodiment, a digital analysis (e.g., a digital amplification and subsequent analysis) may be performed on at least a sufficient number of molecules to obtain a statistically significant result. Certain digital techniques are known in the art, see for example, US

- 9 -

Patent No. 6,440,706 and US Patent No. 6,753,147, the entire contents of which are incorporated herein by reference. Similarly, an emulsion-based method for amplifying and/or sequencing individual nucleic acid molecules may be used (e.g., BEAMing technology).

5 In one embodiment, a sequencing method that can sequence single molecules in a biological sample may be used. Sequencing methods are known and being developed for high throughput (e.g., parallel) sequencing of complex genomes by sequencing a large number of single molecules (often having overlapping sequences) and compiling the information to obtain the sequence of an entire genome or a significant portion thereof.

10 According to the invention, such methods, although designed for complex sequence analysis, may be particularly suited to sequence a large number of substantially identical molecules in order to identify the rare one(s) that contain a mutation or alteration.

 High complexity analytical or sequencing techniques may involve high speed parallel molecular nucleic acid sequencing as described in PCT publication

15 WO01/16375, U.S. application 60/151,580 and U.S. published application 20050014175, the entire contents of which are incorporated herein by reference. Other non-limiting sequencing techniques are described in PCT publications WO05/73410, WO05/54431, WO05/39389, WO05/03375, WO05/010145, WO04/069849, WO04/70005, WO04/69849, and WO04/70007, and U.S. published application 20050100932, the

20 entire contents of which are incorporated herein by reference.

 High complexity analytical or sequencing techniques may involve exposing a nucleic acid molecule to an oligonucleotide primer and a polymerase in the presence of a mixture of nucleotides. Changes in the fluorescence of individual nucleic acid molecules in response to polymerase activity may be detected and recorded. The specific labels

25 attached to each nucleic acid and/or nucleotide may provide an emission spectrum allowing for the detection of sequence information for individual template nucleic acid molecules. In certain embodiments, a label is attached to the primer/template and a different label is attached to each type of nucleotide (e.g., A, T/U, C, or G). Each label emits a distinct signal which is distinguished from the other labels.

30 High complexity analytical or sequencing techniques may involve or be based on methods or technology described in Shendure et al., Nature Reviews/Genetics, Volume 5, May 2004, pages 335-344; Braslavsky et al., PNAS, April 1, 2003, Volume 100, No. 7, pages 3960-3964; the entire disclosures of which are incorporated herein by reference.

- 10 -

In other embodiments, high complexity analytical or sequencing techniques may involve providing a primed target polynucleotide linked to a microfabricated synthesis channel, and flowing a first nucleotide through the synthesis channel under conditions such as to allow the first nucleotide to attach to the primer. The presence or absence of a
5 signal is determined, the presence indicating that the first nucleotide was incorporated into the primer and the identity of the complementary base that served as a template in the target polynucleotide is determined. The signal is then removed or reduced and the process repeated with a second nucleotide. The second nucleotide can be either the same as the first nucleotide or a different nucleotide. The specific labels attached to each
10 nucleic acid provide an emission spectra allowing for detection of sequence information of the nucleic acid molecule. In other embodiments, a plurality of different primed target polynucleotides linked to different synthesis channels may be used. In further embodiments, the polynucleotide may be attached to a surface. In some embodiments, a label is attached to the nucleotide.

15 In certain embodiments, a high complexity analytical or sequencing technique may be provided by Helicos BioSciences Corporation (Cambridge, MA). In some embodiments, a nucleic acid polymerase and a fluorescently labeled nucleotide may be added to an assay to bind to immobilized templates (e.g., bound to appropriate primers). The sample may be washed to remove unbound nucleotides and excess polymerase. The
20 sample may be analyzed and the positions of the incorporated nucleotides recorded. The fluorescent label may be removed and a second labeled nucleotide may be added. The process may be repeated several times until a desired length is reached.

Other useful genome/complex sequencing methods include high throughput sequencing using the 454 Life Sciences Instrument System. Briefly, a sample of single
25 stranded DNA may be prepared and added to an excess of DNA capture beads which are then emulsified. Clonal amplification may be performed to produce a sample of enriched DNA on the capture beads (the beads are enriched with millions of copies of a single clonal fragment). The DNA enriched beads may be then transferred into PicoTiterPlate™ and enzyme beads and sequencing reagents may be added. The
30 samples may be analyzed and the sequence data recorded. Pyrophosphate and luciferin are examples of the labels that can be used to generate the signal.

In other embodiments, single molecule sequencing technology available from US Genomics, Mass., may be used. For example, technology described, at least in part, in

- 11 -

one or more of U.S. patents 6,790,671; 6,772,070; 6,762,059; 6,696,022; 6,403,311; 6,355,420; 6,263,286; and 6,210,896, the disclosures of which are incorporated herein may be used.

5 Similar assays can be performed on other heterogeneous biological samples including fluids and mucus (e.g., urine, blood, serum, sputum, semen, breast nipple aspirate, or other bodily fluids).

 Assays also may be performed on one or more tissue biopsies (e.g., colon biopsies or biopsies of other tissues or organs). Tissue biopsies are expected to contain
10 more abnormal genetic material if they are positive for an adenoma. Accordingly, assays performed on tissue biopsies may not be as specific and sensitive as assays performed on heterogeneous biological samples.

 Regardless of the source of nucleic acid (e.g., a biological sample, a tissue biopsy, etc.), an analysis may involve a nucleic acid capture step, a nucleic acid
15 amplification step, and/or a nucleic acid analysis step (e.g., a using mutation-specific detection technique or a scanning technique, etc.). A capture probe that is complementary to one of the strands in the region of a sequence being assayed (e.g., the locus or position being interrogated for the presence of one or more point mutations, deletions, insertions, etc., or any combination thereof) may be used to capture nucleic
20 acid fragments for subsequent analysis. A capture probe may be between about 20 and about 50 nucleotides long (e.g., between about 25 and about 45, or about 30, about 35, or about 40 nucleotides long). However, in some embodiments shorter or longer capture probes may be used. A capture probe may be designed to be complementary to a sequence that is found in the vicinity of the region being assayed (e.g., within about 10,
25 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, or 1,000 nucleotides or more of the nucleic acid region being analyzed). It should be appreciated that in some embodiments, the distance between the capture probe binding site and the region of interest may be determined in part by the size of the fragment that is being amplified after capture. However, in many embodiments, a capture probe is more effective when it
30 is close to a sequence being assayed. Accordingly, a capture probe may be identical or complementary to one or more sequences of any of SEQ ID NO: 1-14. A capture probe may be attached to a solid support and used to hybrid capture sample nucleic acid of

- 12 -

interest (e.g., via electrophoresis, repeated electrophoresis, chromatography, repeated chromatography, mixing, etc.)

In some embodiments, a captured nucleic acid may be amplified. For example, a PCR amplification may be performed using a pair of amplification primers designed to amplify a region containing the sequence being assayed. In some embodiments, the amplification primers may amplify the region that hybridized to the capture probe. In some embodiments, the amplified region may be adjacent to (but does not include) the sequence that was bound during hybrid capture. Amplification primers may be between about 10 and about 50 nucleotides long (e.g., between about 15 and about 45, or about 20, about 25, about 30, about 35, about 40, or about 45 nucleotides long). However, in some embodiments shorter or longer amplification primers may be used. It should be appreciated that each primer in a pair of amplification primers may be complementary to a different (complementary) strand of the nucleic acid region being amplified. Accordingly, an amplification primer may be identical or complementary to one or more sequences of any of SEQ ID NO: 1-14. The amplification primers may be designed to amplify regions of different sizes. In some embodiments, amplification products may be from about 30 to about 5,000 nucleotides long (e.g., about 40, 50, 75, 100, 150, 200, 250, 500, 750, or 1,000 or more) depending, in part, on the assay format and the number and spacing of markers that are to be analyzed on the single amplification product. However, amplification products of shorter, longer, or intermediate sizes also may be analyzed. In some embodiments, one or more amplification products may be used to assay multiple positions on a predetermined gene or genetic region (e.g., within the APC, Kras, p53 or other locus). In some embodiments, a separate amplified region may be assayed for each marker in the panel. In some embodiments, several markers (e.g., 2, 3, 4, 5, 5-10, or more) may be assayed on a single amplified nucleic acid. In some embodiments, one or both amplification primers may be methylation specific primers (primers that are specific for methylated C by having Gs to pair with methylated C that is not modified by bisulfite treatment in a methylation detection assay).

Oligonucleotide primers also may be used for marker detection (e.g., on an amplified nucleic acid). Primers may be used for many different hybridization and/or extension based assays (e.g., one or more extension assays, single base extension assays, sequencing assays, scanning assays, methylation detection assays, etc.). Assay primers may be between about 10 and about 50 nucleotides long (e.g., between about 15 and

- 13 -

about 45, or about 20, about 25, about 30, about 35, about 40, or about 45 nucleotides long). However, in some embodiments shorter or longer amplification primers may be used. An assay primer may be designed to anneal to a target region adjacent or near a site of interest. For example, the 3' end of an assay primer may anneal immediately upstream of a position opposite a target position that is suspected of being altered (e.g., mutant) on the nucleic acid that is being assayed. In some embodiments, the 3' end of the assay primer may anneal 1 or more nucleotides upstream from the target position (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 10-15, 15-20, 20-25, 25-50, 50-100, or more nucleotides upstream from the target position). It should be appreciated that the presence of a marker (e.g., a mutation) may be assayed on either strand of a nucleic acid region of interest using appropriate assay primers. In some embodiments, an assay primer may be methylation specific.

Accordingly, in certain embodiments, a specific hybrid capture method may involve using a capture probe to bind to a target nucleic acid. The bound product then may be isolated. In one embodiment, a capture probe may be bound to a solid surface thereby acting as an anchor for isolating a target molecule. In other embodiments, a capture probe may be modified in a manner that allows it to be isolated or purified from a sample. For example, a capture probe may be biotinylated, attached to an antigen, attached to a magnetic particle, attached to a molecular weight marker, attached to a charged particle, attached to another particle or other molecular "hook" that can be used to isolate that capture probe and thereby isolate a target molecule that is hybridized to the probe.

In aspects of the invention, a nucleic acid preparation may be captured by repeated exposure of a biological sample (for example, a processed biological sample) to a capture probe on a solid support or in a medium, for example, by the rapid flow of the sample past a capture probe for the target nucleic acid molecule. The repetitive nature of such a method allows for a target molecule to bind and enhances the total number of molecules bound to the capture probe, providing a high yield capture. The solid support may be an electrophoretic medium (e.g., gel or beads) and the repetitive exposure of the sample to the capture probe may involve exposure to repeated cycles of electrophoresis in alternate directions (back and forth across a solid support region containing one or more different types of capture probe). In some aspects, a sample is added to a portion of an electrophoretic medium having at least two regions arranged consecutively in a

- 14 -

first spatial dimension. In some aspects, at least one of the at least two regions includes a first capture probe which is immobilized within that region. An electric field is applied to the electrophoretic medium in a first direction which is parallel to the first dimension. An electric field is then applied to the electrophoretic medium in a second direction
5 which is opposite to the first direction. In further aspects, the electric field is applied repeatedly in each direction. For further details see for example U.S. published application 20050247563 or PCT publication WO2005/047881, the entire contents of which are incorporated herein by reference.

In aspects of the invention, a sample may be a biological sample. A biological
10 sample may be, but is not limited to, stool, whole blood, serum, plasma, tears, saliva, nasal fluid, sputum, ear fluid, genital fluid, breast fluid, milk, colostrum, placental fluid, amniotic fluid, perspirate, synovial fluid, ascites fluid, cerebrospinal fluid, bile, gastric fluid, aqueous humor, vitreous humor, gastrointestinal fluid, exudate, transudate, pleural fluid, pericardial fluid, semen, upper airway fluid, peritoneal fluid, fluid harvested from a
15 site of an immune response, fluid harvested from a pooled collection site, bronchial lavage, urine, biopsy material, a nucleated cell sample, a fluid associated with a mucosal surface, hair, or skin. A sample also may be a pooled sample containing biological material and or isolated nucleic acids from a plurality of subjects (e.g., 2, 3, 4, 5, about 10, or more).

In aspects of the invention, a large amount of sample may be processed in order
20 to increase the confidence level of isolating or capturing a rare event indicative of very early stage disease (e.g., an adenoma, an early stage cancer, etc.). For example, about 10g, about 20g, about 30g, about 40g, about 50g, about 60g, about 70g, about 80g, about 90g, about 100g, about 150g, about 200g, or more stool sample may be processed using a
25 capture technique described herein.

In embodiments of the invention, exposure of a biological sample (for example a crude preparation of total nucleic acid from a biological sample) to immobilized capture probe(s) may be repeated between 2 and 100 times, e.g., between about 5 and about 50
30 times, between about 10 and about 40 times, or about 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, etc. times, including about 25, 30, or 35 times.

In aspects of the invention, a sample may be exposed repeatedly to a capture probe using chromatographic methods, for example high performance liquid

- 15 -

chromatography (HPLC), fast performance liquid chromatography (FPLC), etc., or any combination thereof.

A captured preparation of target nucleic acid molecules (e.g., of low genomic complexity) may be eluted using any suitable technique and prepared (e.g., single
5 stranded molecules may be prepared) for subsequent analysis using a technique for analyzing nucleic acid samples of high genomic complexity. Sample capture techniques described herein may be used to analyze DNA and/or RNA.

Methods of the invention may be useful for screening an individual as part of a routine cancer screen. Methods of the invention may be useful as part of a population
10 screen to identify individuals with early stages of cancer (e.g., colon cancer). Methods of the invention may be used to test patients suspected of having colon cancer (e.g., patients with polyps, a family history of colon cancer, or other indicators of cancer such as blood in the stool, etc.). Polyps may include non neoplastic polyps with a diameter of 1 cm or more (potentially significant polyps) and non neoplastic polyps with a diameter of less
15 than 1 cm or an unknown diameter.

In certain embodiments, a general population screen may be performed with markers that are greater than 60% informative. For example, a general population screen may be performed using a panel of multiple mutations (e.g., a multiple mutation panel described herein). In another example, a general population screen may be performed
20 using an assay for hypermethylation. In yet a further embodiment, a general population screen may be based on a scanning assay. Of course, any combination of the above types of assays may be used to obtain higher informativeness.

Analyses of individual patients may be performed using assays of different informativeness. If a patient has one or more signs of colon cancer (e.g., blood in the stool, a history of polyps, etc.) or other risk factors (e.g., age, diet, exposure to
25 carcinogens) an analysis of a marker panel that is sufficient for obtaining a sensitivity of 90%, 95% or more may be recommended. In other embodiments, a screen of a subset of markers (e.g., a panel of markers with lower informativeness) may be recommended for a general health screen of younger individuals (e.g., younger than 50, 40, 30, 20 etc.)
30 with no other risk factors or indicia of cancer.

Accordingly, aspects of the invention may be useful for detecting adenomas with greater than about 60%, greater than about 70%, greater than about 80%, greater than about 90%, or greater than about 95% confidence. Therefore, aspects of the invention

- 16 -

are useful for assaying patient samples for adenomas and reducing the number of false negatives to fewer than about 40%, fewer than about 30%, fewer than about 20%, fewer than about 10%, or fewer than about 5%.

5 In one aspect, an assay may be performed on a regular basis (e.g., annual medical checkup). The informativeness of this assay may increase over time, because additional genetic markers may become positive as adenomas develop over time. Therefore, initial false negatives may show up as positives in subsequent genetic tests. These later identifications still be may early enough for the patient to have a positive prognosis.

10 In general, the detection of one or more indicators of adenomas in a screen of the invention may be followed up by one or more subsequent analyses to locate and/or confirm the presence of an adenoma, and ultimately to treat any cancerous or precancerous lesions that may be detected. Invasive procedures such as colonoscopies or sigmoidoscopies may be used to locate and sample tissue for further analysis. Other less invasive procedures may include virtual colonoscopies. Treatments may include surgical
15 removal (e.g., of the lesion or of a region of the colon or other organ or tissue that contains the lesion), radiation, chemotherapy, or any combination thereof.

It should be appreciated that aspects of the invention described herein in the context of colonic adenomas and colon cancer may be used to screen patient samples for the presence of adenoma in other tissues.

20 According to aspects of the invention, biopsies that are removed in order to identify the source of mutant genetic material may be assayed using methods described herein. As discussed above, it should be appreciated that different cutoff levels may be used for tissue biopsy samples than for non-tissue samples, because biopsy samples may contain relatively more adenoma cells than biological fluid or solid samples (e.g., mucus, stool, etc.).
25

In another aspect, the invention provides groups of nucleic acid probes or primers that are useful for assaying a biological sample for the presence of each genetic abnormality included in a genetic abnormality panel that is at least 60% informative for adenoma. Certain nucleic acid probes or primers may be useful for amplifying regions
30 of the genome that contain one or more genetic loci of interest. Certain nucleic acid probes or primers may be useful for performing primer extension reactions on amplified or non-amplified template nucleic acid in order to assay for the presence of one or more genetic abnormalities. Primers may be oligonucleotides ranging from about 10

- 17 -

nucleotides to about 100 nucleotides in length, and preferably from about 20 to about 50 nucleotides in length. A primer for a single base extension reaction may be complementary to a genomic sequence that is adjacent to the position of a genetic abnormality included in an informative panel of genetic markers. Accordingly, aspects of the invention include a panel or group of oligonucleotides designed to interrogate a biological sample for the presence of each genetic abnormality that belongs to a panel that is at least about 60% informative for adenoma (and preferably at least about 70%, at least about 80%, at least about 90%, or at least about 95% informative for adenoma). Accordingly, aspects of the invention include a panel of assays designed to interrogate each of a panel of genetic loci for the presence of a genetic abnormality indicative of adenoma. Therefore, a kit of the invention may include one or more of a capture probe, an amplification primer pair, and/or an extension primer (or any combination of two or more thereof) for each locus that is being analyzed. In some embodiments, the kit may include a capture probe and/or an amplification primer pair and/or an extension primer for each assay included in a panel of assays described herein. It should be appreciated that in some embodiments, a single capture probe and/or a single amplification primer pair may be used to capture and/or amplify a single region that may be assayed for two or more markers (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 12-20 or more). For example, a single Kras region may be captured and/or amplified and subsequently assayed for mutations at two or more of the positions that are included in a chosen panel of markers.

Examples

Example 1

A commercially available 22 marker multiple mutation panel (V1) was developed for the detection of colorectal carcinomas and analysis of 147 cancer tissues revealed informativeness of ~71%. V2 markers including two hypermethylation sites, mutation site BRAF and scanning sites APC-MCR, PIK3CA-Ex9 and Ex20 and B-catenin were tested in the same tissues with resulting informativeness of 88% and an overall informativeness (i.e., V1+V2) of ~97%. To determine informativeness in adenoma, fifty tissue samples were tested with all V1 and V2 markers.

Results for V1 and V2 markers in 50 adenoma tissues are shown in Table 1. Two V1 cutoffs are presented that represent cutoffs based on stool data (Cutoff 1) and more

- 18 -

conservative cutoffs based on tissues that were > 2 fold over stool cutoffs (Cutoff 2). The APC-MCR scan had the highest informative value of 74% and V2 markers alone detected 94% of the adenomas suggesting this would be a good panel for early stage CRC detection.

5

Table 1

	Cutoff 1	Cutoff 2
V1	62.0%	54.0%
V2	94.0%	94.0%
V1+V2	98.0%	94.0%

	Alg-1A	Iq-2
KRAS	38.0%	32.0%
APC	38.0%	34.0%
P53	4.0%	2.0%
BAT-26	4.0%	4.0%
Total MuMu	62.0%	54.0%
APC-MCR	74.0%	
PIK3CA-9	14.0%	
PIK3CA-20	4.0%	
B-CAT	4.0%	
BRAF	4.0%	
Total Scan	90.0%	
HLTF	38.0%	
V29	50.0%	
Total Methylation	60.0%	

Example 2: Scanning (base tracking)

10 Scanning or base tracking methods of the invention may be used to screen a nucleic acid region (e.g., the APC-MCR) for the presence of one or more mutations at different positions within the nucleic acid region. Current methods of nucleotide sequencing use a single sequencing reaction containing a mixture of all four terminator nucleotides in the same reaction, where each terminator base is differentially labeled and
15 detected. The signal from an altered sequence present at low concentrations in a sequencing reaction is often masked by the signal of the wild type base at the same location. Variant sequences must be at least about 10% of the DNA being sequenced before their presence can be readily detected. In contrast, methods according to the invention increase the sensitivity of assays to detect nucleic acid alterations that are
20 present at a relatively low level in a sample, especially, *e.g.*, in a heterogeneous sample.

The present invention includes methods of screening nucleic acids for at least one genetic variation through the application of a novel modification of a DNA sequencing

- 19 -

reaction. Methods of the invention modify current sequence reactions such that only one terminator nucleotide, and not all four terminator nucleotides, is provided in the primer extension reaction to allow for single base scanning, which is also referred to herein as single base tracking. The modified reaction is herein referred to as a single base tracking
5 reaction.

Sensitivity is increased in single base tracking at least because signals from bases at any one position in a sequence being scanned are no longer masked by signals from an alternate base in the wild type sequences present at higher concentrations in the sample. Therefore, methods of the invention detect the presence of nucleotide sequences with
10 altered residues as compared to a control "wild type" nucleotide sequence, where the nucleotides with altered sequence make up about 50%, about 25%, about 10%, about 5%, about 4%, about 3%, about 2.5%, about 2%, about 1.5% or especially about 1% of the sample being analyzed. Such an increase in sensitivity has at least several uses. For example, methods according to the invention can be used to screen the human genome,
15 providing for increased sensitivity for detection of low frequency genetic variations.

In a preferred reaction, the terminator nucleotide is labeled. A preferred label is a fluorescent label, although it is within the skill of an artisan to use substitute labels of equal or higher sensitivity in signal detection, and/or equal or lower background signal noise. The DNA single base tracking reaction utilizes sensitive labeling techniques in
20 order that the resulting sequence fragments may be analyzed and, *e.g.*, compared to a known normal control sample to determine whether at least one genetic variation exists between the sample and normal control.

Additionally, methods of the invention can be used to screen for mutations that are predicative of a disease state. Often, these mutations are present in a sample at a
25 relatively low level, *e.g.*, where the mutation is a somatic mutation in a nucleic acid population obtained from biopsied tissue. Accordingly, methods according to the invention are more sensitive than current methods and can detect relatively low frequency mutations in a heterogeneous sample.

One aspect of the invention includes a method for detecting a difference between
30 two nucleic acids. The method includes extending a first primer complementary to a target nucleic acid in the presence of a first nucleotide and a second nucleotide to produce at least one product. The first nucleotide is at least one deoxynucleotide, and

- 20 -

more preferably is a mixture of the four deoxynucleotides, namely dATP, dCTP, dGTP and dTTP ("dNTP mixture"), used for the elongation step of the primer extension reaction. The second nucleotide is a terminator nucleotide, preferably includes a detectable label, and has the same base as one of the first deoxynucleotides. The method
5 also includes detecting a signal from the at least one product and comparing the signal from the at least one product with a signal that is generated from a comparison nucleic acid in substantially the same manner as the signal is generated from the target nucleic acid. A difference between the signals indicates at least one difference between the target nucleic acid and the comparison nucleic acid. Signal differences include the
10 addition of at least one peak, the deletion of at least one peak, or a shift in the position of at least one peak present in the sample as compared to the control.

The embodiments described above and below can have any or all of the following features. The method may include the step of amplifying a nucleic acid to form the target nucleic acid. The extending step can include extending the primer in the presence
15 of the deoxynucleotides dATP, dCTP, dGTP, and dTTP. The target nucleic acid can be a nucleic acid suspected of containing a mutation. The target nucleotides to be screened in the methods of the invention may be genomic DNA, complementary DNA (cDNA), or RNA. Where the initial sample is RNA, it is preferred that the RNA is converted into DNA prior to further processing. The extending and comparing steps can be repeated.
20 The extending and comparing steps can be conducted at least four times with the same primer, each time using a different one of adenine (A), cytosine (C), guanine (G) or thymidine (T) for the base of the second "terminating" nucleotide (*i.e.*, each extension reaction contains only one type of extension terminating nucleotide, where the terminating nucleotide may be a dideoxynucleotide or an acyclonucleotide, and the base
25 of the terminating nucleotide is chosen from A, C, G, or T.

The comparison nucleic acid can be a wild type nucleic acid. The signal from the comparison nucleic acid can be determined prior to, at the same time as, or after the signal from the target nucleic acid. The signal can include a fluorescent light emission. Alternatively, the signal results of the control sequence may be obtained from a database
30 of nucleotide sequences. The comparison step may be done manually or by automation.

The methods described above or below can also have any or all of the following features. In certain embodiments, the method includes extending a second primer

- 21 -

complementary to the target nucleic acid in the presence of the first nucleotide and the second nucleotide to produce at least one secondary product. In a preferred embodiment, the first nucleotide is a deoxynucleotide (dNTP) mixture, the second nucleotide is a terminator nucleotide (dideoxynucleotide or acyclonucleotide) of only one base selected from A, C, G or T, and the at least one secondary product is the product of a primer extension reaction. The method may also include detecting a signal from the at least one secondary product and comparing the signal from the at least one secondary product with a signal that was generated from a comparison nucleic acid in substantially the same manner as the signal was generated from the target nucleic acid. A difference between the signals indicates at least one difference between the target nucleic acid and the comparison nucleic acid.

The methods described above or below may also include the following features. In one embodiment, a second primer complementary to a strand complementary to the target nucleic acid is extended in the presence of the first nucleotide and the second nucleotide to produce at least one secondary product. In a preferred embodiment, the first nucleotide is a deoxynucleotide (dNTP) mixture, the second nucleotide is a terminator nucleotide (dideoxynucleotide or acyclonucleotide) of only one base selected from A, C, G or T, and the at least one secondary product is the product of a primer extension reaction. The method can then include detecting a signal from the at least one secondary product and comparing the signal from the at least one secondary product with a signal that is generated from a comparison nucleic acid in substantially the same manner as the signal is generated from the target nucleic acid. A difference between the signals indicates at least one difference between the target nucleic acid and the comparison nucleic acid.

In another aspect of the invention, a method for detecting a difference between two nucleic acids includes extending a first primer complementary to a target nucleic acid in the presence of a first nucleotide including a detectable label and a second nucleotide to produce at least one product. In a preferred embodiment, the first nucleotide is a deoxynucleotide (dNTP) mixture, the second nucleotide is a terminator nucleotide (dideoxynucleotide or acyclonucleotide) of only one base selected from A, C, G or T, and the at least one product is the product of a primer extension reaction. The method also includes detecting a signal from the at least one product and comparing the signal from the at least one product with a signal that is generated from a comparison

- 22 -

nucleic acid in substantially the same manner as the signal is generated from the target nucleic acid. A difference between the signals indicates at least one difference between the target nucleic acid and the comparison nucleic acid.

In another aspect of the invention, a method for detecting a difference between
5 two nucleic acids includes extending a first primer including a detectable label and being complementary to a target nucleic acid in the presence of a first nucleotide and a second nucleotide to produce at least one product. In a preferred embodiment, the first nucleotide is a deoxynucleotide (dNTP) mixture, the second nucleotide is a terminator nucleotide (dideoxynucleotide or acyclonucleotide) of only one base selected from A, C,
10 G or T, and the at least one product is the product of a primer extension reaction. The method also includes detecting a signal from the at least one product and comparing the signal from the at least one product with a signal that was generated from a comparison nucleic acid in substantially the same manner as the signal was generated from the target nucleic acid. A difference between the signals indicates at least one difference between
15 the target nucleic acid and the comparison nucleic acid.

In another aspect of the invention, a method for detecting a difference between two nucleic acids includes extending a first primer complementary to a target nucleic acid in the presence of a first nucleotide and a second nucleotide to produce at least one product. The second nucleotide is a terminator nucleotide and includes the same base as
20 the first nucleotide. In a preferred embodiment, the first nucleotide is a deoxynucleotide (dNTP) mixture, the second nucleotide is a terminator nucleotide (dideoxynucleotide or acyclonucleotide) of only one base selected from A, C, G or T, and the at least one product is the product of a primer extension reaction. The method also includes detecting a mass of the at least one product and comparing the mass of the at least one
25 product with a mass that is generated from a comparison nucleic acid in substantially the same manner as the mass is generated from the target nucleic acid. A difference between the masses indicates at least one difference between the target nucleic acid and the comparison nucleic acid.

30 **Example 3: Stool sample preparation**

The following example illustrates a method for preparing a DNA sample from a stool sample, see for example U.S. published applications 2004-0043467 and 2004-0014104, the entire contents of which are incorporated herein by reference.

- 23 -

A stool sample is collected and may be stored at -80°C before use. The sample is thawed and resuspended in buffer, for example 10mM Tris-Cl pH8.0, 1mM EDTA and 150 mM NaCl, or other suitable buffer as known to those of ordinary skill in the art. In one embodiment, the buffer may contain between 100 mM and 200 mM EDTA, for example about 150 mM EDTA. A suitable ratio of buffer to sample may be used, for example between 5:1 and 20:1 (mls/g of sample), for example about 7:1. The sample is then homogenized utilizing an EXACTOR stool shaker (EXACT Laboratories Marlborough, MA). Following homogenization, the stool sample is centrifuged to remove all particulate matter, and the supernatants are incubated at 37°C. Proteinase K (0.5 µg/µL) and SDS (0.5%) may be added at this point. The DNA is extracted from the supernatant using Tris saturated phenol (Gibco/BRL, Grand Island, N.Y.), phenol/chloroform/isoamyl alcohol (25:24:1), and chloroform. The DNA is then precipitated (1/10 volume 3M NaAc and an equal volume isopropanol), removed from solution by centrifugation, and resuspended in TE (0.01M Tris pH 7.4, 0.001M EDTA) buffer containing RNase A (2.5 µg/mL), or other suitable buffer.

Equivalents

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by examples provided, since the examples are intended as a single illustration of one aspect of the invention and other functionally equivalent embodiments are within the scope of the invention. Various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims. The advantages and objects of the invention are not necessarily encompassed by each embodiment of the invention. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Incorporation by Reference

All publications, patents and sequence database entries mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or

- 24 -

patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

- 25 -

Claims

1. A method of screening a subject for the presence of an adenoma, the method comprising
interrogating a biological sample obtained from a subject suspected of having an
5 adenoma for the presence of each of a panel of genetic markers, wherein the panel of
genetic markers is more than 60% informative for adenoma, and wherein the presence of
one or more of the genetic markers in the biological sample is indicative of an adenoma
in the subject.
- 10 2. The method of claim 1, wherein the panel of genetic markers is more than 70%
informative.
3. The method of claim 1, wherein the panel of genetic markers is more than 80%
informative.
- 15 4. The method of claim 1, wherein the panel of genetic markers is more than 90%
informative.
5. The method of claim 1, wherein the panel of genetic markers is more than 95%
20 informative.
6. The method of claim 1, wherein the panel of genetic markers comprises a
hypermethylated genetic locus.
- 25 7. The method of claim 6, wherein the biological sample is interrogated for the
presence of the hypermethylated genetic locus using a methylation specific primer
extension or amplification assay.
8. The method of claim 6, wherein the hypermethylated genetic locus is the HLTF
30 or V29 locus.
9. The method of claim 1, wherein the panel of genetic markers comprises a
predetermined mutation in a cancer-associated genetic locus.

- 26 -

10. The method of claim 9, wherein the biological sample is interrogated for the presence of the predetermined mutation using a primer extension assay.

11. The method of claim 9, wherein the predetermined mutation is a mutation at the
5 KRAS, APC, P53, BAT-26 or BRAF genetic locus.

12. The method of claim 1, wherein the panel of genetic markers comprises a mutation in a mutation cluster region.

10 13. The method of claim 12, wherein the biological sample is interrogated for the presence of a mutation in the mutation cluster region using a scanning assay.

14. The method of claim 6, wherein the mutation cluster region is APC-MCR, exon 9 of PIK3CA, exon 20 of PIK3CA, or an exon of B-catenin.

15

15. The method of claim 1, wherein the panel of genetic markers comprises a mutation in a mutation cluster region, a predetermined mutation in a cancer-associated genetic locus, and a hypermethylated genetic locus.

20 16. The method of claim 1, wherein the biological sample is a stool sample.

17. The method of claim 1, wherein the biological sample is a tissue biopsy sample.

18. The method of claim 1, wherein the adenoma is a colonic adenoma.

25

19. The method of claim 1, wherein the adenoma is an invasive adenoma.

20. A method of detecting indicia of adenoma in a biological sample, the method comprising assaying the biological sample for the presence of one or more genetic
30 abnormalities from a group of genetic abnormalities that is more than 60% informative for adenoma.

- 27 -

21. A method of detecting adenoma in a subject, the method comprising performing, on a biological sample obtained from a subject suspected of having an adenoma, an assay that is more than 60% informative for adenoma.
- 5 22. A kit comprising a group of oligonucleotides, wherein each oligonucleotide is adapted for interrogating a genetic locus for the presence of a genetic marker belonging to panel that is at least 60% informative for adenoma.

SEQUENCE LISTING

<110> Exact Sciences, Corp.

<120> METHODS AND COMPOSITIONS FOR DETECTING ADENOMA

<130> E0411.70100W000

<140> Not yet assigned

<141> 2006-03-01

<150> US 60/657,841

<151> 2005-03-01

<160> 14

<170> PatentIn version 3.3

<210> 1

<211> 10719

<212> DNA

<213> Homo sapiens

<400> 1

tggagacaga atggaggtgc tgccggactc ggaaatgggg tccaagggtg gccaggatg	60
gctgcagctt catatgatca gttgttaaag caagttgagg cactgaagat ggagaactca	120
aatcttcgac aagagctaga agataattcc aatcatctta caaaactgga aactgaggca	180
tctaatatga aggaagtact taaacaacta caaggaagta ttgaagatga agctatggct	240
tcttctggac agattgattt attagagcgt cttaaagagc ttaacttaga tagcagtaat	300
ttccctggag taaaactgcg gtcaaaaatg tccctccgtt cttatggaag ccgggaagga	360
tctgtatcaa gccgttcttg agagtgcagt cctgttccta tgggttcatt tccaagaaga	420
gggtttgtaa atggaagcag agaaagtact ggatathtag aagaacttga gaaagagagg	480
tcattgcttc ttgctgatct tgacaaagaa gaaaaggaaa aagactggta ttacgctcaa	540
cttcagaatc tactaaaag aatagatagt cttcctttaa ctgaaaattt ttcocttaca	600
acagatatga ccagaaggca attggaatat gaagcaaggc aaatcagagt tgcgatggaa	660
gaacaactag gtacctgcca ggatatggaa aaacgagcac agcgaagaat agccagaatt	720
cagcaaatcg aaaaggacat acttcgtata cgacagcttt tacagtccca agcaacagaa	780
gcagagaggt catctcagaa caagcatgaa accggctcac atgatgctga gcggcagaat	840
gaagggtcaag gagggggaga aatcaacatg gcaacttctg gtaatggtca gggttcaact	900
acacgaatgg accatgaaac agccagtgtt ttgagttcta gtagcacaca ctctgcacct	960
ogaaggctga caagtcatct gggaaccaag gtggaaatgg tgtattcatt gttgtcaatg	1020
cttggtactc atgataagga tgatatgtcg cgaactttgc tagctatgtc tagctcccaa	1080
gacagctgta tatccatgcg acagtctgga tgtcttcctc tctcatcca gcttttacat	1140

ggcaatgaca aagactctgt attgttggga aattcccggg gcagtaaaga ggctcgggcc 1200
agggccagtg cagcactcca caacatcatt cactcacagc ctgatgacaa gagaggcagg 1260
cgtgaaatcc gagtccttca tcttttggaa cagatacgcg cttactgtga aacctgttgg 1320
gagtggcagg aagctcatga accaggcatg gaccaggaca aaaatccaat gccagctcct 1380
gttgaacatc agatctgtcc tgctgtgtgt gttctaataa aactttcatt tgatgaagag 1440
catagacatg caatgaatga actaggggga ctacaggcca ttgcagaatt attgcaagtg 1500
gactgtgaaa tgtatgggct tactaatgac cactacagta ttacactaag acgatatgct 1560
ggaatggctt tgacaaactt gacttttggga gatgtagcca acaaggctac gctatgctct 1620
atgaaaggct gcatgagagc acttgtggcc caactaaaat ctgaaagtga agacttacag 1680
caggttattg cgagtgtttt gaggaatttg tcttggcgag cagatgtaaa tagtaaaaag 1740
acgttgcgag aagttggaag tgtgaaagca ttgatggaat gtgctttaga agttaaaaag 1800
gaatcaaccc tcaaaagcgt attgagtgcc ttatggaatt tgtcagcaca ttgcaactgag 1860
aataaagctg atatatgtgc tgtagatggt gcacttgcac ttttggttgg cactcttact 1920
taccggagcc agacaaacac tttagccatt attgaaagtg gaggtgggat attacggaat 1980
gtgtccagct tgatagctac aaatgaggac cacaggcaaa tcctaagaga gaacaactgt 2040
ctacaaactt tattacaaca cttaaaatct catagtttga caatagtcag taatgcatgt 2100
ggaactttgt ggaatctctc agcaagaaat cctaaagacc aggaagcatt atgggacatg 2160
ggggcagtta gcatgctcaa gaacctcatt cattcaaagc acaaaatgat tgctatggga 2220
agtgtgtcag ctttaaggaa tctcatggca aataggcctg cgaagtacaa ggatgccaat 2280
attatgtctc ctggctcaag cttgccatct cttcatgtta ggaaacaaaa agccctagaa 2340
gcagaattag atgctcagca cttatcagaa acttttgaca atatagacaa ttttaagtccc 2400
aaggcatctc atcgtagtaa gcagagacac aagcaaagtc tctatggtga ttatgttttt 2460
gacaccaatc gacatgatga taataggtca gacaatttta atactggcaa catgactgtc 2520
ctttcaccat atttgaatac tacagtgtta cccagctcct cttcatcaag aggaagctta 2580
gatagttctc gttctgaaaa agatagaagt ttggagagag aacgcggaat tggcttaggc 2640
aactaccatc cagcaacaga aaatccagga acttcttcaa agcgagggtt gcagatctcc 2700
accactgcag ccagattgc caaagtcatt gaagaagtgt cagccattca tacctctcag 2760
gaagacagaa gttctgggtc taccactgaa ttacattgtg tgacagatga gagaaatgca 2820
cttagaagaa gctctgctgc ccatacacat tcaaacactt acaatttcac taagtccgaa 2880
aattcaaata ggacatgttc tatgccttat gccaaattag aatacaagag atcttcaaat 2940

gatagtttaa	atagtgtcag	tagtagtgat	ggttatggta	aaagaggtca	aatgaaaccc	3000
togattgaat	cctattctga	agatgatgaa	agtaagtttt	gcagttatgg	tcaataccca	3060
gccgacctag	cccataaaat	acatagtgcg	aatcatatgg	atgataatga	tggagaacta	3120
gatacaccaa	taaattatag	tcttaaatat	tcagatgagc	agttgaactc	tgggaaggcaa	3180
agtccttcac	agaatgaaag	atgggcaaga	cccaaacaca	taatagaaga	tgaataaaaa	3240
caaagtgagc	aaagacaatc	aaggaaatcaa	agtacaactt	atcctgttta	tactgagagc	3300
actgatgata	aacacctcaa	gttccaacca	cattttggac	agcaggaatg	tgttttctcca	3360
tacagggtcac	ggggagccaa	tggttcagaa	acaaatcgag	tgggttctaa	tcatggaatt	3420
aatcaaaatg	taagccagtc	tttgtgtcaa	gaagatgact	atgaagatga	taagcctacc	3480
aattatagtg	aacgttactc	tgaagaagaa	cagcatgaag	aagaagagag	accaacaaat	3540
tatagcataa	aataataatga	agagaaacgt	catgtggatc	agcctattga	ttatagttta	3600
aaatatgccg	cagatatttc	ttcatcacag	aaacagtcac	tttcattctc	aaagagttca	3660
tctggacaaa	gcagtaaaac	cgaacatatg	tcttcaagca	gtgagaatac	gtccacacct	3720
tcatctaata	ccaagaggca	gaatcagctc	catccaagtt	ctgcacagag	tagaagtggg	3780
cagcctcaaa	aggctgccac	ttgcaaagtt	tcttctatta	accaagaaac	aatacagact	3840
tattgtgtag	aagatactcc	aatatgtttt	tcaagatgta	gttcattatc	atctttgtca	3900
tcagctgaag	atgaaatagg	atgtaatcag	acgacacagg	aagcagattc	tgctaatacc	3960
ctgcaaatag	cagaaataaa	agaaaagatt	ggaactaggt	cagctgaaga	tcctgtgagc	4020
gaagttccag	cagtgtcaca	gcaccctaga	accaaataca	gcagactgca	gggttctagt	4080
ttatcttcag	aatcagccag	gcacaaagct	gttgaatttt	cttcaggagc	gaaatctccc	4140
tccaaaagtg	gtgctcagac	acccaaaagt	ccacctgaac	actatgttca	ggagacccca	4200
ctcatgttta	gcagatgtac	ttctgtcagt	tcaattgata	gttttgagag	tcgttcgatt	4260
gccagctccg	ttcagagtga	accatgcagt	ggaatggtaa	gtggcattat	aagccccagt	4320
gatcttccag	atagccctgg	acaaaccatg	ccaccaagca	gaagtaaaac	acctccacca	4380
cctcctcaaa	cagctcaaac	caagcgagaa	gtacctaaaa	ataaagcacc	tactgctgaa	4440
aagagagaga	gtggacctaa	gcaagctgca	gtaaatgctg	cagttcagag	ggtccagggt	4500
cttcagatg	ctgatacttt	attacatttt	gccacggaaa	gtactccaga	tggattttct	4560
tgttcatcca	gcctgagtgc	tctgagcctc	gatgagccat	ttatacagaa	agatgtggaa	4620
ttaagaataa	tgccctccagt	tcaggaaaat	gacaatggga	atgaaacaga	atcagagcag	4680
cctaaagaat	caaatgaaaa	ccaagagaaa	gaggcagaaa	aaactattga	ttctgaaaag	4740
gacctattag	atgattcaga	tgatgatgat	attgaaatac	tagaagaatg	tattattttct	4800

gccatgccaa	caaagtcac	acgtaaagca	aaaaagccag	cccagactgc	ttcaaaatta	4860
cctccacctg	tggcaaggaa	accaagtcag	ctgcctgtgt	acaaacttct	accatcacia	4920
aacagggttg	aacccccaaa	gcatgttagt	tttacaccgg	gggatgat	gccacgggtg	4980
tattgtgttg	aaggacacc	tataaacttt	tccacagcta	catctctaag	tgatctaaca	5040
atcgaatccc	ctccaaatga	gtagctgct	ggagaaggag	ttagaggagg	ggcacagtca	5100
ggtgaatttg	aaaaacgaga	taccattcct	acagaaggca	gaagtacaga	tgaggctcaa	5160
ggaggaaaaa	cctcatctgt	aaccatacct	gaattggatg	acaataaagc	agagggaagg	5220
gatattcttg	cagaatgcat	taattctgct	atgcccagag	ggaaaagtca	caagcctttc	5280
cgtgtgaaaa	agataatgga	ccaggctcag	caagcatctg	cgtcttcttc	tgaccccaac	5340
aaaaatcagt	tagatggtaa	gaaaaagaaa	ccaacttcac	cagtaaaacc	tataccacia	5400
aatactgaat	ataggacacg	tgtaagaaaa	aatgcagact	caaaaaataa	tttaaagtgt	5460
gagagagttt	tctcagacia	caaagattca	aagaaacaga	atttgaaaaa	taattccaag	5520
gtcttcaatg	ataagctccc	aaataatgaa	gatagagtca	gagggaagtt	tgcttttgat	5580
tcacctcatc	attacacgcc	tattgaagga	actccttact	gtttttcacg	aaatgattct	5640
ttgagttctc	tagattttga	tgatgatgat	gttgaccttt	ccagggaaaa	ggctgaatta	5700
agaaaggcaa	aagaaaataa	ggaatcagag	gctaaagtta	ccagccacac	agaactaacc	5760
tccaaccaac	aatcagctaa	taagacacia	gctattgcaa	agcagccaat	aaatcgagg	5820
cagcctaaac	ccatacttca	gaaacaatcc	acttttcccc	agtcattcaa	agacatacca	5880
gacagagggg	cagcaactga	tgaaaagtta	cagaattttg	ctattgaaaa	tactccggtt	5940
tgcttttctc	ataattcctc	tctgagttct	ctcagtgaca	ttgaccaaga	aaacaacaat	6000
aaagaaaatg	aacctatcaa	agagactgag	cccctgact	cacagggaga	accaagtaaa	6060
cctcaagcat	caggctatgc	tcctaaatca	tttcatgttg	aagatacccc	agtttgtttc	6120
tcaagaaaaa	gttctctcag	ttctcttagt	attgactctg	aagatgacct	gttgacaggaa	6180
tgtataagct	ccgcaatgcc	aaaaaagaaa	aagccttcaa	gactcaaggg	tgataatgaa	6240
aaacatagtc	ccagaaatat	gggtggcata	ttaggtgaag	atctgacact	tgatttgaaa	6300
gatatacaga	gaccagattc	agaacatgg	ctatccctg	attcagaaaa	ttttgattgg	6360
aaagctattc	aggaagggtg	aaattccata	gtaagtagtt	tacatcaagc	tgctgctgct	6420
gcatgtttat	ctagacaagc	ttcgtctgat	tcagattcca	tcctttccct	gaaatcagga	6480
atctctctgg	gatcaccatt	tcattctaca	cctgatcaag	aagaaaaacc	ctttacaagt	6540
aataaaggcc	cacgaattct	aaaaccaggg	gagaaaagta	cattggaaac	taaaaagata	6600

gaatctgaaa gtaaaggaat caaaggagga aaaaaagttt ataaaagttt gattactgga	6660
aaagttcgat ctaattcaga aatttcaggc caaatgaaac agccccttca agcaaactg	6720
ccttcaatct ctogaggcag gacaatgatt catattccag gagttcgaaa tagctcctca	6780
agtacaagtc ctgttttctaa aaaaggccca ccccttaaga ctccagcctc caaaagccct	6840
agtgaaggtc aaacagccac cacttctcct agaggagcca agccatctgt gaaatcagaa	6900
ttaagccctg ttgccaggca gacatcccaa ataggtgggt caagtaaagc accttctaga	6960
tcaggatcta gagattcgac cccttcaaga cctgcccagc aaccattaag tagacctata	7020
cagtctcctg gccgaaactc aatttcccct ggtagaaatg gaataagtcc tctaacaaa	7080
ttatctcaac ttccaaggac atcatcccct agtactgctt caactaagtc ctgaggttct	7140
ggaaaaatgt catatacatc tccaggtaga cagatgagcc aacagaacct taccaaacia	7200
acaggtttat ccaagaatgc cagtagtatt ccaagaagtg agtctgcctc caaaggacta	7260
aatcagatga ataatggtaa tggagccaat aaaaaggtag aactttctag aatgtcttca	7320
actaaatcaa gtggaagtga atctgataga tcagaaagac ctgtattagt acgccagtca	7380
actttcatca aagaagctcc aagcccaacc ttaagaagaa aattggagga atctgcttca	7440
tttgaatctc tttctccatc atctagacca gcttctocca ctaggtcca ggcacaaact	7500
ccagttttta gtccttccct tctgatatg tctctatcca cacattcgtc tgttcaggct	7560
ggtggatggc gaaaactccc acctaatctc agtcccacta tagagtataa tgatggaaga	7620
ccagcaaagc gccatgatat tgcacggtct cattctgaaa gtccctctag acttccaatc	7680
aataggtcag gaacctggaa acgtgagcac agcaaacatt catcatccct tctcgagta	7740
agcacttggga gaagaactgg aagttcatct tcaattcttt ctgcttcac agaatccagt	7800
gaaaaagcaa aaagtgagga tgaaaaacat gtgaactcta tttcaggaac caaacaagt	7860
aaagaaaacc aagtatccgc aaaaggaaca tggagaaaaa taaaagaaa tgaattttct	7920
cccacaaata gtacttctca gaccgtttcc tcagggtgcta caaatggtgc tgaatcaaag	7980
actctaattt atcaaatggc acctgctgtt tctaaaacag aggatgtttg ggtgagaatt	8040
gaggactgtc ccattaacaa tcctagatct ggaagatctc ccacaggtaa tactcccccg	8100
gtgattgaca gtgtttcaga aaaggcaaat ccaaactta aagattcaaa agataatcag	8160
gcaaaacaaa atgtgggtaa tggcagtggt cccatgcgta ccgtgggttt ggaaaatgc	8220
ctgaactcct ttattcaggt ggatgccct gacaaaaag gaactgagat aaaaccagga	8280
caaaataatc ctgtccctgt atcagagact aatgaaagtt ctatagtga acgtaccca	8340
ttcagttcta gcagctcaag caaacacagt tcacctagt ggactgttgc tgccagagt	8400
actcctttta attacaacc aagccctagg aaaagcagcg cagatagcac ttcagctcgg	8460

ccatctcaga tcccaactcc agtgaataac aacacaaaga agcgagattc caaaactgac 8520
agcacagaat ccagtggaac ccaaagtcc t aagcgccatt ctgggtctta ccttgtgaca 8580
tctgtttaaa agagaggaag aatgaaacta agaaaattct atgttaatta caactgctat 8640
atagacattt tgtttcaa at gaaacttta aagactgaaa aattttgtaa ataggtttga 8700
ttcttgtagt aggggtttttg ttctggaagc catatttgat agtatacttt gtcttcaactg 8760
gtcttattttt gggaggcact cttgatgggt agggaaaaaa tagtaaagcc aagtatgttt 8820
gtacagtatg ttttacctgt atttaaagta gcatcccatc ccaacttcct ttaattattg 8880
cttgtcttaa aataatgaac actacagata gaaaatatga tatattgctg ttatcaatca 8940
tttctagatt ataaactgac taaacttaca tcagggaaaa attgggtattt atgcaaaaaa 9000
aaatgttttt gtcttctgtg gtccatctaa catcataatt aatcatgtgg ctgtgaaatt 9060
cacagtaata tgggtcccgga tgaacaagtt taaccagcct gctttgcttt actgoatgaa 9120
tgaaactgat ggttcaattt cagaagtaat gattaacagt tatgtgggtca catgatgtgc 9180
atagagatag ctacagtgtg ataatttaca ctattttgtg ctccaaacaa aacaaaaatc 9240
tgtgtaactg taaaacattg aatgaaacta ttttacctga actagatttt atctgaaagt 9300
aggtagaatt tttgctatgc tgtaatttgt tgtatattct ggtatttgag gtgagatggc 9360
tgctctttta ttaatgagac atgaattgtg tctcaacaga aactaaatga acatttcaga 9420
ataaattatt gctgtatgta aactgttact gaaattggta tttgtttgaa gggctctgtt 9480
tcacatttgt attaataatt gtttaaaatg cctcttttaa aagcttatat aaattttttt 9540
cttcagcttc tatgcattaa gagtaaaatt cctcttactg taataaaaaac aattgaagaa 9600
gactgttgcc acttaaccat tccatgcgtt ggcaacttatc tattcctgaa atttctttta 9660
tgtgattagc tcatcttgat ttttaatat tttccactta aacttttttt tottactcca 9720
ctggagctca gtaaaagtaa attcatgtaa tagcaatgca agcagcctag cacagactaa 9780
gcattgagca taataggccc acataatttc ctctttctta atattataga attctgtact 9840
tgaaattgat tcttagacat tgcagtctct tcgaggcttt acagtgtaaa ctgtcttgcc 9900
ccttcatctt cttgttgcaa ctgggtctga catgaacact ttttatcacc ctgtatgtta 9960
gggcaagatc tcagcagtga agtataatca gcactttgcc atgctcagaa aattcaaact 10020
acatggaact ttagaggtag atttaatacg attaagatat tcagaagtat attttagaat 10080
ccctgcctgt taaggaaact ttatttgttg taggtacagt tctggggtac atgttaagtg 10140
tccccttata cagtggaggg aagtcttcct tcctgaagga aaataaactg acacttatta 10200
actaagataa ttacttaat atatcttccc tgatttgttt taaaagatca gagggtgact 10260

gatgatacat gcatacatat ttgttgaata aatgaaaatt ttttttagt gataagattc 10320
atacactctg ttttgggga gggaaaacct ttttaagcat ggtggggcac tcagatagga 10380
gtgaatacac ctacctggtg ctttgaaaat cacatcaagt agttaattat ctaccoccta 10440
cctgtgttta taacttccag gtaatgagaa tgattttttt taaagctaaa atgccagtaa 10500
ataaaagtgc tatgacttga gctaagatat ttgactccaa tgcctgtact gtgtctactg 10560
caccactttg taaacacttc aatttactat ctttgaaatg attgaccttt aaatttttgc 10620
caaatgttat ctgaaattgt ctatgaatac catctacttc tgttgttttc ccaggcttcc 10680
ataaacaatg gagatacatg caaaaaaaaa aaaaaaaaaa 10719

<210> 2
<211> 648
<212> DNA
<213> Homo sapiens

<400> 2
actaggaaaa ctgtaacaat aagagtggag atagctgtca gcaacttttg tgaggggtgtg 60
ctacaggggtg tagagcactg tgaagtctct acatgagtga agtcatgata tgatcctttg 120
agagccttta gcgcgcgag aacagcagtc tggctattta gatagaacaa cttgatttta 180
agataaaaga actgtctatg tagcatttat gcatttttct taagcgtcga tggaggagtt 240
tgtaaatgaa gtacagttca ttacgataca cgtctgcagt caactggaat tttcatgatt 300
gaattttgta aggtattttg aaataatttt tcatataaag gtgagtttgt attaaaaggt 360
actggtggag tttttgatag tgtattaacc ttatgtgtga catgttctaa tatagtcaca 420
ttttcattat ttttattata aggcctgctg aaaatgactg aatataaact tgtggtagtt 480
ggagctggtg gcgtaggcaa gagtgccttg acgatacagc taattcagaa tcattttgtg 540
gacgaatatg atccaacaat agaggtaaatt cttgttttaa tatgcatatt actggtgcag 600
gaccattctt tgatacagat aaaggtttct ctgaccattt tcatgagt 648

<210> 3
<211> 20303
<212> DNA
<213> Homo sapiens

<400> 3
ttcccatcaa gccctagggc tcctogtggc tgctgggagt tgtagtctga acgcttctat 60
cttggcgaga agcgcctacg ctccccctac cgagtccgcg ggtaattctt aaagcacctg 120
caccgcccc cgcgcgcctg cagaggggcg agcagggtctt gcacctcttc tgcattctcat 180
tctccaggct tcagacctgt ctccctcatt caaaaaatat ttattatcga gctcttactt 240
gctaccagc actgatatag gcactcagga atacaacaat gaataagata gtagaaaaat 300

tctatatcct cataaggctt acgtttccat gtactgaaag caatgaacaa ataaatctta 360
tcagagtgat aagggttggtg aaggagatta aataagatgg tgtgatataa agtatctggg 420
agaaaacgtt aggggtgtgat attacggaaa gccttcctaa aaaatgacat tttaactgat 480
gagaagaaaag gatccagctg agagcaaacg caaaagcttt cttccttcca cccttcatat 540
ttgacacaat gcaggattcc tccaaaatga tttccaccaa ttctgccctc acagctctgg 600
cttgcagaat tttccacccc aaaatgttag tatctacggc accaggtcgg cgagaatcct 660
gactctgcac cctcctcccc aactccattt cctttgcttc ctccggcagg cggattactt 720
gcccttactt gtcattggcg ctgtccagct ttgtgccagg agcctcgcag gggttgatgg 780
gattgggggtt ttccccctccc atgtgctcaa gactggcgct aaaagttttg agcttctcaa 840
aagtctagag ccaccgtcca gggagcagggt agctgctggg ctccggggac actttgcgtt 900
cgggctggga gcgtgctttc cacgacgggtg acacgcttcc ctggattggg taagctcctg 960
actgaacttg atgagtcctc tctgagtcac gggctctcgg ctccgtgtat tttcagctcg 1020
ggaaaatcgc tggggctggg ggtggggcag tggggactta gcgagtttg gggtagtggt 1080
gatggaagct tggctagagg gatcatcata ggagttgcat tgttgggaga cctgggtgta 1140
gatgatgggg atgttaggac catccgaact caaagttgaa cgcctaggca gaggagtgga 1200
gctttgggga accttgagcc ggctaaagc gtacttcttt gcacatccac ccggtgctgg 1260
gcgtagggaa tccctgaaat aaaagatgca caaagcattg aggtctgaga cttttggatc 1320
tcgaaacatt gagaactcat agctgtatat tttagagccc atggcatcct agtgaaaact 1380
ggggctocat tccgaaatga tcatttgggg gtgatccggg gagcccaagc tgctaaggtc 1440
ccacaacttc cggacctttg tccttcctgg agcgtcttt ccaggcagcc cccggctccg 1500
ctagatggag aaaatccaat tgaaggctgt cagtcgtgga agtgagaagt gctaaaccag 1560
gggtttgccc gccaggccga ggaggaccgt cgcaatctga gaggcccggc agccctgtta 1620
ttgtttggct ccacatttac atttctgcct cttgcagcag catttccggg ttctttttgc 1680
cggagcagct cactattcac ccgatgagag gggaggagag agagagaaaa tgccttttag 1740
gccggttcct cttaacttggc agagggaggc tgctattctc cgctgcatt tctttttctg 1800
gattacttag ttatggcctt tgcaaaggca ggggtatttg ttttgatgca aacctcaatc 1860
cctccccctt tttgaatggt gtgccccacc cccgggtcg cctgcaacct aggcggacgc 1920
taccatggcg tagacaggga gggaaagaag tgtgcagaag gcaagcccg aggcactttc 1980
aagaatgagc atatctcatc ttcccggaga aaaaaaaaaa agaatggtac gtctgagaat 2040
gaaattttga aagagtgoaa tgatgggtcg tttgataatt tgcgggaaa aacaatctac 2100

ctgttatcta gctttgggct aggccattcc agttccagac gcaggctgaa cgtcgtgaag	2160
cggaaggggc gggcccgcag gcgtccgtgt ggtccctccgt gcagccctcg gcccgagccg	2220
gttcttcctg gtaggaggcg gaactogaat tcattttctcc cgctgccccca tctcttagct	2280
cgcggttgtt tcattccgca gtttcttccc atgcacctgc cgcgtaccgg ccactttgtg	2340
ccgtacttac gtcattctttt tcctaaatcg aggtggcatt tacacacagc gccagtgcac	2400
acagcaagtg cacaggaaga tgagttttgg cccctaaccg ctccgtgatg cctaccaagt	2460
cacagaccct tttcatcgtc ccagaaacgt ttcattcacgt ctcttcccag tcgattcccg	2520
acccacactt tattttgatc tccataacca ttttgccctgt tggagaactt catatagaat	2580
ggaatcagga tgggcgctgt ggctcacgcc tgcactttgg ctcacgcctg cactttggga	2640
ggccgaggcg ggcggtattac ttgaggatag gaggttccaga ccagcgtggc caacgtgggtg	2700
aatccccgtc tctactaaaa aatacaaaaa ttagctgggc gtggtgggtg cctgtaatcc	2760
cagctattcg ggaggggtgag gcaggagaat cgcttgaacc cgggaggcag aggttgacgt	2820
gagccaagat cgtgccacta cactccagcc tgggcgacaa gaacgaaact ccgtctcaaa	2880
aaaaaggggg gaatcataca ttatgtgctc atttttgcg ggcttctgtc cttcaatgta	2940
ctgtctgaca ttcgttcatg ttgtatatat cagtattttg ctcccttttca tttagtatag	3000
tccatcgatt gtatatccgt ccttttgatg gccttttgag ttgtttccca tttgcggtta	3060
tgaaataaag ctgctataaa cattcttgta caattctttt tgtgatcata tgttttcgtg	3120
tttcttgag aaataacttag gaggggaatt gtggaggaag taaaaagtag ctgtattttg	3180
aactttttca gaagctctga gttttccaga gcggtgttac cattttacac tccaactagc	3240
aaggatatgg agttattatg gttgtgccac agccttccgg acattaggta tgtcagtcctt	3300
tctaattgtg tatatccttg tggttgtaat ttacagttct ctattgacta aggatgttca	3360
gcattttttc atgtgcctat tggccattcg tattttgttt gtaaagtagc tcttcgagtc	3420
ttttacctgt tattttgggt tttttgtttg tttttattgt tcagttgtgg gactgcttta	3480
tactttctgg atacaagtcc ttatcagatc catgagtcgt gaatgttttc ttctgatctg	3540
ttgcgggcct atttgtttgc tttacagagt ttacagaatc ttaagaggag tggattaatc	3600
ttttttatgt tcagtatttg ccttgcctcg tttaggacat cttttttttt ttttttaacc	3660
ccagggtcat gaagatataa tottacattt tcttttagga cctttatggg ggtaagtttt	3720
acagtaaggc ccttaagcca ttaattaatt cttaaaatta attgtttatg gtgtgagggtg	3780
taggagtcag tctctggtat ctttctctgta tggaaatcca gttattotgt ctccacttgt	3840
tgaaataggc ttcctttctc tactgaatgc ttttaatttt aattatttta cagttggagt	3900
atagggtac catttttagtg ctattttctt tttttctttg ttaatttttg agacagggac	3960

tcacactgtt gccaggcta gagtacaatg gcacaatcaa ggcttactgc agcctogaac 4020
ccctgggctc aagcagtcct ctagcagcct cacgagtagc tgggattact ccaccacacc 4080
cagctaacta ttttattttt ttgtattgac aggatctcac tatgttgccc aggctggtct 4140
caaactgctg gcctcaagct ttcattcccat ctgggcctcc caaagtgctg ggattacagg 4200
tgtgagccac catgcctgac ctcttagtgc tattttctat ttatctctc tgttctctgc 4260
tctctttaa cgttggagga agaaacagta cccattctac acaaactctt cagaaaacag 4320
aggaacagac tgggcgcggt ggctcatacc tgtaatctca gcactttggt acgctgaggc 4380
aggggatcat ttgaggtcgg gagttcgaga ccagcctggc caacacggcg aaaccccatc 4440
tctactaaaa tacaaaagta gctaggcgtg caccatacct gtaatgccag ttactcagga 4500
ggctgaggca caagaatccc ttgaacctgg gaagcggagg ttgcagttag ccgagattgc 4560
gccactgcac tccagcctgg gcaacagagt gagacctgt ctcaaaaaa aaaagaaaga 4620
aagaaaaaat agaggaatat ttcccaactt gttttogaag ccaggataat cctggtacca 4680
aaaccaaaca aggacattat aagaaaagaa aatatagacc aatattcctg ttagcataga 4740
catgcaacag ctaaccaatt ttagcaaacc aaacctggta atatagaaaa aaggataaat 4800
aggccagtog cgggtggctca cgctgtaat ccagcactt tgggaggctg aggcaggcag 4860
atcacttgag gtcaggagtt tgagaccagc ctgaccaaca tggtgaaacc ccgtttctaa 4920
taaaaataca aaaatcaggc tgggcacggt ggctcacgcc tgtaatcca gcactttggg 4980
aggccgaggt gggcagatca cgaggtcagg agttcaagac cagcctgacc aatgtggtga 5040
aacgccatct ctactaaaaa tacgaaaatc agccggtgtg gtggcacctg cctgtaatcc 5100
cagctactca ggaggctgag gcagaattgc ttgaaccgga gaggcagagg ttgcagttag 5160
ccaagatcgt gccactgcac tccagcctgg ggcacagagc aagacttcat ctcaaaaaa 5220
aaaaaaatta gctgggcatg gtggtgggca cctgaaatcc cagctactcg ggagtctgag 5280
gcaggagaat cgcttgaacc caggaggcag aagttgcact gagctgggat cacaccattg 5340
cactccagcc tgggcaacag agtgagactc catctcaaaa aaagaaaaag aaaaaggata 5400
aatacattct aaccaaataa tgtttatctc atgattgtag ctgattcaac attcaaaaat 5460
tggcctggtg cagtagctca ggcctgtaat cccaacattt taggaggctg aggcaggaag 5520
atctcttgag ccaggaattt caagaccagc ctgggcaaca tagtcagact ggtctttact 5580
ggggggaaaa aaatcagtct gtgtaattca ccacattaac aaagggaac ataaaaaccc 5640
tatgatcatt tcaacagatg tagcaaaagc agttaatgat atcaacacat atgcatgatt 5700
acaaaccaac caacctccta gaaactagg gaaaggaaac ttaactagtt tgataacagg 5760

gcgtccacag tcggagttcc actagcagca tacataatgg tagaaaactc agtgctgctg 5820
ggggcggttg ctacgcctg taatgccagc gctttgggag gcctaggcgg gcggatcacg 5880
aggtcaggag atcgagactg tcctgactag catgctgaaa ccccgctctc actaaaaata 5940
caaaaacaaa aaattagccg ggcatggttg cgggcgccta tagtgccagc tactogggag 6000
gctgaggcga gagaatggcg tgaaccggg aggcggagct tgcagagcct agatcgtgcc 6060
actgcactcc agcctgggtg acagagttag acttcgtctc aaaaaaaaaa aaaaaaaaaa 6120
aagaaaagaa aactcaacgc tttttcctct aagatcagga actagaaaag gatttgactc 6180
tcacaacgtt gataccatac tggagggtttt aaccaggcaa gaaaaagaaa taatgagggc 6240
cgggtgcggt ggctcaggcc tgtaatccca gcactttggg aagccgagac gggtggtatca 6300
cgaggtcagg agatcgagcc atcctggcta acacggtgaa accctgtctc tactaaatat 6360
acaaaaaatt agccgggctg ggtggcgggc gcctgtagtc ccagctactc gggaggctga 6420
ggcaggagaa tggcgtgaac tcagggggcg gagcttgagc tgagctgaga tcgagccact 6480
gcactccagc ctgggcgaca gagcaagact gtgtctcaaa aaaaaaaaaa gaaaaagaaa 6540
taatgattag tggcccgatg tctcaogcca gtaatcccag cactttggga ggccgaggtg 6600
ggcagatcac ctgaggtctg gagttggaga ccagcctgac aaagatggtg aaacctcgtc 6660
tctattaaaaa tattaaaaaa atagccaggc gttggccggg tacagtggct catgcctgta 6720
acccagcac tttgggaggc cgaggtgggt ggatcacctg aggtcaggag ttcaacacca 6780
gcctggccaa catggtgaaa ccccatctct actaaaaata caaaattagc cgggcgtagt 6840
ggcgggcgcc tgtaatccca gctacttggg aggcttaggc aggagaatcg cttgaacctg 6900
ggaggcggag gttgtagtga gccgagattg caccattgca ctccagcctg ggtgacaaaa 6960
gcaaaaactc cgtctcaaaa aaaaaagaat tagccagggg tagtggtgaa cgcctgtagt 7020
cccagctact caggaggcag aggcaggaga atcacttgaa ccccgaggc agaggttgca 7080
gtgagccgag attgtcccat tgcactccag cctaggcgag aagagcaaaa ttccatgtca 7140
aaaaaaaaaa aaaaaaagga aagaaaaaaa ataacgatta gaaaggaaga aatcaaacac 7200
attcacagcc agtatgattc tatacatacc atggtcctaa tggggccagg cgtggtggct 7260
catgctgtaa tcctagcaat tttaggaggc tgaggcagggt ggcttccctg ggaccagctg 7320
gccaacatgg tgaaacccca actctaataa aaatacaaaa aatcagccag gcgtggtgag 7380
ggcacctcta atcccagcta ctcaggaggc tgaggcagga gaattgcttg gacctgggag 7440
gcagagggtg cagtgagcog agatcgcgct attgcactcc agcctgggca acaagagtga 7500
aactccggca ggggtgtggtc ttacgcctgt aatcccagca cttcgggagg ctgagccagg 7560
ccgatcacct gaggtcagga gtttgagacc aacctaacat ggtgaaaccc cgtctctact 7620

aaaaatacaa gaattagctg ggtgtagtgg tgggcgcctg taatcccagc tacttgggag	7680
gctgagacag aagaattgct tgaaccagg aggtggaggt tgcagtgagc tgagatcatg	7740
ccattgcaca ccacgccggg caacagagcg agattccgtc tcaaaaaaaaa aaaaaagatg	7800
aaactctatc tcaaaaaaaaa aaaaaagtcc taatggaaaa tccataaaaa gctaccaaaa	7860
ctaataaata aatatagcag ggttgcaggt tacagggcaa tatagttatc cctctatctg	7920
taggggcttg gttctgggac tcctcacaca ccaaaccac agatgtctaa gtcccatata	7980
taagacggaa tagtatttaa cctacacata tcctccata tagtttaaat tatctagatt	8040
acttacatta ccccatata atgaaaatgc taatgtacat gcaagtatgt atgtaagtac	8100
ttgtactata ttgtttaggg aatcactgga cagataggcc ttcaagactg ataccagcag	8160
ccactgttaa gattctgtc aggcctgcc ctgtttggg tctcagttga totcattgcc	8220
ttcccacca gccagggca cctgcatttc tcttggtcc ctggccattt ggaaggccta	8280
gttcagcctg gcacatttgt atcctggccc actgatgctg gtaccctgg gaaggctctg	8340
ctctgaaaaa cacggagatt ttagttgcta ctgaagattt gagagataaa gacagggaga	8400
cctgtctgta gacctgtgtc cctccaagtg ggattgagac tttgggcccc ccatttcagg	8460
acagcacctc ctggcctgtt gactgaatag atccctgaag gaggtgtagt tgcattttag	8520
gagtgggggt gggagcagta ccactgatcc gactaaca tccacagtt ctctctagaa	8580
taataatata gaacaagtga aatagaacaa ttgcagaaag agctaacctt tgttgagctc	8640
ttactgtgtg ccagcactt tcctcaactc tacatttccc ataatacata gagtactagg	8700
taggcggggc ttgggggctc aogcctgtaa tcccagcact ttaggaggcc aaggggggtg	8760
gatcacctga ggtcgggagt tcaagaccag cctgactaac atggtgaaac cccgtctcta	8820
ctagaagtac aaaattagcc aggtgtggtg gcacatgctt gtagtcctag ctactcagca	8880
ggctgaggca ggagaatcat ttgaatccgg gaggagggtg cagtaagcgg agatagtgcc	8940
actgtactcc agcctgggca ataagagctg agactccgtc tcaaaataaa ataaaataaa	9000
ataaaataaa ataaaataaa ataaaaaaag aaaagagcct gccattaaag gagctgtttg	9060
gtaggggatg ttttgtcagt gcaaacaaca gaaaagtggg ctgggcacag tggttcatgc	9120
ctgtaatccc agcactttgg gaggccaagg cgggcggatc acctgaagtt gggagttcaa	9180
gaccagcctg accaatatgg agaaaccccg tctctactaa aaatacaaaa ttagccgggc	9240
gcagtggccg atgcctgtaa tcccagctac togggaggct gaggcaggag aatcgcttga	9300
acctgggagg cagaggttgc ggtgagccga gatcgacca ttgcactcca gcctggacga	9360
gagcaaaact ctgtctcaaa aaaaaaaaaa aacagaaaag tgtaacaaac acttacagta	9420

ggcatgtttc ttagcaaadc tgatgacaaa tttggcataa agaaagagag catccctgaa 9480
aaaaaaaaaa agaaaaagaa agagagcatc ctgcctgggc aacatagtga aaccctgcct 9540
ctacaaaaaa actcaaaaat tggccgggtg cagtggctca cacctgtaat ccagcactt 9600
tgaggagtcg aggggggagg atcacctgag gtcaggagtt cgaaaccagc ctggccaaca 9660
tggaacaaac ccatctctac taaaaataca aaaaattaat caggcgcatg ggtgggcgcc 9720
tgtaatccca gctactcagg aagttgaggc aagaggatcg cttgatactg ggaggtggag 9780
gttacagtga gtcgagatca caccactgca ctctagcctg ggtgacaggc cgagactccg 9840
tctccaaaaa aaaaaagaaa aagaaaaaga ctaaaaaatt agccaggcag gcctctgtgg 9900
tcccagctac ttgggaggct gaggcaggag aatcactgag ccaggagtg gcaggctgta 9960
gtgagccatg attgcaccac tgtaccctag cttgggcttc aaagcaagac cctgcctcaa 10020
aagaaaaaag aaagaaagaa agaacatggc gggccaggca cagtggctca cacctgtaat 10080
cccagcgctt tgagaggccg aggcagggtg atcacaaggt caggagttcc acaccagcct 10140
ggccaacatg gtgaaccct gtctctacta aaaatacaaa aaatcagcag gcagggtggt 10200
aggggcctgt aatcccagct actgggagg ctgaggcagg agaattgctt gaaaccagaa 10260
ggcagagggt gcagtgagcc tagactgcac cactgcactc cagcctgggc gaaaagagcc 10320
aaactccatc tcaaaaaaca aacaaaaaaa caaaacaaaa aaaaacatgg cagcctttga 10380
aagcttgtct gggagaaggt gcgatgatgg ttgcataact tcgtgcaaga tgctggtcca 10440
cacaggggct gccccttgc ctttctcgct ctottaacct ctcatataac aggccttgtg 10500
gttatgcaca tttattgagc ccaagcagg gcaaggcatt gtgatctaact actttggtca 10560
gcaagacaac aagatagatc actgcctgc ccttaggaag tgtatatgct attagaggaa 10620
acagataaaa taaacaagga aaagtatcag acaatgtaag tgctatgaga atgcaaatga 10680
ggtgatgtga attaaaatag gatgacttaa gtctgcacgg aaggcccta ccccatggt 10740
cctggctagc caaggaacca ccagttgatt agcagagaag ggcagcccg ctagctagag 10800
cttttgggga agaggagtg gttgttaaga gatgagatta aagaagccga gacgggcct 10860
tcgtgagggg gggttgtaat gcagggtga ggagtgtccg aagagaatgg gcaggtagc 10920
ggtgagacag ttgttcttcc agaagcttg cagtgaagg aatcaaagaa atggagccgt 10980
gtatcagggt ggggaagggt ggggccaagg ggggtgcctt cccatacag agattgcagg 11040
ctgagaatga ctatatcctt gttacagga ggtgggagca gggcacggt gctcacacct 11100
gtaatcttg cacttttagga ggcggaggcg ggccgatcac ctgaagtaag gattogaga 11160
ccagcctggc caacatgcaa agccctgtct ctactaaaaa taaaaaatt agctgggtgt 11220
ggtggtactc gcctgtaatc ccagctactc gggagactga ggcaggagaa tggcttgaa 11280

ccggaaggta gaggttgacg tgagctgaga tcatgccact gtgctccagc ctaggtgaca 11340
gagagagact ccatctcaaa aaaaaaaaaa aatacaggaa gggagttggg aataggggtgc 11400
acatttagga agtcttgggg atttagtggt gggaaggttg gaagtccttc tctgattgtc 11460
ttttcctcaa agaagtgcac ggctggtgtg ggggtggggca ggagtgcttg ggttggtgtg 11520
aaacattgga agagagaatg tgaagcagcc attcttttcc tgctccacag gaagccgagc 11580
tgtctcagac actggcatgg tggtggggga gggggttcct tctctgcagg cccaggtgac 11640
ccagggtttg aagtgtctca tgctggatcc ccacttttcc tcttgacgca gccagactgc 11700
cttcgggttc actgccatgg aggagccgca gtcagatcct agcgtcgagc cccctctgag 11760
tcaggaaaca ttttcagacc tatggaaact gtgagtggat ccattggaag ggcaggccac 11820
caccgccacc ccaaccccag ccccttagca gagacctgtg ggaagcgaaa attcatggga 11880
ctgactttct gctcttgtct ttcagacttc ctgaaaacaa cgttctggta aggacaaggg 11940
ttgggctggg gacctggagg gctggggggc tggggggtg aggacctggt cctctgactg 12000
ctcttttcac ccatctacag tcccccttgc cgtcccaagc aatggatgat ttgatgctgt 12060
ccccggacga tattgaacaa tggttcactg aagaccagg tccagatgaa gctcccagaa 12120
tgccagaggc tgctccccgc gtggcccctg caccagcagc tcctacaccg gcggcccctg 12180
caccagcccc ctctggccc ctgtcatctt ctgtcccttc ccagaaaacc taccagggca 12240
gctacggttt cgtctgggc ttcttgcatc ctgggacagc caagtctgtg acttgacagg 12300
tcagttgcc tgaggggctg gcttccatga gacttcaatg cctggccgta tccccctgca 12360
tttcttttgt ttggaacttt gggattcctc ttcacctta ggcttcctgt cagtgttttt 12420
ttatagttta ccacttaat gtgtgatctc tgactcctgt ccaaagttg aatattcccc 12480
ccttgaattt gggcttttat ccatcccatc acacctcag catctctcct ggggatgcag 12540
aacttttctt tttcttcac cactgtatt ccttggcttt tgaaaataag ctctgacca 12600
ggcttggtgg ctcacacctg caatcccagc actctcaaag aggccaaggc aggagatca 12660
cctgagcccc aggagttcaa gaccagcctg ggtaacatga tgaaacctg tctctacaaa 12720
aaaatacaaa aaattagcca ggcattgttg tgcacacctc tagtcccagc cactcaggag 12780
gctgaggtgg gaagatcact tgaggccagg agatggaggc tgcagtgagc tgtgatcaca 12840
ccactgtgct ccagcctgag tgacagagca agacctatc tcaaaaaaaaa aaaaaaagaa 12900
aagctcctga ggtgtagacg ccaactctct ctagctcgct agtgggttgc aggaggtgct 12960
tacacatggt tgtttctttg ctgccgtgtt ccagttgctt tatctgttca cttgtgccct 13020
gactttcaac tctgtctcct tcctcttcct acagtactcc cctgcctca acaagatggt 13080

ttgccaactg gccaaagacct gccctgtgca gctgtggggtt gattccacac ccccgcccg 13140
 caccgcgctc cgcgccatgg ccatctacaa gcagtcacag cacatgacgg aggttgtgag 13200
 gcgctgcccc caccatgagc gctgctcaga tagcgatggt gagcagctgg ggctggagag 13260
 acgacagggc tggttgcccc gggccccag gcctctgatt cctcactgat tgctcttagg 13320
 totggcccct cctcagcatc ttatccgagt ggaaggaaat ttgcgtgtgg agtatgttga 13380
 tgacagaaac acttttgcac atagtgtggt ggtgccctat gagccgcctg aggtctgggtt 13440
 tgcaactggg gtctctggga ggaggggtta aggggtggtt tcagtggccc tccgggtgag 13500
 cagtaggggg gctttctcct gctgcttatt tgacctccct ataaccccat gagatgtgca 13560
 aagtaaattg gtttaactat tgcacagttg aaaaaactga agcttacgag gctaagggcc 13620
 tcccctgctt ggctggggcg agtggctcat gcctgtaatc ccagcacttt gggaggccaa 13680
 ggcaggcgga tcacgaggtt gggagatoga gaccatcctg gctaacggtg aaaccccgctc 13740
 tctactgaaa aatacaaaaa aaaattagcc gggcgtggtg ctgggcacct gtagtcccag 13800
 ctactcggga ggctgaggaa ggagaatggc gtgaacctgg gcggtggagc ttgcagttag 13860
 ctgagatcac gccactgcac tccagcctgg gcgacagagc gagattccat ctcaaaaaaa 13920
 aaaaaaaaaa gcctcccctg cttgccacag gtctcccaa ggcgcactgg cctcatcttg 13980
 ggctgtgtt atctcctagg ttggtctga ctgtaccacc atccactaca actacatgtg 14040
 taacagttcc tgcattggcg gcatgaaccg gaggccatc ctcaccatca tcacactgga 14100
 agactccagg tcaggagcca cttgccaccc tgcacactgg cctgctgtgc ccagcctct 14160
 gcttgccgct gaccctggg ccacactctt accgatttct tccatactac taccatcca 14220
 cctctcatca catctccggc gggaatctcc ttactgtctc cactcagttt ccttttctct 14280
 ggctttggga cctcttaacc tgtggcttct cctcccacct cctggagctg gagcttaggc 14340
 tccagaaagg acaaggggtg ttgggagtag atggagcctg gttttttaa tgggacaggt 14400
 aggacctgat ttccttactg cctcttgctt ctcttttct atcctgagta gtggtaatct 14460
 actgggacgg aacagctttg aggtgcgtgt ttgtgcctgt cctgggagag accggcgcac 14520
 agaggaagag aatctccgca agaaagggga gcctcaccac gagctgcccc caggagcac 14580
 taagcgaggt aagcaagcag gacaagaagc ggtggaggag accaaggtg cagttatgcc 14640
 tcagattcac ttttatcacc tttccttgcc tctttctag cactgcccc caacaccagc 14700
 tcctctcccc agccaaagaa gaaaccactg gatggagaat atttcacct tcaggtacta 14760
 agtcttggga cctcttatca agtggaagt ttccagtcta aactcaaaa tgccgttttc 14820
 ttcttgactg ttttacctgc aattggggca tttgccatca gggggcagtg atgcctcaaa 14880
 gacaatggct cctggttgta gctaactaac ttcagaacac caacttatac cataatatat 14940

atlttaaaagg accagaccag ctttcaaaaa gaaaatagtt aaagagagca tgaaaatggg 15000
tctatgactt tgccgtgatac agatgctact tgacttacga tggagttact tctgataact 15060
cgctgtaagt tgaaatattg aaatattgta agttgaaaat ggatttaata cacctaactt 15120
aaggaacatc atagcttagc ctagcctgct tttttttttt tttttttttt ggagacagag 15180
tctcactctg ctaccaggc tggagtgcag tggcgggac tcggctcact gcaacctccg 15240
ccttctgggt tcaagcgatt ctctgcctc agccactga gtagctggga ttacaggcac 15300
ctgccccgac gccagctaa ttttttgta tttatttctt ttttttttag tagagataga 15360
atltcaccat gttggccagg ctagtctcga actcctgacc ttgtgatctg cctgccttgg 15420
cctcccaaag tgctgggatt acaggcgtga gccaccgcac ctggcctgcc tagcctactt 15480
ttattttatt tttaatggag acagcatctt gctctgttgc ccaggctgga ttacagtgat 15540
gtgatcatag ctcatatac cctcctgggc tcaagcaatc cccctaactc tgccctccca 15600
gtagctagga ccacaggcat acaccaccat accagctaa tttttaaaat tttttgtaga 15660
tagatagagt ctactatgt tgccaggct ggtctctagc ctactttttt gagacaagg 15720
cttgctctgt caccaggct ggatagagt cagtagtgca gtcacagctc actgcagcct 15780
ccacctccca ggctccatcc atcctcccag ctcagcctcc caagttgctt caactacagg 15840
cctgcaccac catgcctggc taatttttat ttattttatt ttattttatt ttattttatt 15900
ttttgagact cagtctcact ctgtcgcctt aggttgagg gtagtggcat gatctcggct 15960
cactgctaac ctctgcctcc tgggtttcaa gtgattctcc tgccctcagcc tcccgaatag 16020
ctaggactac aagcgcctgc taccacgccc ggctaatttg tgtattttta gtagagacag 16080
ggtttcacca tgttgccag gctggctctg aacttctgac catgtgatcg ccgcctcggc 16140
ctcccaaagt gctgggatta cagggtgtgag ccaccacgcc cggctaattt ttattttatt 16200
atltaaagac agagtctcac tctgtcactc aggttagagt gcagtggcac catctcagct 16260
cactgcagcc ttgacctcc tgggctccg tgatttcacc ctcccaagta gctaggacta 16320
caggcacatg ccacgacacc cagctaattt tttattttct gtgaagtcaa ggtcttgcta 16380
cgttgcccac gctggatatca aacctctgg ctcaatcaat ccttccacct cagcctcccc 16440
aagtattggg gttacaggca tgagctacca cactcagccc tagcctactt gaaacgtgtt 16500
cagagcattt aagttaccct acagttgggc aaagtcatct aacacaaagc cttttttata 16560
gtaataaaat gttgtatatc tcatgtgatt tattagatat tgttactaaa agtgagaaac 16620
agcatggttg catgaaagga ggcacagtcg aagccaggca cagcctgggc gcagagcgag 16680
actcaaaaaa agaaaaggcc aggcgcactc tcacgcctgt aatccagca tttcgggagg 16740

ctgaggcggg tggatcacct gaggtcagga gttcaagacc agcctagcca acatggtgaa 16800
accccgctct tactaaaaa caaaaattaa ccgggcgtga tggcagggtgc ctgtaatccc 16860
agctacttgg gaggtgagg caggagaatc gcttgaacca ggaggcggag gttgcaggga 16920
gccaagacgg cgccactgca ctccagcctg ggcgatagag tgagactccg tctcagaaaa 16980
aaaagaaaag aaacgaggca cagtgcgatg cacatgtagt ccagttact tgagaggcta 17040
aggcaggagg atctcttgag cccaagagtt tgagtccagc ctgaacaaca tagcaagaca 17100
tcatctctaa aatttaaaaa agggccgggc acagtggctc acacctgtaa tcccagcact 17160
ttgggagggtg gaggtgggta gatcacctga cgtcaggagt tggaaaccag cctggctaac 17220
atggtgaagc cccatctcta ctaaaaacac aaaaattagc cagtgtgaga cacgttgagt 17280
ccacgtactc ggaggctgag gcacaagaat cacttgaacc ccagaggcgg agattcgaat 17340
cagccaagat tgcaccattg cactcccgcc tgggcgacga gagtgaacc ccatctcaa 17400
ataaataaat aaatatTTTT aaaagtcagc tgtataggta cttgaagtgc agtttctact 17460
aaatcgatgt tgcttttgat ccgtcataaa gtcaacaat tgtaacttga accatctttt 17520
aactcaggta ctgtgtatat acttacttct cccctcctc tgttgctgca gatccgtggg 17580
cgtgagcgtc togagatggt ccgagagctg aatgaggcct tggaactcaa ggatgcccag 17640
gctgggaagg agccaggggg gagcagggt cactccagggt gagtgaacct agccccttcc 17700
tggccctact cccctgcctt cctagggttg aaagccatag gattccattc tcatcctgcc 17760
ttcatggtca aaggcagctg acccatctc attgggtccc agccctgcac agacattttt 17820
ttagtcttcc tccggttgaa tcctataacc acattcttgc ctccacgtag tatccacaga 17880
acatccaaac ccaggacga gtgtggatac ttctttgcca ttctccgcca actcccagc 17940
ccagagctgg agggctctca ggggcctaata aattgtgtaa tactgaatac agccagagtt 18000
tcaggtcata tactcagccc tgccatgcac cggcagggtcc taggtgaccc ccgtcaaact 18060
cagtttcctt atatataaaa tggggtaagg gggccgggag cagtggctca cgaatccac 18120
actctgggag gccaaggcga gtggatcacc tgagggtcggg agtttgagcc cagcctgacc 18180
aacatggaga aacccatct ctactaaaaa taaaaagta gccgggcgtg gtgatgcatg 18240
cctgtaatcc cagctaccta ctgggaggc tgaggcagga gaatcgctt aaccgggag 18300
gcagaggttg cgggtgagctg agatctcacc attacactcc agcctgggca acaagagtga 18360
aactccgtct caaaaaagta taataaagta aaatggggta agggaagatt acgagactaa 18420
tacacactaa tactctgagg tgctcagtaa acatatttgc atggggtgtg gccaccatct 18480
tgatttgaat tcccgttgtc ccagccttag gcccttcaaa gcattggtca gggaaaaggg 18540
gcacagacco tctcactcat gtgatgtcat ctctctccc tgcttctgtc tctacagcc 18600

acctgaagtc caaaaaggggt cagtctacct cccgccataa aaaactcatg ttcaagacag 18660
aagggcctga ctcagactga cattctccac ttcttggtcc ccactgacag cctcccaccc 18720
ccatctctcc ctccccctgcc attttggggtt ttgggtcttt gaacccttgc ttgcaatagg 18780
tgtgcgtcag aagcaccacag gacttccatt tgctttgtcc cggggctcca ctgaacaagt 18840
tggcctgcac tgggtgttttg ttgtggggag gaggatgggg agtaggacat accagcttag 18900
attttaaggt ttttactgtg agggatgttt gggagatgta agaaatgttc ttgcagttaa 18960
gggttagttt acaatcagcc acattctagg taggtagggg ccacttcac cgtactaacc 19020
aggggaagctg tccctcatgt tgaattttct ctaacttcaa ggcccatatc tgtgaaatgc 19080
tggcatttgc acctacctca cagagtgcac tgtgaggggt aatgaaataa tgtacatctg 19140
gccttgaaac caccttttat tacatgggggt ctaaaacttg acccccttga ggggtgcctgt 19200
tccctctccc tctccctgtt ggctgggtggg ttggtagttt ctacagttgg gcagctgggt 19260
aggtagaggg agttgtcaag tcttgcctggc ccagccaaac cctgtctgac aacctcttgg 19320
tcgaccttag tacctaaaag gaaatctcac ccatccac accctggagg atttcatctc 19380
ttgtatatga tgatctggat ccaccaagac ttgttttatg ctcaggggtca atttcttttt 19440
tctttttttt tttttttttt ctttttcttt gagactgggt ctgcctttgt tgcccaggct 19500
ggagtggagt ggcgtgatct tggcttactg cagcctttgc ctccccggct cgagcagtc 19560
tgcctcagcc tccggagtag ctgggaccac aggttcatgc caccatggcc agccaacttt 19620
tgcatgtttt gtagagatgg ggtctcacag tgttgcccag gctggtctca aactcctggg 19680
ctcaggcgat ccacctgtct cagcctcca gagtgcctggg attacaattg tgagccacca 19740
cgtggagctg gaagggtcaa catcttttac attctgcaag cacatctgca ttttcacccc 19800
accttcccc tcttctccc tttttatata ccatttttat atcgatctct tattttacaa 19860
taaaactttg ctgccacctg tgtgtctgag ggggtgaacgc cagtgcaggc tactggggtc 19920
agcagggtgca ggggtgagtg aggaggtgct gggaagcagc cacctgagtc tgcaatgagt 19980
gtggactggg gggccagtg cccgggttcc gggaggggaa caaaggctgg agactgggtc 20040
agtctgcggg ctgcatgaca acaaggaggg ggggtggctcc attcataact caggaaacca 20100
ccgtccctcc tcccctccgc ccacggctgg cacaagggtc tctgcctccc ctgcttctag 20160
gattgggctg cttccccctc ggcagcctct caccaaggat tacgggattt aaatgtcgtg 20220
atttacgaag gctgagcctc cagggtggcc atcttcgtcc atcagaagtg gcaggatacc 20280
tgggttccaa gggaacaggg tgg 20303

<210> 4
<211> 44505
<212> DNA
<213> Homo sapiens

<400> 4
ttggtaatgg aacagagaaa tacaaatata gaaacgtttt tatccttttag atcttctaata 60
ggtgcagata tttcttaacc aatttcaaga gtgccttatt caaaaacaaa aaacctgcac 120
atagaaaaag caaacatggt tttttaaaaa aagccctgta aaagtggta atctgtcgtc 180
tgattttcaa tgtatgactt taatcaattg catattcctt aattctgcaa atgacaatta 240
aaattaggaa tcagtacctg aaaaogccat ttatactttt gagtttatat tagaaacacg 300
gttttgatga aatacctttt tcggttcaat ctttttgaat tgtgaccaca accaatagct 360
acacttagat totacagaat tagtaggaaa gatacaaatg agaaggtott tctagacaat 420
ggatctcact gaacatcatc ttagtgtgta gactttttcc actgcagtgg ccacatgtg 480
ctatggtgac agacggcagt tggcattacc acttatatag taaaatatct ttaagacaaa 540
aataaaaagt atttaagaa tacgctggcc ctgaaacatg agctgtgcct taaactacat 600
atcctacctc catgagtga tgaataagt ggttttccgc ctgcatctat gtgttaagac 660
ctattgtttg cttgtattaa ttgtagtcct gggataaaag cacagaggta actttcactg 720
ctgctttttg tactctctcc aatgttttgg aggaaaaata agcacaaca atagcctaga 780
gaaactgaat cgatcact tgttgacgt tcgacaaacg tcaattttgc tgcattagaa 840
tgggaaacat ttttcagtct attgaaatga attacaaacg tttttaaata gaaaattagt 900
taaaaaattg gaggctgctt aatcgatagc tttctctata aacatacttg gatttcacaa 960
ataagtaata ccgtaaaaat cttcttctcc aaagaaaaat cccacaaaat aaatctattg 1020
atacctagtg acaagtggaa ccagataaaa atggaatcta taagaattaa cctaattgac 1080
agcgtcttg agctaatacca tttccattag ttatttgttc acagtaggta ctcttaagga 1140
cttgttgaat tgcgggcttg gcgcccgttc tacggagagt tcacagcctt cgtgagtggg 1200
gacagaaggc ggctcggcc ggtgattcag gtcgaaatc aagctgaaca gcctgctgag 1260
aggtgggato caccatccgg acagtgggg gctttgggg tgctgtgaga ctgggctgag 1320
accaggtcc agcaggaggt gtgcggcaca gaccacaagg tcggcgaggc ccctaacco 1380
gcgcccggcc gggaaccgc agaccagcga cggggcagct gcggggccag gagcgccca 1440
agacgggagg gcgctgaacc cgagcccctg ccgcccgcct ggcccgaac ttccgcctc 1500
ccaggacctg tcccggcgc cccgagcgtt actcgaaggc cggggccgag atgccacctt 1560
ccgcaggccg cgggaaaggc gcgccgagtc ctgcagctgc tctcccgtt cgggaaacgc 1620
gcggggcggg ggcgtcgggc ttgggacagg ggaggatacc agggccacct tcccaacco 1680

aggccgcggg	ggcccggcct	ccccgatgca	gaccacagcg	ccctcacggg	ctgccctcag	1740
gccgcgcagc	gggcagccgc	cagccgtcac	cccggggagc	gtccgtgggg	tgcccaggca	1800
ccccaccccg	gcccggggcg	ctcagacggc	agcagactgc	tgggcggcgc	ggggactact	1860
ttccaccgcc	ccctcgcgcc	ccgccccttg	tcctcgcgcg	gcggaacgct	ccgcgctgcg	1920
ccggtggcgg	caggatacag	cggttctctg	gcgacttata	agagctcctt	gtgcggcgcc	1980
attttaagcc	tctcggctctg	tggcagcagc	gttgcccccg	cccggggagc	ggagagcgag	2040
gggaggcgga	gacggaggaa	ggtctgagga	gcagcttcag	tccccgccga	gccgccaccg	2100
caggctcagg	acggtcggac	tcccgcggcg	ggaggagcct	gttccctga	ggtgcttggg	2160
cgctcctttc	cttatccttc	cggggctgct	cccgttctct	ctcggagcca	aacttcgtag	2220
caggcgcgcg	gtccggggcg	cgggctgggc	gcagccggga	ggcctggggg	tgggagcggg	2280
gagctcaggt	gggggacggg	gaggggtggc	cgcgcccggg	gcgcggaggg	cggcggccgg	2340
gcccgggttc	cggtcgcgct	gcctctcttg	ggccctgggg	gcatcgcttg	cggggagggg	2400
gcgcgcggcg	ggcgcgta	ggagcccgga	tggcaggcgg	ggtgggggtg	gggggtgggg	2460
tctgtggttt	ccgtccgggg	ctctggcctt	ggccgagttt	gggggagggg	cccgggtgcct	2520
cgggatgcgc	cgggcccttg	gtggggggcg	gggtggggac	ggggggctcc	gccttctcag	2580
ctcttgcggc	gagttggggg	tcgggcgctg	aggcagagac	gccaccctaa	gtcccatcag	2640
tcctggggat	cggaccagtg	gactttctct	taagatttct	tctttcattc	ttaagaatag	2700
aagtgttatt	atTTTTTTTa	atgccctggc	tatgtgagtt	tgaatcgaag	caactttaaa	2760
ccttagagca	actaaactct	aagtgcagcg	ggtgcgatgc	gtcagtaggg	tgagcacata	2820
aaaaatccat	gtcttgacc	tgtatttttag	cgtactatgc	aggtgagtga	aagcagtgga	2880
taatgtactg	ggagtcttat	ggatttatgg	tagtgggtat	gagaccctgg	tgaataagg	2940
gggtggagga	aggcgaaggt	gatggcttac	tgtttcttac	caagtgaact	gcaggattca	3000
gcctctgact	cagaccgctt	cgagaatttt	gttcgtagaa	ataattttaaa	tttattcaaa	3060
tagtttgatg	gcagctaaaa	ttgaattata	gagcacgttt	tcttttcagc	ggagtgaatt	3120
tttccttcgc	tccaaagctg	gccaaatgga	attcaagcat	tgcaacttct	ttcagtgttt	3180
tgtctggaga	gaggactttg	aaccgagact	tttogaagtt	aagttcctat	agcctgcttc	3240
tgaatctgcc	aagcttgaaa	gctttggcag	ttgggtgtat	gtagttgttg	ccttcgttct	3300
cttccctttt	ggaggggagc	ttgtctccta	ctttgtatct	tccagacatc	tgtggtcttc	3360
ccccacccc	togagtttgt	gagtggtgaa	tgaagaaaga	ctaggctgct	ggtatgcaga	3420
ggtcggcaaa	aggaaatcga	ggagtggttt	tagtgaaatg	agagctttgt	atcatgaata	3480

atgggtggcctt aggctagaca tcaacttgaa gagacggcag catttccttt cataaagtct	3540
aggctaattgt ttttcagatc gctaagttgt agtttgtctg gaatttagga agccatttca	3600
gtatttgtca cttggtgaac gaacattcaa taccttcaga tgtcttcgtg ttgacttgta	3660
ttcatcctaa gaaatagtaa atatagtctc aagtgttatt tatgttatac tgctggttta	3720
ttctctgctt aaattattga cataaatttc tactttggag gcttttcgtt tgaactaagg	3780
ctgtgcggaa tttattttac ttttatattt aaatctttga aaaatctctg attaaaaaa	3840
aagtaccctt aaaggtttga ggatgtcctt tcacaccaga caaaatttgg ttaatttgcg	3900
cccaatattc attactttga cctaacctt gttctgaagg ccgtgtacaa ggacaaggcc	3960
ctgagattat tgcaacagta acttgaaaaa ctttcagaag tctattctgt aggattaaag	4020
gaatgctgag actattcaag tttgaagtcc tgggggtggg gaaaaataaa aaacctgtgc	4080
tagaaagctt agtatagcat gtaactttag agtcctgtgg agtcctgagt ctcccacaga	4140
ccagaacagt catttaaaag ttttcaggaa aaaccaactt aaaaaaaaat aagggtggcta	4200
attaaaaaaa aatgaagcat ttaacagtgt tcaggtttca gagtatggaa gaggggtttt	4260
ttaaactggt atctgattat ttcttttacc aacatgatat agaaaagtgt atttccagta	4320
ttaaaattta tcagactgag cttactgttc ctgttaatga ctggaataaa aattggcata	4380
aatgagggtc tgtatgcttg ttttaataac accaccacca agatagaaaa cgaggaggca	4440
agtttctoca agggatattt gaaatgtgtt agcaaaacta ttgcagatac tcgtttttgt	4500
tataggggtga ggtggggaga ggogcatgct aagtattgtt gaaactaggg atgtagagaa	4560
ttaaaagttt gaatataatt attttgtagt tataagtagc agtgaaatta aatctcctgc	4620
aatagactat agaagtatat ttagccaaat gaaacttcag tgttattgaa atgaaataat	4680
acatctgtcc tgttacaaga ttatttttat ttctcttctg gtttcttagc ttctgataat	4740
caataattgt agatgagtag gtggtaagtt ttaagtttgt actttgagct tagtcggaag	4800
catgcttgac tgccaacccg gggcacaaag gatgaaggct tttagaactg gacaaacttc	4860
taacaaaagg tatttgcaac tctttttagt tgtgtcatgt tgatttgtga cattgttttt	4920
gaaaaatatgt gttaacttag ttttcttgta gccctctttt tattggaact gtggtatcta	4980
ttgttgaaac tgcttgactg agaacatttt tataccataa aagtaaatag taaacatagc	5040
ccaggagcgg cttctggttt gtccatcgta tgtagccatt gcctccttgt actctcattg	5100
agaagatact gatttgcaga ttcagttgtc cttctctaac agactattta tgtaatatgt	5160
cagttgtgat tgtgataggt aagtggacca gtcggttaaa ataaatactc aggtttcaca	5220
aaaggaaaat aatatgattt gtgttgatct aaatgagtat aggagttaac tcctatagtt	5280
tttcatcact taaactcagg ggaaagttct ttatttcctc tgtttactta agaatgctgc	5340

ttttgtgttt catgcaagac tgagcttgac tcagtttgaa acctaggctc atctgttgag 5400
gcctgaaccc tgctgtcctt gaagtatgca tataatttgc ttcttccta agggaaaata 5460
agctcttgaa agataaagtc aatcacatta ggaaccatt tttagggttt agccactttt 5520
tttttttttt ttttttaact catgggcac tcctctgtta agagacattc cccactctcc 5580
aagtttccct caagcctgaa gcagcagagt gagtagtggt ggagcatggt ttcattgcat 5640
gcttggttca tgttgagtgc cctccagtgg atatagtata atgcttgatga tttttttttt 5700
tttaattcca aacaagttaa tgtgggatat atttaggaat agttctgatg agggagaatc 5760
aactaagaaa cctttgattt ctaaaataat taatatcatt actgctaatt aaaatacagg 5820
cttgagaaaa tgtcttctca gccaatattt gcagtagaaa agtcgggagg ttttttaagg 5880
tcactttgag taggcagttc tgcttaaata tatcataatg ataaaccaga atctcagtat 5940
agtactttag gaggtaaaag atcataatat tcagttatat tgatgaatta cagcaactga 6000
aattctcaga aaaaaattaa tgaaaatgtg aattgtcaat ttgtctaaaa tcattcacag 6060
agtaaacat aagtgtcaa cttgattata ttaggaaata gatagaaata aaggtaattg 6120
agccagtgtg tgtgacctaa aatataatgc ccttagtgac cataggggtg gtctcatttg 6180
tacatagtgg tgggcoatga tgaactgtgt tttgcccttt gaatttttcc ttaaaaagct 6240
ttctctaggc tcctatgttc atggtttttc tgttagtaat attattttct gaaaatccat 6300
gtttcaaatc agaataat tagcaacagg aatgaagctt attctaaatt agtttttgga 6360
agttaaacgg tcagcatatg gaaatttttc agggtttaga tttttaaaaa tttgttttcc 6420
agaatatgtt gctggaatga aaacgttagc gtagggacgg aaaatgacac ttaccagtga 6480
ttgctttact ttgcctgtgg aattcagtggt aattttgtgg aaacattggt atatgatttt 6540
ttactactta agaaatgtat tgctatagtt agggtttttt tttttttaa ggcaagaatg 6600
cctcaagtgc tttatgtgaa tgattatttc aggatggatt aaatattcct ccatcaagga 6660
ccatacttgt aaatcagtga tttccaagtt ggtgcttagt atttacagca ttactgtct 6720
ataagcttct gttctgattt ttcaagagtt ttctgagaaa tgagagtagg cttaaaagtt 6780
ctttgaaaaa ttatgtacat acaacttact gaaaaaaatt gctaccgggg acttaatttg 6840
tctcttgaaa tgggctactt gccttcatta atgtagcata ctacaatttg atgttcaaga 6900
tatgttacta agaataagat cgctttcaga agccttatat aggattgggt ttactacatt 6960
gtagtgggaa tggctactca aatgtctcca gggccagtta ggtattgggt aaatgggacc 7020
atgcagacta ttaaaaattg aagtgcacat gaagcagcca gtcataagca gctccagcca 7080
ctgtgtggga atatagttaa tgttgccaga tcatctgatt tctttcccct aagtgggaaa 7140

tocagatcaa	tgtacatctc	ttgattttgca	agtgtttggtg	aacaaaattc	atatttttaag	7200
atgctgtatt	cagcacaaat	taaatacact	tattttgctga	atactgccag	tttgtccctc	7260
tgcagtagta	ccatttgaag	tacagtgttt	tcataatgat	tctgtgaaat	gactggttct	7320
gtgaatgtac	ataatttagc	agataacatt	gttaaattat	taggtttgta	tttatttagg	7380
cacttgggaa	atgccttgtg	tcaattgatt	atagattagg	agcttaaaag	caagatttat	7440
attatcaact	tattttgtgaa	gactgggaaa	cccacatttt	taaagttagg	aattaagatg	7500
gccaggttca	aggaaaaggg	ggagaagtaa	ctttcttatt	actcaaccat	cttaaataga	7560
gttctttaag	tgtattttta	agaggtctca	aaacttaatc	tgaagggacg	tcaaatgctg	7620
gacaaattct	gtgtatacaa	ctcaagtcag	cccccaattt	tactggtctt	taaatcatgt	7680
cctttttacc	agaagtttgc	atttctaagc	taaactatta	ctgttagact	agatccaaaa	7740
cttaaaaaca	gttttaggtaa	ttaaaaatta	attgaatata	aacgttttac	ttaaattaat	7800
ggcaaatggc	tttttgcca	atttaagttt	atgtaggcag	ttaaatcgat	tttggttaaa	7860
tcttttgctg	ctaacaaggt	atttcagat	tttgaaaagt	ggggtggcct	ggtgcctgta	7920
gtaccagcac	tttgggaggc	tggggagggt	ggatcacctg	aggtcaggag	ttogagacta	7980
gcctggccga	cgtggtgaat	acaaaaatta	gccaggcatg	gtggcagggt	cctgtgatcc	8040
cagctgcttg	gaagtctgaa	gcatgagaat	tgcttgaacc	tgggaagcgg	aggttgcagt	8100
gagctgagat	cacgcactg	cactccagct	ggggcaacag	agcgagactc	catctcaaga	8160
aagaaaagtg	gggtgttttag	tcttcaaact	ccgtgtttta	gtgactggag	tgaaaatgta	8220
aatcataggc	cgggtgttgg	ttaaaaagca	tcatctgaaa	ataatgctgt	agtctgcaat	8280
tattttttatt	acgatacgat	ggtgtaaaat	acaagcagat	cagtgaacca	ttcatgaaac	8340
attaatccta	aaggcgtctc	acccaagtc	tatcccacaa	totccatgag	acttctgtga	8400
accactgtaa	agtttcttgt	gtaatatccc	agaagtttcc	tacctctggt	atcttttgaa	8460
cttggtgaaa	aggcttttcc	acccctctt	tatgatgggt	tgaagagtgt	gaacatctga	8520
atgatgctgg	ggtgaaactg	cttcataaca	cttccatttt	ctccctatt	tatttccata	8580
tttttatttt	ttcactaata	tccccacgg	tttacttctg	tttttagtaat	tcacatgttg	8640
ctggactaat	tctttttaac	tgacttgtaa	cagatatggt	aaaccgttta	aaacttgggg	8700
ggtattttta	acctacttta	agttagtcca	agttaatcag	tctacatggc	atataaacct	8760
tatgattaat	aaatcttaaa	tgtgtgtagc	tgagttggaa	gccaaagacg	tacaaaaaag	8820
ctgaagtgtt	aggtttagtg	tgataagctt	ctcttactaa	cagggttttg	taatagcaga	8880
aatagatata	tgcataatata	tgtgcatata	tatagcatat	cttattggat	gtccatataa	8940
aaatgtgtaa	gaagttaa	ttactgcaaa	atttcttggg	agtgcaattt	gaagatgatc	9000

ttaagtgggtg atagtagttt gctacactgg gggatagttg ttgcaaactg ctcctaattt 9060
tcctttactg tgaagtaaac tgaacagctg taatagggat taggaactgt actccctctc 9120
tctctttttt aagtataatt aagtggtttt ggggtaaggg tgtagggagt gagtgtcttt 9180
gaagttttgc atatactaga tgaatgccac atgtataagg gaggaacaag ggattcttgg 9240
aaatattttt caatccaagt aactttggag gcttccaagt ggagttcatt cccctgtgta 9300
ggaaagtgct ggggtagacc cttaaattcc tttctgagcc attgaaagaa tgtcctcaaa 9360
cttcgcttat actttatagt tcatttagat acaaaagtta caaactgaat gctatttagg 9420
aaacgtaata cactgacata ccgctcttta aatagattat aaatttagta tatcaatttt 9480
ctggcatttt gctgaatttt attgtttagt tttcaagccc aactatcttg ttactttgta 9540
tatcgtagtt gtccccgtt gatcactgtt tcctgcttaa ttgtgtgtgc gtttttcctg 9600
ggctctgatt cagagtgtca gcattctgtt ccccatagaa taagaagagg ctagaaagtt 9660
tacagatgag atatctagga atgccagaag atcaggggtc accgttgagg cagagtaatt 9720
aattatgggtt aaaatgggtg tgctgataag tgggtgctgg gaaataatta aaatttgatt 9780
ttttagaaga atacttctca tgcttgaaga gcgcctcat tatatgctaa agggcctcag 9840
gtttttcctt attgccatta tgctgcagat totattacat ttgtctgaaa agatctaaga 9900
cagaaggggt gtttaatacc ttcccttttc tcctgaactt cccctctcct ctcccccatc 9960
aggagctaag taggaacccc ttacacttgt taocatcaga ttcatcaat ggtctgtctt 10020
tacaatgaag gaagtagtac tgcattctgg gcagaggcca gtcctgaggc atgccttttc 10080
aaggacattg ttactttagt tacactggct cttctgtttt aactottatc cccagactc 10140
taatcctgtt gctttttttg gtcccatct cccacctttc atcatctgaa atccattcat 10200
tgtaacttct ggaactcagt cgttagaaaa tcctttatat tctcaatctt gtgaatgttc 10260
ctttctttct tattccagct gtaacctagc cttctcccca agaatgctac ttcccttgca 10320
gctctctcaa gtggtgaatt tttcccttct tgcacacctt ataactga actaggaggt 10380
gtgtggacta aatgtctgt tttgttcctt attgtcactt cttgacctt attttccaaa 10440
acttcaagct ttgaactttca tgtgatcaaa ttataccacc cactgcctgt ctttatttca 10500
agcacctgca aaccttctg ggtcattcac atccttcttt gttcacttca ttagctcttg 10560
gctcattgtc actgtctctt atttctgtca taattcttgg tgacatcagt atctatgtag 10620
agcaatacta gtgaagatgt ggtctggtaa ctgttacctg tatgaattaa gataaggagt 10680
tatgccagaa tataagtcac ctgtgtcact aagtttactg tttagcttac tttttttgta 10740
gcaagatttt gatgaaggac gcaatatgtt gatttacagt ctggtacaaa ttttgatgta 10800

gaagatgctt ccaatatcct ggtctcttag ttcttggatt tcttctccag tgatcttatt 10860
ttctacccta actcaactac atattcccat tgtcatatcc tagaatattt tgtcttttat 10920
ctgtaactct gctctcttcc cccaatctca tttcaagcat ccactttct aattcctcta 10980
gtaaatacgt cagttccaac agcccatcaa tcccatggg acctacagtt tatctatcca 11040
agcttttccc tgttctcac cctcacttct atacagctga agtttcatac tgaattataa 11100
tcactttctc gtatacacgt ttaacaatct tgtccctccc tggcttcatg ccagtgatc 11160
tcttgatct atgaccatgt cctttatctt ctctctgtc actggatgaa ctgtagcctt 11220
ccaagataag gccactcagt tcatttgtag agcagattcc atccctctt gctctcaaga 11280
atattactgt ggtatctctc tttcttgtc tctactggct ctttccatga gcaaactagg 11340
tattatccca ttacaaaaa aattttttct cgtctctcc ttccactcac caoctcagtc 11400
tctgcttctc tttcccgcaa aataacctg aaaaattgct ttatgtactc cgttttctt 11460
ttgaaccct gccagtgacc accacgttat aaatttgtag ttgtcatctc acttaatctg 11520
ttagtagtat ttggcaccat tgctacagtt gcttgaaatg ctttttcatt ggtttccagg 11580
ccaccatgtc tgttagcagc ttttctctt acttcaactag catttccttc tttgtttttt 11640
ctgttatctt tctgacctct gttggagtgg ctgaagggtt agtccttgaa tctttttttg 11700
ttgtgcataat ttactccagt atcatagctt tatacagatg gtattttacat ctgtttgcta 11760
acgatttcca aattgggtatc cttaaactgg tatccagcta ttttttggtc agcattttgg 11820
atgtctaaga agcttctcaa actaaactga cctcccggtt ttccccaag ctgcatctta 11880
gtcttttccg aaatgcaatt ctgtctttcc agttacctag cttaaaagct tgcagttctt 11940
gactcatctt tctctcatac cacgtatctg aattctctct gcaaaaaatt gtctgttctc 12000
ccttcagaat aaagtcacgt gtcattttat gatggggata cattcagaaa tgcgtcatta 12060
ggagataatc atggttggtg gaacatcaga gtatacatag acaaacctag atggtatagc 12120
ctactacaca tctaggctat atggtgtggc caattactat gatgaatact gtaggtaatt 12180
gtaacataaa ggtaggtatt tttatctaaa cgtattgaaa catagaaaa gtacagtaaa 12240
aaatatggta tcaaaaaataa aaaatgggtac aactgtataa ggcagttgtg atgaatggag 12300
cttgaggat atgttgctct gggtagtca gcgagtgcg attgagggaa cgtgaaagat 12360
gtgggacatc actgtacact actgtagact ttataaacac tgtacacttg ggctacacta 12420
catttttgta aggtttttaa agactttttt ctataataaa ctttaaatta ctgtcacttt 12480
tttactttat gaattcttaa ttttttaaac gtttctoactc ttgtaataac acgtagctta 12540
aaacatacat tgtacagctg tacaaaaatt ttctttatat ctttataagc ttttttatat 12600
ttttaaaatt actttttacc ttttagcttt tttgttgaaa aactaagaca tgggccaggc 12660

goggtggctc acgcctgtaa tcccagcact ttgggaggct gaggcaggcg gatcacgagg 12720
tcaggagata agagaccatc ctggctaaca tgggtgaaacc ccgtctctac taaaaataca 12780
aaaaattagc cgggcgtgggt ggcgggcacc tgtagtccga gctacttggg aggctgaggc 12840
aggagaatgg cgtgaaccca ggaggcggag tttgcagtga gccgagatag cgccactgca 12900
ctccagtctg ggcgacagag cggaactoc gtctcaaaaa aaaacaaaca aaaaactaag 12960
acatgaacac attagcctag gcctacagag ggtcaggatc atcagtatca ctgtatttcc 13020
atctccacat cttgtccttc tggaatgtct tcagaggcag taaacataaa tggagctgcc 13080
acctcctgtg ataacagtgc cttctggaat acctcttgaa ggacctacct gtggctgttt 13140
tatagttaac tttttttttt taagaagtaa cagaaggagt acactctaata gataaaaagt 13200
atagtaagta cataaacctg taacaatcat tatcattatc aagtgtcatg tactggacat 13260
aactgtatat gctatacttt ttttttttga gatggcatct cactctgtca cccaggctgg 13320
agtgcagtgg tgcgaggata gctcactgta acctcagact cctggggtca agtgatcctc 13380
ctacctcagc ctoccaaagta gctgggacta caccaggcac cccaccatgc ctggctaatt 13440
aaaaaaaaatt ttttgtagag acagggctct actctgttgc cagggctggc cttgaattcc 13500
tggcatcaag taatcctccc actttggcct cacaaagtgc gaggattaca ggtaagagcc 13560
accatgtctg gccactgta cttttatata actgaagcac agtaaaccta ctgtgggtttc 13620
gtttacacca gcatcaccac aaacaccatg agtagaacat tgtgctgca cgtaaocgat 13680
ggctacaaca tcactagggtg ataggaattt tttagctcca ttataatctt atgagaccac 13740
tgttgtatgt gcagttcatc atccactgaa atgtccttat gtgatgcatg tottcatatc 13800
caaaaatatt aatcatttct cactgaagcc atgccatgoc atgccatctt ttgcctgtat 13860
tattattttt cagctttttt ttttagattca ggggtgtacat gtgcaggttt gttagaaaga 13920
gtatatcgta tgatgctgaa gtttgggata cagttgaacc agtcaccag gtagtgagca 13980
tagtactcaa tagataacgt tctaacatta ctctccttc cctccctgtt cttgtctctg 14040
tctattgtat ctttatgtcc atgtgtacca aatgttttagc tcattcttgt gagaacatgt 14100
ggcatttgat tttgtttctg tgttaatttg cttacaataa atagtctcca gctgcatcca 14160
cattgctaca aaggacatga ttttgttctt ttttataggc tgcacatata tccatgggtg 14220
ataggtacca ctttttcttg atccagtcta ccgttcatgg gcatttgggt tgattgtatc 14280
tttgctatta tggatggctt ttgcctatat tattggaaag gccttctaac tgggtgtcct 14340
gcttacaccg ttttccccct taaatgtgtt ttcaacatgg tagccagagt aacccttttt 14400
ataacaataa atcgtgtaac ttttttgttc agaaacttac agggcttacc atttcattca 14460

gtaaaagctc aagctcctgt atagtcagac catatccttc atcacctggtt actttttotcc 14520
tctgactctt cagccttttt gtttttcctc aaactgatga agccttcatg gctgatgtca 14580
gatgttttgc ccattgagat ctctcttggt gactcagttg cacttggtca tatgattttc 14640
atatttttgg ggtatctaata cataatctga aagttggcta cttattttta cccctttgag 14700
ggtccttgcc ctgtttttgt atccctgata gcgggacagc cagatatctg gaacttacag 14760
gtgttcaata aagttttgtt gaatgaatat tctggaatca cccaaccttt tttttccct 14820
ccacttattt ttctttctcc tttcacggcc tgaaagatgt cctatgtata tggttccact 14880
tatactctc atcccagttt gtgatatact attccattat attactatta ttaatacaat 14940
tccattgaac ttgctcttgc tgacttcacc actggaccta catgttggcc aaatggatac 15000
tttataattt tagtcttgac ccctgccttt ggcacatttc ttacctctag cacagcactg 15060
tccagtaatc cacactttct gagacagtgg aaatgttcag tatctgtgct gttcagttgg 15120
tagcaaccag ctacccatgc ctattaaaca ttgaaatgt ggctgtgtga ctagtggcaa 15180
ttatgttga gagtacagtt ttagaaactc ctgtttttct tacatggcac tacatttagt 15240
atcacaatct aattgtgcaa gccagatagg taggagtcac ctttattcct gttatttaat 15300
ttttctcatc tactatatcc agttcatcac atcaacagcg cctgttggtt ctacctcta 15360
aatatttctt tagtctaact actacttgtc ctagtgcca ccaccatcta tcagctggaa 15420
tattgctata gctgccttac aggtttccct tctttcctgt tctcttctag ttttttgaat 15480
tttagtcagc acgagatttt aaaaactcaa ataagattgt gttattcacc tgcttaaaac 15540
ctttcatgac tttcagtgtc acgtagaaca gaaaacactt ttcttaccaa aggctagaga 15600
gctctacgtg atctggctat ttttaacgtt tcattgcaact cacccttttc ctctataatc 15660
aaactactct gatctcaagg gttagtctt gaaagatgat catgttcttt aatgacttta 15720
ggtttttgtg tgttattttc tatttctggg atgtttattc tctgttcctt acatgctggc 15780
ccttttgcat ccttcttcag gtctcagctt acatgttacc ttcaagaagc ctttgaccac 15840
tctaagtggg cccttccttc caattctgct gtgtaatccc actcccttct cccacttggt 15900
aattagttac atactttttt gtaattgttt atttggttgc tgtctccctc tcaagaatgc 15960
agggaccatg tctgcattct gcagtaatca ctactgcaca cccagaatct attacagatc 16020
ctggcatgta gctgatgcat aaatatttgt tgaatgaaag tctgtacatt gtatttatgc 16080
tattggtatt gctatgacct gaaactaaaa ggagttgtgg aaaagatttc ttatggaaca 16140
gaaatatccc ttttgattaa tatcacaatc tcgtaaattg agaaaacaaa aaaatatata 16200
ctactggagc attcatgtat agttggagat tatgactcat ttattggtgt gtttttggac 16260
tcagaacaaa gatgagggaa tattccttaa agctctgtat tgaaataacg aaaagcagtc 16320

acattttaat aatagaagct tcctagctta ctctttctgt aatcttcttt tcctaaatgt 16380
aagagagcct cataattatg aggcattatta ctagagtaag gctgtcaaag gcagcaaaat 16440
gtctttctgt ttggaagaat aacataaact tgacatgtat ggtgggggac agaaggtttc 16500
aaaagtttaa gaatctgtgt tgtottaaca aatagatgct tctcaaggag ottacgctag 16560
tggttactct gtccagtcag ggttttttct tctttaactt gggttcattt cctgatggca 16620
cacatgaagt ttggatcata tggtttgact ttagctatgg tccttagcta tggggagcag 16680
catcagcgac ctgtgacatg taaattaaaa atacaatgcc agggcccttc ccagcccct 16740
ctgatagaga aaccttggc catctgtatt ttagatggt ccaggttagt ctgattaaca 16800
cccttggtta agaaccattg ggaggatctg attgccagtt taaggggacc ttcaagcctg 16860
taggtcttta tagttaaaaa aaaaaaaga ttttaaaaat catgcatatg ttgtggctga 16920
attctggttt agcacatact gcttttaatg gcctgaaatg tttttccaa ataaattgtc 16980
ttgttatagc tttcatgtgt gatttgggtcc agcttcttgt tttgaagata ottacggggg 17040
ggaacacttt gtgatttctc ttagtaacat attaaaccac ttaaaaaccc tttctattac 17100
aggtcttcac atttaggctt aatgtgctta attcaaagt aaaaatacac ctgcctttgt 17160
tctcagtga agtatgtaat aaataaatga ggggttgga aactactgcc caccatctgt 17220
ttttttatgg cctatgaact aagaatcgtt ttggatagct aaaaaaaaa atcaaaagga 17280
taattatttt gtgacgtga aattatatga aattcaaatt tcagtttctg tgaatgaagt 17340
tttaatggaa cacagccatc catgcttatg taagtgtgca tattctctgg ctgttttcac 17400
tgcaatagca gagttgagta gttgtgacaa agagtttatg gccacaaaa cctaaaatat 17460
ttactttctg atgctttaca gaaaaagttt cctgaacctt attctagcta tatgttggtc 17520
ataaatgaat ctttctgtgt totgaaggca ttaagaatc tcttaggtta taaattggct 17580
gggcgcagtg gctcacgcct gtaatccag cactttggga ggccgaggct ggtggatcac 17640
gagggcagga gttcaagatc agcctagcca agatggtgaa accctgtctc cattaaaaaa 17700
aaaaaaaaa aaaaaaaaaat agctgggggt ggtgggtggc agtaatcca gctactcggg 17760
aggctgaggc agagaattgc ttaaaoccag gaggcggagg atgcagtgag ccaagatcgc 17820
gccactgcac tccagcctgg gcaacagagt gaaacaccat ctcaaaaaa aaaaaaaaaa 17880
aaaaaactc ttaggttata aataattgtt gttagctctc caagcctcca tattacattt 17940
tgtgtgttct cctgttcaca ttttgagcat tttatttttt attagcacat tcagttcatc 18000
aggatattaa gagcttaata tatgcaaag catatatata gcgagaagct gtttctaaat 18060
gtactgtctc agccctcaca gagttcactt cattaggctc tttaaaattt ctttctttaa 18120

aaggctcagcg tgctgggtata gtggggaagg gaaactctta caacacgtcg agtagaggaa 18180
ggttatcatt atgggatata atttggaagt cattgagtag ctgccattaa ttctgcctgt 18240
agtctgaatg tagagattaa catgtagaaa cttttttgaa ataaaatctt caattttcttt 18300
ggcatatcta gtactgtcta gctaggcata tagtcaaagt atgggtgtata tttcaagtat 18360
taaaagtttt tttgggctgt agtcactgtt gaaaggatat agttctttac tattacatgt 18420
gataccttta tataaaattg gctaaoccct gtcttttcatt tatctgcaac actgactgtt 18480
accagttgtc tctaactttg gtatgggggg tggaaatatg attagattga aagggtacat 18540
gactgagcca caagcagacc tggatttgaa ttttaactga acggtttatt agctattctt 18600
acattaatac tgctaatacag ttttcttctg atatgaggaa tgatgtcttc tttatgaggt 18660
tgctaggaag attcaatgag ataacatact aggtcagaa ctgaagttgc taggaattta 18720
attatgttac cttgttaaag tatgtcaaag gcagaattca gtgttttagct gataccacaa 18780
ggcagtatcc taaaattatg ctgtaaaaga tataaagatg ctgtaagtga ctcaaaaacc 18840
tagtgacttt gtaatgcagt tgattcttag aatactgtca ctttaacaga ataggagcta 18900
ggaatgaaga aatagttatt aaattactaa aatagaaaat ttattgacac atgtaaagtg 18960
acatttgctt aaatattgaa aaatttgtag tactatttcc ttgctttaga aaacattggt 19020
taccactttt tttattttata gcagtttgtt tttgcottga ggcaagatgg ttgactgagt 19080
agttgccaca tttcttttgt acaaagtcca tttcataggc catctagctt ttatgcttag 19140
aaacatttcc ttaacgttat atttcagtat ttggctaacc tatatagggg taaattatat 19200
aggctaactt ctcgacaga tatttctaata aatttatgta tttggttctg caaatgtatg 19260
caaaaatata tgtacaaagg tatgcagatg ccttgcatac ttgatatatg ttaattttt 19320
tttaatgtag acctttttcg ttctctttta tgactatatg gtattccacc atccccgct 19380
cacctggaca actacagtaa cctcctaaat ggtgtttcta ctttgcattt gcccttatt 19440
gtcttttttc ccctttatag ctgctggagt gaattttaga aagcctaagt catacatcac 19500
attgcttcat gggcatccca gtacactttg gattttattt tacatcctta ctgatctgat 19560
tctcatctct gtctcttcat gggtctctgc cttctagtta cactggtgac ctttcaaac 19620
ctttaccaca ttgagttcat tccttacttt tcaactcttc totgcctgga gtgttctgcc 19680
ccatctttac gtggccagct gctcctctc tgatgaaatg tctcttctc acaggccttc 19740
cctgaccacc cactagagta gcacatcttc tacctcataa acttgtttat tagtatttt 19800
tactctaaat tttcttttaa attgcttaat tccctaacag tagaatataa gcttcaactgt 19860
atgtatgac ttgttgactc tcttactcat tgttattgta ataccagtaa caaagggtgt 19920
ttaaattttg ttcagtgggt gaatatatgt tccatttaat ggataaatta ttttttattc 19980

agtcctcctgt tgatggacat ttgaataatt tccatctttt tctctatgaa tgcctcactt 20040
ggcatgcttc tgacagtatt gccacagaat acattttctgt tataaaaaatt gaatttttaa 20100
gtcaaagggg agttacactt taatggatag tggcagctta ctatcaaaag tttctgctag 20160
tttcaccata tocttattag cagtagatat tatcaatctt ttcaatcttt gccaatctga 20220
taagcaaaaa gtaaattggg ttaaacaatcc tttgtatata ttcatgtctc actttatggt 20280
tttcctttga aatgttattt cttgttcttt ccctgcagta tgattctttc ttttttgac 20340
ttgttcccag ttttttgtgt actatggata ttagccttta attatgttac ggatgttcta 20400
gtatgttatt ttttgaatta cttcaaagt gatttgttgc tcagatttta aaaactacat 20460
acacaaatta tctcatgttt cccttttttg tttcaatttc gactcatgct taatcagttc 20520
atcgattggg catggtttta ttcttaatat ataccogtat tttatctcat tttatttttt 20580
tacgtgtaaa tatttggtga atatagggtt aattttaatg taaaataagg atgaaaaatg 20640
atagttggaa ttacaagccc atttctccta ataactttta tcaagtaatc cactaattga 20700
aatattacct tottcattta tgaaattgcc acattatato tgggtgtttt tctgcctact 20760
acagtctctt acccatttct ttctaataa tacaataact gaattgctgt ggttggtgat 20820
ttataatgtt atcttaatga taacattata aatgtgatgg aactggttcc tocttatagt 20880
tcttcttaaa tcaagaacaa gacatatctt cccatttact ctctgatgta tctcatttta 20940
ctgttatgaa tgaaatctgt cctatttgtg tataggaaaa tagtttttgt atgtaattgt 21000
gatatggcca gttttattaa aaatttggtt aaactaagag ttgtttctg ttcagcctta 21060
tcatactata aaatccacat aaaatgggta taaaagtgtc gcaggacact gggctcagat 21120
gattctccca cctcagcttc ccaagtagct gggactacag cggcatatgc caccacaccc 21180
agccaatttt taaataagtt ttaaaaatag tatttttagt agagacaggg tttcaccatg 21240
ttgccaggc tggctctgaa ctctggact cagacaatcc acctgccttg gcttcccaaa 21300
gtgttgggat tacagggtgt agccaccaca ccttgccgaa ttgcagccat atttaatact 21360
ttttccatc ctattccctt tgctgcccc aggctctctg tattgatagc ccgctattaa 21420
gaagctagtgt tatattcttt gcatactttt acttcataaa ctatatgaag cattgttctg 21480
ttttttaact taattggtat aaaattatat tttggaaatt cagtatatct tgtgaaaatt 21540
atttagaaaa tgtgcctctg agataaagcc tattcaggat gtatcttaaa ggagatagct 21600
gtgctttaac attatcagtc tttttggctg cttatgttaa tataagttgg agaaaaacag 21660
tctgcttttt gtgataatat gttcttgag atggagtga agattgttta aaaacattgt 21720
cttttttttc ccctgaagta ccagtattta ttttaggatt atgttactga tcaaagatgc 21780

tgtgtggagt tactcattgg tgagactaac aataaatcac acatgcaaag gatgttacca 21840
taatctaatt attttaaaca gtaaaattat attctaagac atccagttgg cctatatgtg 21900
ctatatcaat gactatcaag gggcttttta tgtatactgt atacatgtac ttcacaaaaa 21960
tataaaagga tgacatcaaa aatctggcaa gccaaaagcc tacattacat gtagcaaata 22020
aataagcata tgaacttatt ggaattttaa accctgtagg atgggcgggt gatggtatgt 22080
atgttagatg tgtggacata tctattaaaa gttgtgtcag ataacagctg gtgctgacaa 22140
gcccttggtg agatggcagc atgttcaata tgttctgtga aaattatctc agtttatgat 22200
ctgtcagtat tgtggagcta tgcatgaaag gacttaaaat tottaccctt aaactcagta 22260
acagtgtttc tagaacttct ggtgatatgg gaaattaaga gaattattta tatgcaaagg 22320
tgtttattgc agcattgttg gaataataga caaaatgggg aagaacaagc tcagaatgga 22380
ggaggtagct tatagtatag acatacgata caatccagat gataatattt tataatagtc 22440
ttcacaagga attttatatt tttattttta aaaatacata gcagtgagtt taatatacca 22500
aacataccaa aatgtcatca tttactgtgt ggtggactca tatgatggag atgataaata 22560
aaaatattaa tttatttgag gcatatatatt atggctgagg aaggaagaca gttatgaaga 22620
acagctcatt ctggaaacat actaattttt cccagccata aagagatttc ctatttcttt 22680
tttttttcca tttaccttct gtttcctaoc tgagaagatt tcatacttct aataaccatt 22740
tgtgtacctt tttaaagaca gtaccaaagg catacatttt agtgtttgga ggaccaaggg 22800
tcatttgatg tttgatgctt attgactatt cgaggatgac aagacacctt gagaacacac 22860
acaccacac ccacaccac accctcacc acccaccoca cccccctccc cgaagaaagc 22920
tgtgaaggaa gaaagcagaa aagaacctgg agtgagttgt aacttaaaat gttagtgttg 22980
catgaagtgt gttaaaacag gaagatttga ggaaattgca tacattttct agatggcaaa 23040
gtattactgg tgacagttaa tgaaaatgca tatgcatgtg ttttttagatt taaaaattt 23100
actaagaact ttttaaaaat ccctgaagggt gtatcaaaag tttatcatgc ttatgaaata 23160
gagtagcact ttctaacttt aaaacgggga ataattcttt ggatcttgat tattggaaaa 23220
gtgaattatg aattgctagt ataaaactgt ggttttaaaa tatgtctgct ttatatattt 23280
atgtagcaga tttactcta gttaataata ctcaaactta ctgaaaacta aggtaattaa 23340
gataattctg toctgatggg aagaggaaaa ataacttcag tgtgaaatct attatatatt 23400
agttgtggca agatttctcc cattgaacttt gactggagac atttataggg taaaaatcgg 23460
aaatagcacg gtgaattttg aagtatcctt gtagttggaa agagtattat gttcatattg 23520
ccaaaaaaaa gatgcatgga tgcattagac tggatggaaa atacatgaga agttggctag 23580
ccccctcttt gtcaaaacat cacttggtgg tgataaagct gttggaaaac acagcattct 23640

aatgtagtct gtagtttaat gataatctgt gtcttgaaac atttagcgta gtacttatac 23700
aaacctagat ggcatagtgt actgcatgcc tagcctatat agtatagcct gttgcttcta 23760
gggtgtaaag ctgtatagcg tgttactata ggcagttgaa acagtggat ttatgtatcc 23820
tttttttttt ttttaaatc ttttaagaga cagggctcttg ctctgttgcc caggctggat 23880
gcattgggtgt gatcatagct cactataacc ttgaactcct aagtgatcct ctttgcccca 23940
gcctcccag tggctaggac tacaggcaca tactaccaca cctggctaatt ttttaacatt 24000
tttttgtaga gatggaattt cgctgtgttg tccaggctgg tcttggaact cttgtgctgc 24060
agcaatccac ccgcctcca aagtgttaga attacaagcc acttcgcctg gcttgcttac 24120
ctaaacatag aaaagatcca gtaaaaatac agaattaaaa tcttggtggg ccactgtagc 24180
atatgtagtc catcttgact gaaatgtcct tatgcagtgc atgattgtac ttcataattt 24240
ttaagcactc ctccctcttg attggtactt agtggatttt atcatttttg tttcttcata 24300
attctttctg aaatgtctac tggttggacc tttgatctcc tgaattgatc gtgatttctt 24360
ctggtgtatt ttttgtcttt gtcatttttt tgtactctag gcagttttct caatttttagt 24420
ttctattcaa ctttttgctt ttattttatc tctccagtat ttatggagat actaaattga 24480
agtgttctgt ttctctctcc accotatccc tagtttcaag ttttatctca gtttctatgg 24540
agtcagtttt ttogttgctt taaaaaaaa ttttctgaa gtgattggta agttttggct 24600
aattgggagc actagaattg ggcccttaat ggttggcagg gtgtgggtga ggagagacag 24660
cccttagtcc aaaggctcag gccagaaaaa gaaagaggaa ggctttcctt ttcccttccg 24720
gagcaggggt ctgccctagg tcttgcttgg cagtctatct gatttcttta gcagttaatg 24780
ctcagttttt tggcatatgt ggatctgcct ccagagcagg tacaaggtga gtgagtctat 24840
gctgttacct aattagatcc ccatttctac cctttgtttt tacttctcta totactgata 24900
ggtttttacc ctccctcacc tcataggggt gcagtgaaga gcaagatgaa tttttattta 24960
tgttgcataa atttttaaag ctaaaaaata tatatgtaat gttgggaagt ccagtgtagc 25020
aaatggctat tgtaaatttg gaacatgaac ttgctttttt ccattgtaaa aatgaaatca 25080
ttataaattg cgggtcaagt actaggtcag cccacacaga gtttaccag taatatgcgt 25140
aaatgttttg cctttgcatc aacaacaagg aaaaacagta ctataaaaa atgttcctgg 25200
aagccggatg tatcaaagca cttctgaaat agctatatag cctatagaca tgaccagttg 25260
gtttctgagt ctgttgacat tggccaaagg agaagctcag tgtagaacat gtttgagtc 25320
tccttttgca gaaatacatt ggaggctgga gtggggaacc aatttttcag aaagggtgtg 25380
aagtagttac atagccactc ttttaagac agtcaaaaga tagaaactaa ggccaggtgt 25440

tggtcacat ctgagatagg aaaatcactt gaacctggga ggaggagggt gcagtgcgc 25500
cagtatgcac ctctgcactc cagcctgggt tggcaagaga ccaaaactct gtctcaaaaa 25560
aaaacaaaac atagtccaca cttaaataatt ttattccata tctttacata cccaatatgt 25620
taatttatag ttcaagatga acttggttg gacagatttt gtaataaagg aaatcgtgtt 25680
attagaaata tctagaggcc atgagccctt aaactgttct aatttgcaag tagttccctg 25740
tgtgatgcag tttttttcaa tattgcacaa taaaggcaaa atacggacaa attagatgat 25800
aagatttata taaattttta aaatattgat caaaatatgt atccatattg gtaatatattg 25860
tattttataat aaatcattgc tgtaaatattg aacttagaaa aattttacta ataaagggtgc 25920
ttttgtgttg caaactttca ttgaaaagt aatttttctt tgtaccaaaa aatctaaaat 25980
tcgctattct agtcacaaa atttgcttta tgaaaaataa tttttgatgg cactatatca 26040
gaaaacaact tgttaaagaa aatgtggagt ttttaaatac ccactgtacc tctgttatcc 26100
aaaggggatc tgtgaatttt tctgtgaaag gttaaaaaag gagagacctt taggaattca 26160
gagagcagct gatttttgaa tagtgttttc ccctccctgg cttttattat tacaactctg 26220
tgctttttca tcaccatcct gaatatctat aattaatatt tatactatta ataaaaagac 26280
atttttggta aggaggagt ttactgaag ttcagcagt atggagctgt ggttgagggtg 26340
tctggaggag accatgaggt ctgcgtttca ctaacctggg aaaagaggat atgggttttt 26400
tttggtgtg taatagtgc atttaacagg tatcccagt acttaggagt attaatcaag 26460
ctaaatttaa atcctaata cttttgatta acttttttta ggtatttga agtataccat 26520
acaactgttt tgaaaatcca gcgtggacaa tggctactca aggtttgtgt cattaaatct 26580
ttagttactg aattggggct ctgcttcgtt gccattaagc cagtctggct gagatcccc 26640
tgctttctc totccctgct tacttgctag gctacctttt gctccatttt ctgctcactc 26700
ctcctaatagg cttggtgaaa tagcaaaca gccaccagca ggaatctagt ctggatgact 26760
gcttctggag cctggatgca gtaccattct tccactgatt cagtgcgtta ctgttaggtg 26820
gttcctaag ggattaggt tttcatcact gagctaacc tggctatcat tctgcttttc 26880
ttggctgtct ttcagatttg actttatttc taaaaatatt tcaatgggtc atatcacaga 26940
ttcttttttt ttaaattaaa gtaacatttc caatctacta atgctaatac tgtttcgtat 27000
ttatagctga tttgatggag ttggacatgg catggaacc agacagaaaa gcggctgtta 27060
gtcactggca gcaacagtct tacctggact ctggaatcca ttctggtgcc actaccacag 27120
ctccttctct gagtggtaaa ggcaatcctg aggaagagga tgtggatacc tccaagtcc 27180
tgtatgagt ggaacaggga ttttctcagt cttcactca agaacaagta gctggtaaga 27240
gtattatttt tcattgcctt actgaaagtc agaatgcagt tttgagaact aaaaaggtag 27300

tgtataatag tttaaataaa atgttgtggt gaagaaaaga gagtaatagc aatgtcactt 27360
ttaccattta ggatagcaaa tacttaggta aatgctgaac tgtggatagt gagtggtgaa 27420
ttaacctttt ccagatattg atggacagta tgcaatgact cgagctcaga gggtaacgagc 27480
tgctatgttc cctgagacat tagatgaggg catgcagatc ccatctacac agtttcatgc 27540
tgctcatccc actaatgtcc agcgtttggc tgaaccatca cagatgctga aacatgcagt 27600
tgtaaaccttg attaactatc aagatgatgc agaacttgcc acacgtgcaa tcoctgaact 27660
gacaaaactg ctaaatagac aggaaccaggt aagcaatgac atagctagct ttttagtctg 27720
ctttgaagta aatgctcaag gggagtagtt tcagaatgtc taccataac cagtacttga 27780
aaactaacga tgtttctgaa ttctgtatt acaggtggtg gttaataagg ctgcagttat 27840
ggtccatcag ctttctaaaa aggaagcttc cagacacgct atcatgcgtt ctccctcagat 27900
ggtgtctgct attgtacgta ccatgcagaa tacaatgat gtagaacag ctggtgttac 27960
cgctgggacc ttgcataacc tttcccatca tcgtgagggc ttactggcca tctttaagtc 28020
tggaggcatt cctgccctgg tgaaaatgct tgggtaagaa aacatgtcag aatgcttgaa 28080
gctaaaaagt agaagagtat actcacaata tttctgatga ggcttttttc ttcttccag 28140
ttcaccagtg gattctgtgt tgttttatgc cattacaact ctccacaacc ttttattaca 28200
tcaagaagga gctaaaatgg cagtgcgttt agctggtggg ctgcagaaaa tggttgcctt 28260
gctcaacaaa acaaatgtta aattcttggc tattacgaca gactgccttc aaattttagc 28320
ttatggcaac caagaaagca aggtaagaga attattcttt atgtggtttt catggagcat 28380
tggaacacct cagtgtcatg tcattccatg cagtgttctt aacctttttg gcaccaggga 28440
ccagtttctg gaaaacagt tttccatga atgggtgtg ggaatgggtt ctggatgaca 28500
ccattccacc tcagataatc aggcattaga ttctcatagg gagcgtgcag cctagatccc 28560
tcgcatgtgc agtccacact agggtttcta ctctcatgag actctcatgg tgcagttgat 28620
ctgacaggag gtagagctca agccaggtaa tgctcgctca cctgccactt acctoctgct 28680
gtgcagccca gttoatttct gttcttttaa atttttgagt ttccatatgt aaagcactat 28740
gcgaagtagt agggatatgg taggcaagct tctcttcaca cttttgttct taggtgggat 28800
gtagatgttg ggaataataa cctaataatt aatttgtgta gtgggaagaa gtggggctat 28860
gagggcacat aacacaagtt gaaactgact ctttttgagg gttcaaggag acctcttgga 28920
ggaagtgata gttgagttca gtgttcaagg atgagaaggg attcactagg tgaaggtag 28980
gtgagaaaac aacatctttg aaacgaagga aggagatgga aagttttggg aatttaagaa 29040
atactaatag taaggaggaa gaaaggtttg aggtgaggct attgagatag acttagcaga 29100

tctcataggg ctttgtagag catgttttaa agcacaatgg gaaatttcag cagaagcctg 29160
aatgatgaa atttgttttt agaaaattgg ggcagtgttg aaaggaaga tatacaggga 29220
atgaaaggac aagcatgaat gatcatttta tggatctgtg ttttaagggtg gatataatta 29280
ggaaaattaa agggccaaat gatgaggagt taagtgccag ttctggttca aattttcagt 29340
gaatcagttt tgatataact ttcattcttag ggcattactc ttgcctacca acatagtttc 29400
taaatttttt tcttttgggtg tgatcactgt gggaagaagg aaattgggcc caaactgata 29460
cattgtttgg aggactggga tgtctgaatt tgagtggaaat gctttaaaag gacaagttgg 29520
atagggcccc agtatggggg tctgagtgat ggggtccagg aatacattta ggtccaatgg 29580
caagctggct gaaattcttg tataataaaa taggttggtg atatggctct tctcagacat 29640
gtgatcaaga ttcottgact aacaagatat atatataat ctttctagct catcactg 29700
gctagtgggtg gacccaagc tttagtaaata ataagtagga cctatactta cgaaaaacta 29760
ctgtggacca caagcagagt gctgaagggtg ctatctgtct gctctagtaa taagccggct 29820
attgtagaag ctggtaagta tatgtatcta ttctgagtct tgtgtatagc atctgcagtt 29880
ctaattagat tacttttctt aggaaaagggt ggtagaactt taactactga aaataaatgg 29940
tcctattcag tttgcagcca agatttacat tcagagtacc tgtcatctgg attgtagcta 30000
aatatttaag gctagtttag gtagagtctt tattatccat caaaaatgat ggcatatgtt 30060
ttgcttaata aaatttggtt gtaatttcag ttttgagtaa acctaagatt tgctaacaga 30120
gctgtgaatt tataggagaa aagacaaatt ctaatatagt acagttttat gtaaagtgat 30180
tgctttatta gtagatgctc atgagcagtt tttgttttgt ttttaactttt aggttccggg 30240
taatgtgcag gcttggtata taggtaaatt gcatgtcaca ggggttttgt gtgcagatta 30300
ttttgtcacc caggcagtaa gtattgtacc caataggtag tttttcagtt ctttacctcc 30360
caccgtaag taggccccag tgtctgttgt tcccttcttt gtgcccgtgt gtactcagtg 30420
tttacctccc acttataagt gagaacatgt ggtatttggt tttctattcc tatgttagtt 30480
tgcttaggat aatggcctcc agctccatcc atgttgctga ggaagacatc ttggatattt 30540
tttatggctg cttagtattc catagtatat atgtaccaca ttttctttat ctagtctacc 30600
attgatgggc atttaggtta attccatata tttgctattg tgaataatgc tgcagtgaac 30660
atatgcatgc atgtgtcttt atggtaaaaa gatttctttt tctttgggca tatacctaata 30720
aataggattg ctggattgaa tggttaattct gtcagggtttt ttgagaaatc accaaaattgc 30780
tttccacaat ggctgaacta atttactttc ccaccagcag tgtataagca ttctcttttc 30840
tcagcaacct caccagcatc tgtcattttt tgacttttta ttagtagcca ttctaactgg 30900
tgtgagacgg tatctcattg tggttttgat ttgcatttct ctaatgatca gtgatgtcga 30960

gctttttcttc atatgtttct tggccacttg tatgtcttct tttgaaaagt gtotgttcat 31020
gtccttttgcc cactttttaa tgggggttgt cttttttgct tgttaattta agttttattgt 31080
aaactctgga tattagacct ttgtcagatg catagtttgc cagtactttc tcccatgcca 31140
gtactttctc ccattctgta ggttgtctgt ttactctgtt gatttctttt gctgcgcaga 31200
agctctttat actgtcccat ttgtcagttt ttgtttttgt tgcaacttct cttggcatct 31260
tcgtcatgaa atctttgcca ggtottatgt ccagaatggg atttcctagg ttatcttgca 31320
gagtttttac agttttaagt tttatattta agtctttaat ccattctgag ttgatttttg 31380
tacatcatgt aaggatgggg tgcagtttca atcttggatg tggctagcca gttatccag 31440
caccatttat tgaataggga gtcccttccc cattgottgt ttttgtttac ttgttaggtg 31500
tgcggcctaa cttctgggct ttcttttctg ttccattggg ctctgtgtct gtttgtatac 31560
cagtaccatg ctgtgattgt aacctgtat taacagtata gcttgaagtt gggtaaagtg 31620
attcctocag ttttgttctt ttgtcttagg attgccttgg ctattcaggg tcttttttgg 31680
gttcatatga atttttaaat agtttttttt taattatgtg aagaatgcca ttggtagttt 31740
ggtaggaata gcattgaatc tgtgaattgc tttgggctgt atggccattt taacaatatt 31800
gattcttctt gccatgaaat agaatgtttt ttcatctgtt ggtgtcatct ctgatttctt 31860
tgagcagtgt tttttgtaat tctcattgta gagatctttc acctccctgg ttagttgtat 31920
tcctaggat tttattcttt ttgtggcttt ggtaaaggg attgcattct tgatttggct 31980
tgcagcttgg atgttgttgg tgtctagaaa tgcttctgac ttttgtacat tgatttttat 32040
atcctgaaac tttgtgaag tttattggat caaggagctt ttgggcagag attatggggg 32100
tttctaggta tagaatcata ttgtttgcaa acagacttcc tatttggatg cattttcttt 32160
ctcttgctg attatgagca gtgttttgcc ctgataattct gtattctcag tgaatagatg 32220
tcgtctaagt atgagaaaca attttttctt attctgagta tttttaagaa ggcaacttat 32280
atgtgggtact ttgtatatgt tgtatgttgg caattgggga aaagaataga tggtttgtac 32340
tagggcctct tgggttctgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtca tgaaaacagt 32400
tacttttttag ctaccaagca ttttttctcc tttcagtaac ccacctaaaca acatttactc 32460
agaatttcaa agcaagcttc aaatcagtat tgaaagaagg aaaaatataa aggcatttaa 32520
tggaagaaaa tgttgggaat aaagtatagg gctggcaaca cttacttttc toacttattg 32580
agagtaattt tacttgggaa tttatgagag agaagacat tatgattgct ccaggtaact 32640
actggcagag gaaccatagt cttggggata gacaaatgtg gctgagttca tatagaatga 32700
ggggatggga tgtaaattct gtcagctgtt ccagcagtaa cctgtaattgt aggctaaaaa 32760

tacagatttt gagattttatt taatcagaat ccctggagtg ttaattttta tatcaagatc 32820
tcatagtgtt ttatttgaag tgacagggag gtctgtagat agctggacat gtatgggact 32880
ggaagcttag gaatctttaa gttcttccag gttattctta tgttcatttg tttattctga 32940
aaatagcatc taatgtattt taagaaatgg aataggcaca tagtatacat tgggtaacac 33000
aacagatagg gtccccgtgc ttaattctta gtctgtgaa ggtgacaaaa ataactaaaa 33060
atatgtgac ctaaattaga atgagtgtta tgggagaaat gacagcaa atgtgatgaga 33120
attaatgggg aggggaattg tctagatgag agggaaaagg tctccttgaa aaggggatgt 33180
taagtgggac tgcaggatga gagggaaccg tctcttgtct atatgagaag tgagggttaa 33240
acgttttcca ggtagagaaa aggaacacca tgtgctatgt cttagaacca gggatatcca 33300
gtcttttggc ttccctgggc cacattggaa gaagaataat tgtcttgggc tacacaccaa 33360
atacactaat gatagctgat gagctaaaac aaaaaaaat tgcaaaagaa tctcataatg 33420
tttaagaaaag ttacgaatt tgtgttgggc tacattcaga gctgtcctag gccatgtggc 33480
ccatgggctg caggttggac aagcttgct tagaaggaaa gagattggtc aggcacgggtg 33540
gctcacgcct gtaattccag cactttggga ggctgagggtg ggcggatcat gaggtcagga 33600
gatcgagacc agcctggcta acacagtga accccatctc tactaaaaat acaaaaagtt 33660
agccgggctg ggtggcaggc gcctgtagtc ccagctactt gggaggctga ggcaggagaa 33720
tggtgtgaac ccgggaggcg gagcttgtag tgagctgaga tagcgccact gcacttcagc 33780
ctgggcgaca gagtgagact ctatctcaaa aaaaaaaaaa agggaaagag attgtggaga 33840
tccaggtgct gaagagaagg tctgcataaa cagaacttag taatgaggtg gatggcctgg 33900
tatgaggttg aggttaggta agcagagcca taacatgcag gactttctag gttcctataa 33960
gatagttact actcatggag tttattcatg ctttattcca gctttggagc catagatata 34020
gaatactttg gtcagtttg aaggctaggt gggatccaaa ttctaaacgg ttcctcaggg 34080
ttatactaaa gtatttctat tatcttaaaa ggatgctgag acactttcga tggttgttta 34140
tcaatagcaa agcatcacag tgggtgtgtt aaaatattaa taatagcatt gtatagatta 34200
acagtttgaa tgacaaaag ctagaagacc agactactga gatgttacag gcttttagga 34260
atgaaatagt ttgcttttag aactcaatag caaagggcag atgtctgaga tgcctgaaag 34320
aatcatagaa tgtaataata taggagctaa gggagcaacc aaaaacggtt tgtggagggg 34380
acaacattgg taccatgaag ataaatggaa ccctcagaag gcatccttaa tttttgaaca 34440
taataattta agaagctgac ttaaagtac ttaaagggtc agtaggtagc tggaaatgta 34500
tgatactaga atgcaagaga ggcaggctag agatttggaa gtttccctct tagtatatag 34560
gggtaagggc agcagggag gggaggtaga ggtgccacag agtcactctg atgggacttt 34620

tttttttacc ctagaactgc tgaatcagaa tgtgtgtgtt ttaaagtctc tgtaggcoat 34680
tctgatggac atctgggggtt aaaatccatt ctcttagagt taatagttat gtaaaggag 34740
ggaatgaagt cttaaagagg ggaagaagg tagtcatttc acaatactg agcatcctga 34800
tcatcagtct tacgcagatc attctattag tagctggagc tactatgaaa aaggaaacca 34860
acagagggtga tctttgtctt gtagggaaag tggagtaact tacactatga aggagaagtg 34920
cagggtacca taagaattac agcagataga cctcatctga ggaaataaaa cagacccgaa 34980
agatgaagga gacaaggaaa agtatctctt actgcattca gaagtgattt aagttgaaga 35040
tggatgagcg aagttaatct actatgtggg cattgggctt ccatttatac tcctttgcca 35100
gagtaaagt ccccatctta agggctcctaa aggatggaag attgtaaacc ttggaacaca 35160
tgtttttag tagtgaatt gtataaagtc cctgacagta agtgttttca tgccgtcttt 35220
ctggattgtt cttaccccag gaatttacct agcttcttta ggtctttagt cagatgtcac 35280
cttcacagtg aggtgacctt attatctatt taaaatcgca gcccactcc attatctttc 35340
tccatagccc tttaatatca totgacatac tgtaagggtt tagtttattg tatatctttc 35400
tgcctcttcc aactagatca taaattctga gggtaggaac ttctgaatat tttgttcac 35460
tgggtctatct gcagctcaga acaggacctg gtactgaata aatatttttg aaatgattga 35520
atggatgaaa agaatgagt aataagaata ttacctaagg gggacagtgg agataacaaa 35580
ggctttttcg gcttaggaaa ggaacagtag ctatttgaga gttgtgact agtgagggtga 35640
actggcaaag tgaaggaaac tgagcaacat tctagaaaat gagaggaaat caaatactta 35700
gggtgaaagga agtaaaactt ggaaatacag aaggacacct cctaaggcta gaaaagatat 35760
ttaggattga taggcacttc tagctaatac ctagggcctt atatcctttt taatctttcta 35820
gggtggaatgc aagctttagg acttcacctg acagatccaa gtcaacgtct tgttcagaa 35880
tgtctttgga ctctcaggaa tctttcagat gctgcaacta aacaggtaaa ttctgagtaa 35940
actggtgcca tgggaataga gtcaagatga gtatgtgctt gtactgacca tctgttttta 36000
tctccatagg aagggatgga aggtctcctt gggactcttg ttcagcttct gggttcagat 36060
gatataaatg tggtcacctg tgcagctgga attctttcta acctcacttg caataattat 36120
aagaacaaga tgatggtctg ccaagtgggt ggtatagagg ctcttggtcg tactgtcctt 36180
cgggctggtg acagggaaga catcactgag cctgccatct gtgctcttcg tcatctgacc 36240
agccgacacc aagaagcaga gatggcccag aatgcagttc gccttcacta tggactacca 36300
gttggtggtta agctottaca cccaccatcc cactggcctc tgataaagggt aaattgtcaa 36360
agtagaattt acctttgttg cagaattgaa aatgaagcat ctctagctgt tggatggctg 36420

tctaagcata gtgatcaata agtaggaatt gtattcctta gtaagtagga agtatggctg 36480
cgataggggt aagattctga aatgtttgtg tagtcagaac tacttttagt tgataccaat 36540
agatttagtg tgggtgggaat tttagggtaa gaaaatgatt ttgttgagtt gtatgccagt 36600
tcttccttct gtttttcagg ctactgttgg attgattcga aatcttgccc tttgtccgc 36660
aaatcatgca cctttgcgtg agcaggggtgc cattccacga ctagttcagt tgcttgttcg 36720
tgcacatcag gatacccagc gccgtacgtc catgggtggg acacagcagc aatttggtgt 36780
aggtaaattc ttacagtgat acctggctat ctaaaaggaa tgcataaatc caaaggatcc 36840
tgaacttctt tcttttgtca ttggttcccc ccatccgtct tcctgaagag ctaatgacaa 36900
agtaaataaa taaataatta cacatttcta tggctgcaga gaaaataagg catagtgtgg 36960
ccccagtgat atttccttgg acacgtcctt cacatggta gtcttacaaa ggttgggtta 37020
ggtgtttcat aaagtgttct catttaattt acacaaaggc ccacttcctt aggaagaggt 37080
agagtcataa tttgagatca aatctgtgta atttcagagc ctcttaccct tgcctcatca 37140
tgcattttga ctataaatat ttagcagtc cttttattat cttttctgtg agttaaactt 37200
ttttcatgga cctaagaata ttcagaaata agtagtagca tttctgtact ctttaaccaca 37260
aaaatctcaa cctgaagctt tgatacaaag tttgtgtctt aaaagtagct tcattaaaag 37320
tatagtctaa tgacatttct gatttctcag actttaagac cttattaggt tagtttagaa 37380
aacaagatg gagcctacca gaacagatgt taggaatctc attttgctgg ttgctttgtg 37440
tatgtactca tattggggct ttggctttct tcatttatta ctgttggtat tggcccatct 37500
ccatgaggtg acttaataga acgttgaggg caccttttat tttaaatctc ttttctagga 37560
agaagagagt ttttgtgtcc ttgtaagaat caagttattt ataaaagctg ctaaatgtag 37620
cagaataata acccctttta aaactcaaat ccagaaacag gagaacaga tggacttac 37680
atattgcaaa agctatcttc cttctataca tgaggctgtc agctgaatag tcttggaaga 37740
gtgaggagtg aatttttctg ctggcaactc ggttagtttt agcagttggt gctaaaactt 37800
ggcaaagttt tcaccaaata catggaagat atacaaaaat agagggggca tgtaaaagaa 37860
aaacgttgac atagtctgag cattactttc tcatcttctc tttttatata ctttttacc 37920
agaatgattg gtgcccttac tgtaggaaag ttgtctttgg gattcagcgc tgtatggaag 37980
ctctgttgca ctgtgtatgg gggaggggtg ctgctttgaa ttagtgctgc caggaggcct 38040
cttttcagtg acattcaagt taatggaatc cttcttctt cctgaactaa ttgcaagtta 38100
cggggaactt cggttatata atgtaaataa ttacagtcta ataattgttc ctcaaacttt 38160
acagaggaga atgccctgtt tgttaaccat gtttcttttg gcaggagggg gtccgcatgg 38220
aagaaatagt tgaaggttgt accggagccc ttcacatcct agctcgggat gttcacaacc 38280

gaattgttat cagaggacta aataccattc cattgtttgt gcaggtatgt ttttaagtga 38340
gtgttctagg ttttatgtcc ataaaatttc cagattgtaa tgactaataa catttcagaa 38400
aattagggac cataataggg ttaccaacat ttaattttat gaaaattccc tacatttttt 38460
ggtcagtaag agaaacattg agacttgaga agagggagga gatttcacat ttacttttta 38520
tgggtgccta gaggggagag ctgacctggg ctgccagagg cagggcatag accccaacc 38580
aattctgggt tttccaaatc ttagatcagt tagagctgcc tctgaagaaa gggtttatag 38640
ctaaaaaata ttatggaaat ccagtgtcc agagcattaa acacccaag acataaaatt 38700
cagagaatat tatttactac agtgtgaatg cctcttgac tctgaattgg gaatgtttgc 38760
accacagtgg ggggcttgcc atgttttagc ttagatttta attaggtttt gtttgtgttt 38820
tctccttagc tgctttattc tcccattgaa aacatccaaa gagtagctgc aggggtcctc 38880
tgtgaacttg ctcaggacaa ggaagctgca gaagctattg aagctgaggg agccacagct 38940
cctctgacag agttacttca ctctaggaat gaaggtgtgg gtaagtaaaa aggaaccaa 39000
gccttttagca gatgtgtaca ttgaagtctc agtttttcct caagggcctt tttctccttg 39060
tctcttagcg acatatgcag ctgctgtttt gttccgaatg tctgaggaca agccacaaga 39120
ttacaagaaa cggctttcag ttgagctgac cagctctctc ttcagaacag agccaatggc 39180
ttggaatgag gtagggaaat gtgagcagtt atttatctgg tagtttccta gagcaggtat 39240
ggcagcttgt tctttcctct caaaacactt agtacacatt catttgcat gatgtttccc 39300
tggcttgagt atttcttctt tatgctgtct agcaactgct ctgaggaaga actataatac 39360
aagctttaa gagtctgttc agaataatta caaataagtt gtgttattta aaattataat 39420
tcataagga gaaagatgaa aaatgttacc agattaaaga agatttttca aaaggatgta 39480
aggaaagagg cagtgttaaa cactgttaag aggacagttt atcagtattt tttactaaac 39540
tttaataaaa cttttctatt tgaatttctg ctatgaattt ttcttcagca tttgtcctca 39600
gtactacagg tggttccttg aaacattgtt tctaataaaa ctagaacatc ctgatatttt 39660
atccattcta tagagatcat tgatgttaca cagacataca gtggattatg tttgttgagt 39720
gaatggaaag agagattgtt aggtttacaa cgatgcagct cttgagaccg gagtttaaga 39780
tcagcctggg caacatagtg aaacccatc tttagctggg catggagatg gatgcctata 39840
gtcctagcta ctggggagac gggggcagga ggattgcttg aaccaggag ttaacagact 39900
gcactcagt acagagccag actccaacac aaaaaaaaaa aaaaaaaaaa agcaaattac 39960
cagtgagtag tgtgttactt gggtttttaa taggcatctt attaacatgt tccaacttga 40020
gcccttaact ttctccacct accccttcc acaaacctgt tttactgtc ttctctgtct 40080

tagttaatgt cagctttgtc tgtccagctg ctcaggctaa aacttttctt tcatataaca 40140
catcctatca gcagctcctg tttgtgggta ggcattttgc cttttttttt tttttttttt 40200
ttaaactgct atatctctag catgtagaac agtgcctggc agcacataat aggtgcttaa 40260
tataatattt gttgaaagaa caagtcagtg agtattttta atgtgagggtg caaagagaaa 40320
aaaaaatgta tctttgaggt gtggagtttt gaagaacttc catttttctaa gcattttgtgt 40380
aatgttggag ttacttggtc cttttgtaat ctgaaagtat gctttaaaaa aaattagtgt 40440
acttttgaga attttcattt tgctttctat tcttccttgc tttgtgcatg tttatctaga 40500
ctgctgatct tggacttgat attggtgccc agggagaacc ccttgगतat cgccaggatg 40560
gtatgtgtct catatttctc gattaactcc agatcaagct aaagttctaa aacttttatc 40620
agaagagccg gtttgctcat ctgggaaacc agtgttggca gaaaagtagt ggcttcaatt 40680
aaaagcagtt cttaaattcc agtcagcaac agtatcttta atggagcaca gggaattcag 40740
agccacacaa tgagtagcag taggattaca ccaccaacaa atacatgcta ctgctaggcc 40800
tctgcagtg caggatgttac aatttaacctg gctttttatt ctcttttttg ccagaggact 40860
cataatacct ttgtctacaa gctacccaag gaagatagga aaactcctgt ttctaggctc 40920
agatctcggg tgggttttta catagttgca ttatcatcag ggttttcttg aaaagctaata 40980
ttaaactctg gtaatgaaca tggaggatgg catagaccac taacaattat aactgtctta 41040
catttataac cgcactctgct tctacctaat tatgaaacca ctaaagcgca gattcttact 41100
gtgagaaata acatgtcaac cctaagataa aatatgttga ggtttcatgg aaatagtgcc 41160
tttccttagt acttttgtgg gtgtcacttg gcctttttgt caagatagat tacacctgcc 41220
agacctcatt attgtcttaa tcctccttcc catgacttct cactgcctag gtggtcacac 41280
agtagattcc tgcttcttct cctcgggaac cccaagtctc ttgacagggg taaatgcaga 41340
gtgttcaggg ttagactaat gatgtgacta ggccctgctg gtgtgcctgt ctgatggaaa 41400
tagatgttat ttgtgtagtc tcatgggtgg cctggcactg agtaattact tggctaaaga 41460
aagctggagg ttgaagaggc tagaaagcgt tgttttctga caagtttgct gctgaacttt 41520
ggatgcccta acctcagtgt taacgtctat gtctgcttct ctctctctc ttttgccttc 41580
cttcttgctt attttgttga cacctgact cttctagatc ctagctatcg ttcttttcac 41640
tctggtggat atggccagga tgccttgggt atggaccca tgatggaaca tgagatgggt 41700
ggccaccacc ctggtgctga ctatccagtt gatgggctgc cagatctggg gcatgccag 41760
gacctcatgg atgggctgcc tccaggtgac agcaatcagc tggcctggtt tgatactgac 41820
ctgtaaatac tccttttaggt aagaagtttt aaaaagccag tttgggtaaa atacttttac 41880
tctgcctaca gaacttcaga aagacttgggt tggtaggggt ggagtgggtt aggctatttg 41940

taaatctgcc acaaaaacag gtatatactt tgaaaggaga tgtcttgga cattggaatg 42000
ttotcagatt tctggttggt atgtgatcat gtgtggaagt tattaacttt aatgtttttt 42060
gccacagctt ttgcaactta atactcaa at gactaacatt tgctgtttta aacattaata 42120
gcagcctttc tctctttata cagctgtatt gtctgaaatt gcattgtgat tggcctgtag 42180
agttgctgag agggctcgag ggggtggctg gtatctcaga aagtgcctga cacactaacc 42240
aagctgagtt tcctatggga acaattgaag taaacttttt gttctggtcc tttttggtcg 42300
aggagtaaca atacaaatgg attttgggag tgactcaaga agtgaagaat gcacaagaat 42360
ggatcacaag atggaattta tcaaacccta gccttgcttg ttaaattttt tttttttttt 42420
ttttaagaat atctgtaatg gtactgactt tgcttgcttt gaagtagctc tttttttttt 42480
tttttttttt tttttgcagt aactgttttt taagtctctc gtagtgtaa gttatagtga 42540
atactgctac agcaatttct aatttttaag aattgagtaa tgggtgtagaa cactaattaa 42600
ttcataatca ctctaattaa ttgtaatctg aataaagtgt aacaattgtg tagccttttt 42660
gtataaaata gacaaataga aaatggtcca attagtttcc tttttaatat gcttaaaata 42720
agcaggtgga tctatttcat gtttttgatc aaaaactatt tgggatatgt atgggtaggg 42780
taaatacagta agaggtgtta tttggaacct tgttttggac agtttaccag ttgcctttta 42840
tcccaaagtt gttgtaacct gctgtgatac gatgcttcaa gagaaaatgc gggtataaaa 42900
aatggttcag aattaaactt ttaattcatt cgattgtgtc actctttctt tttttcttgt 42960
cattaataaa tggtttgtga ggggtgtaag taagtttctc caactcaact ttatcctgaa 43020
ctttatcctg agatagatag atagatacat agtgaagtaa tttgtttag cgatagagag 43080
ggccctacta cttctctaga tgcaatttgc caagtttctt tagcatttgg ccttgatta 43140
cgctggaccc ctaaaaaagt gtttgttca gccttgaca ttgtgattag ttctgggctt 43200
tggtgggaga aagccaagca gactatccat ctatctatct ctacccaat cccctcaga 43260
tgatttgatt ccattaaagt aaagtaatga aataaatagt aactaccct acctggattt 43320
ggatttgaat caccttcacc caagtgtgac tcattgggag aggggaattc cttcaactgg 43380
actttgaact aataataatt aaataaatta aataggttta agtgttgatg tttatctgaa 43440
cataatcaga accagaatga gccttcaca gtacctgca gatgatccta tcaatcatcc 43500
cgggtgtagg tctgacaagc ttagttcatc tcatctcttc aggtagacct gtttgtaaag 43560
acttggtttt tggcttttta aaaagtaaaa gaaagaagtg tttgtcttt agggagagct 43620
ctttctgatt tactgttttt attttgatga tgattgatag tgactattaa ttatgagaaa 43680
aacatacatt cagttctttt tcgttaaaaa tcagcttcat ctggggaggg ctttttgtga 43740

atgcccttgc tttgcaaaaa tgaattcatt gtggtctaaa atgtaattag ggtttgaaaa 43800
gaaacattca catagctgat gatgtggtag gtctcagggt ctgtctcatg tgtgataaac 43860
caaaccagcc agtattttcaa gaagcctagg gctatttgga ggtaaacacc gtgaagatct 43920
tcagtgaatt cagcttttta cataaagagc caccgcacgg tgtttcaatt acagaaagaa 43980
caacttttta aaaccagaga atgttttcag agctttgcta attcattgta ttttattttt 44040
gaagtttaaa aacttaactt aatgtgagag ttaatccagt tcaacaaggt ggctctggaa 44100
gtaccaaggg tgaataatag ccaagacgct taggaaaaat agcacttgct ctagcagctg 44160
tcaaggccta actacttaat acaaagctaa agttagttaa ccctgttaca ggaataggga 44220
tataaataca gattaatata gaataatata caggcattta aagactcttc agagacagcg 44280
ttggagatcc gagtgggagg agagcaactt tttcataaag gtgtgaaaca gccgttggaa 44340
taaaacaata cagaaaaaca tccatgaccc aacaaaatat caagatttta aataaaggca 44400
gtagtactga tttttctttt gcctcaagcg ccagtaaggc ttggtagggc acaacgctga 44460
cagtctacac tggacgcca tgtgtccaac aatctctgtt cacac 44505

<210> 5
<211> 780
<212> DNA
<213> Homo sapiens

<400> 5
atggccttggg gaaatagcaa acaagccacc agcaggaatc tagtctggat gactgcttct 60
ggagcctgga tgcagtagca ttcttccact gattcagtga gtaactgtta ggtgggtccc 120
taagggatta ggtatttcat cactgagcta accctggcta tcattctgct tttcttggct 180
gtctttcaga tttgacttta tttctaaaaa tatttcaatg ggtcatatca cagattcttt 240
ttttttaaat taaagtaaca tttccaatct actaatgcta atactgtttc gtatttatag 300
ctgatttgat ggagttggac atggccatgg aaccagacag aaaagcggct gttagtcact 360
ggcagcaaca gtcttacctg gactctggaa tccattctgg tgccactacc acagctcctt 420
ctctgagtgg taaaggcaat cctgaggaag aggatgtgga tacctccaa gtctgtatg 480
agtgggaaca gggattttct cagtcttca ctcaagaaca agtagctggg aagagtatta 540
tttttcattg ccttactgaa agtcagaatg cagttttgag aactaaaaag ttagtgtata 600
atagtttaaa taaaatgttg tgggaagaa aagagagtaa tagcaatgtc acttttacca 660
tttaggatag caaatactta ggtaaatgct gaactgtgga tagtgagtgt tgaattaacc 720
ttttccagat attgatggac agtatgcaat gactcgagct cagagggtag gagctgctat 780

<210> 6
 <211> 2513
 <212> DNA
 <213> Homo sapiens

<400> 6
 cgcctcccg cccctcccc gcccgacagc ggccgctcgg gccccggctc tcggttataa 60
 gatggcggcg ctgagcgggtg gcggtggtgg cggcgcgagg cggggccagg ctctgttcaa 120
 cggggacatg gagcccagg cggcgccgg cgccggcgcc gcggcctctt cggctgcgga 180
 ccttgccatt ccggaggagg tgtggaatat caaacaatg attaatgtga cacaggaaca 240
 tatagaggcc ctattggaca aatttggtgg ggagcataat ccaccatcaa tatatctgga 300
 ggctatgaa gaatacacca gcaagctaga tgcactcaa caaagagaac aacagttatt 360
 ggaatctctg gggaacggaa ctgatttttc tgtttctagc tctgcatcaa tggataccgt 420
 tacatcttct tcctcttcta gcctttcagt gctaccttca tctctttcag tttttcaaaa 480
 tcccacagat gtggcacgga gcaaccccaa gtcaccacaa aaacctatcg tttagagtctt 540
 cctgccaac aaacagagga cagtggtagc tgcaagggtg ggagttacag tccgagacag 600
 tctaaagaaa gcaactgatg tgagaggtct aatcccagag tgctgtgctg tttacagaat 660
 tcaggatgga gagaagaaac caattggttg ggacactgat atttcctggc ttactggaga 720
 agaattgcat gtggaagtgt tggagaatgt tccacttaca acacacaact ttgtacgaaa 780
 aacgtttttc accttagcat tttgtgactt ttgtcgaaag ctgcttttcc aggggtttccg 840
 ctgtcaaaca tgtggttata aatttcacca gcgttgtagt acagaagttc cactgatgtg 900
 tgttaattat gaccaacttg atttgctgtt tgtctccaag ttctttgaac accaccaat 960
 accacaggaa gaggcgtcct tagcagagac tgccctaaaca tctggatcat ccccttccgc 1020
 accgcctcgt gactctattg ggcccaaat tctcaccagt ccgtctcctt caaaatccat 1080
 tccaattcca cagcccttcc gaccagcaga tgaagatcat cgaaatcaat ttgggcaacg 1140
 agaccgatcc tcatcagctc ccaatgtgca tataaacaca atagaacctg tcaatattga 1200
 tgacttgatt agagaccaag gatttcgtgg tgatggagga tcaaccacag gtttgtctgc 1260
 tccccccct gcctcattac ctggctcact aactaacgtg aaagccttac agaaatctcc 1320
 aggacctcag cgagaaagga agtcatcttc atctcagaa gacaggaatc gaatgaaaac 1380
 acttggtaga cgggactcga gtgatgattg ggagattcct gatgggcaga ttacagtggg 1440
 acaaagaatt ggatctggat catttggaac agtctacaag ggaaagtggc atggtgatgt 1500
 ggcagtgaat atgttgaatg tgacagcacc tacacctcag cagttacaag ccttcaaaaa 1560
 tgaagtagga gtactcagga aaacacgaca tgtgaatatc ctactcttca tgggctattc 1620
 cacaagcca caactggcta ttgttaocca gtggtgtgag ggctccagct tgtatcacca 1680

tctccatata	attgagacca	aatttgagat	gatcaaactt	atagatattg	cacgacagac	1740
tgcacagggc	atggattact	tacacgcaa	gtcaatcatc	cacagagacc	tcaagagtaa	1800
taatataatt	cttcatgaag	acctcacagt	aaaaataggt	gatttttggtc	tagctacagt	1860
gaaatctcga	tggagtggtt	cccatcagtt	tgaacagttg	tctggatcca	ttttgtggat	1920
ggcaccagaa	gtcatcagaa	tgcaagataa	aaatccatac	agctttcagt	cagatgtata	1980
tgcatttggt	attgttttgt	atgaattgat	gactggacag	ttaccttatt	caaacatcaa	2040
caacagggac	cagataattt	ttatggtggg	acgaggatac	ctgtctccag	atctcagtaa	2100
ggtacggagt	aactgtocaa	aagccatgaa	gagattaatg	gcagagtgcc	tcaaaaagaa	2160
aagagatgag	agaccactct	ttccccaat	tctcgctctt	attgagctgc	tggcccgctc	2220
attgccaaaa	attcacgcga	gtgcatcaga	accctccttg	aatcgggctg	gtttccaaac	2280
agaggatttt	agtctatatg	cttgtgcttc	tccaaaaaca	cccatccagg	cagggggata	2340
tgggtgcgtt	cctgtccact	gaaacaaatg	agtgagagag	ttcaggagag	tagcaacaaa	2400
aggaaaataa	atgaacatat	gtttgcttat	atgttaaatt	gaataaaata	ctctcttttt	2460
ttttaagggtg	gaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	ccc	2513

<210> 7
 <211> 3724
 <212> DNA
 <213> Homo sapiens

<400> 7	
tctccctcgg	cgccgcgcgc gccgccgcgg gggctgggac ccgatgcggt tagagccgcg 60
gagcctggaa	gagccccgag cgttttctgt ttggggacaac catacatcta attccttaaa 120
gtagttttat	atgtaaaact tgcaaagaat cagaacaatg cctccacgac catcatcagg 180
tgaactgtgg	ggcatccact tgatgcccc aagaatccta gtagaatgtt tactaccaa 240
tggaatgata	gtgacttttag aatgcctccg tgaggctaca ttaataacca taaagcatga 300
actattttaa	gaagcaagaa aataccccct ccatcaactt cttcaagatg aatcttttta 360
cattttctga	agtgttactc aagaagcaga aagggaagaa ttttttgatg aaacaagacg 420
actttgtgac	cttcggcctt ttcaaccctt tttaaaagta attgaaccag taggcaaccg 480
tgaagaaaag	atocctcaatc gagaaattgg ttttgctatc ggcatgccag tgtgtgaatt 540
tgatatgggt	aaagatccag aagtacagga ctccgaaga aatattctga acgtttgtaa 600
agaagctgtg	gatcttaggg acctcaattc acctcatagt agagcaatgt atgtctatcc 660
tccaaatgta	gaatcttcac cagaattgcc aaagcacata tataataaat tagataaagg 720
gcaaataata	gtgggtgatct gggtaatagt ttctccaaat aatgacaagc agaagtatac 780

tctgaaaatc aaccatgact gtgtaccaga acaagtaatt gctgaagcaa tcaggaaaaa	840
aactogaagt atgttgctat cctctgaaca actaaaactc tgtgttttag aatatcaggg	900
caagtatatt ttaaaagtgt gtggatgtga tgaatacttc ctagaaaaat atcctctgag	960
tcagtataag tatataagaa gctgtataat gcttgggagg atgoccaaatt tgatgttgat	1020
ggctaaagaa agccttttatt ctcaactgcc aatggactgt tttacaatgc catcttattc	1080
cagacgcatt tccacagcta caccatatat gaatggagaa acatctacaa aatccctttg	1140
ggttataaat agtgcactca gaataaaaat tctttgtgca acctacgtga atgtaaatat	1200
togagacatt gataagatct atgttcgaac aggtatctac catggaggag aacccttatg	1260
tgacaatgtg aacactcaaa gactaccttg ttccaatccc aggtggaatg aatggctgaa	1320
ttatgatata tacattcctg atcttcctcg tgctgctcga ctttgccctt ccatttgctc	1380
tgttaaaggc cgaaaggggtg ctaaagagga acactgtcca ttggcatggg gaaatataaa	1440
cttgtttgat tacacagaca ctctagtatc tggaaaaatg gctttgaatc tttggccagt	1500
acctcatgga ttagaagatt tgctgaacct tattggtgtt actggatcaa atccaaataa	1560
agaaactcca tgcttagagt tggagtttga ctgggtcagc agtgtggtaa agttcccaga	1620
tatgtcagtg attgaagagc atgccaattg gtctgtatcc cgagaagcag gatttagcta	1680
ttcccacgca ggactgagta acagactagc tagagacaat gaattaaggg aaaatgacaa	1740
agaacagctc aaagcaatct ctacacgaga tcctctctct gaaatcactg agcaggagaa	1800
agattttcta tggagtcaca gacactattg tgtaactatc cccgaaattc tacccaaatt	1860
gcttctgtct gttaaatgga attctagaga tgaagtagcc cagatgtatt gcttggtaaa	1920
agattggcct ccaatcaaac ctgaacaggc tatggaactt ctggactgta attaccaga	1980
tcctatggtt cgaggttttg ctgttcgggtg cttggaaaaa tatttaacag atgacaaact	2040
ttctcagtat ttaattcagc tagtacaggt cctaaaatat gaacaatatt tggataactt	2100
gcttgtgaga tttttactga agaaagcatt gactaatcaa aggattgggc actttttctt	2160
ttggcattta aaatctgaga tgcacaataa aacagttagc cagaggtttg gcctgctttt	2220
ggagtcctat tgtcgtgcat gtgggatgta tttgaagcac ctgaataggc aagtcgaggc	2280
aatggaaaag ctcatctaact taactgacat tctcaaacag gagaagaagg atgaaacaca	2340
aaaggtacag atgaagtttt tagttgagca aatgaggcga ccagatttca tggatgctct	2400
acagggtctt ctgtctcctc taaaccctgc tcatcaacta ggaaacctca ggcttgaaga	2460
gtgtcgaatt atgtcctctg caaaaaggcc actgtggttg aattgggaga acccagacat	2520
catgtcagag ttactgtttc agaacaatga gatcatcttt aaaaatgggg atgatttacg	2580

gcaagatatg ctaacacttc aaattattog tattatggaa aatatctggc aaaatcaagg 2640
 tottgatctt cgaatgttac cttatggttg tctgtcaatc ggtgactgtg tgggacttat 2700
 tgagggtgtg cgaaattctc acactattat gcaaattcag tgcaaaggcg gottgaaagg 2760
 tgcaactgcag ttcaacagcc acacactaca tcagtggctc aaagacaaga acaaaggaga 2820
 aatatatgat gcagccattg acctgtttac acgttcatgt gctggatact gtgtagctac 2880
 cttcattttg ggaattggag atcgtcacia tagtaacatc atgggtgaaag acgatggaca 2940
 actgtttcat atagattttg gacacttttt ggatcacaag aagaaaaaat ttggttataa 3000
 acgagaacgt gtgccatttg ttttgacaca ggattttctta atagtgatta gtaaaggagc 3060
 ccaagaatgc acaaagacaa gagaatttga gaggtttcag gagatgtgtt acaaggctta 3120
 tctagctatt cgacagcatg ccaatctctt cataaatctt ttctcaatga tgcttggtctc 3180
 tggaatgccg gaactacaat cttttgatga cattgcatac attcgaaaga ccctagcctt 3240
 agataaaaact gagcaagagg ctttgagta tttcatgaaa caaatgaatg atgcacatca 3300
 tgggtggctgg acaacaaaaa tggattggat cttccacaca attaaacagc atgcattgaa 3360
 ctgaaaagat aactgagaaa atgaaagctc actctggatt ccacactgca ctgttaataa 3420
 ctctcagcag gcaaagaccg attgcatagg aattgcacia tccatgaaca gcattagaat 3480
 ttacagcaag aacagaaata aaatactata taattttaa atgttaaacy caaacagggt 3540
 ttgatagcac ttaaactagt tcatttcaaa attaaagcttt agaataatgc gcaatttcat 3600
 gttatgcctt aagtccaaaa aggtaaactt tgaagattgt ttgtatcttt ttttaaaaaa 3660
 caaaacaaaa caaaaatccc caaaatatat agaaatgatg gagaaggaaa aaaaaaaaaa 3720
 aaaa 3724

<210> 8
 <211> 845
 <212> DNA
 <213> Homo sapiens

<400> 8
 caagatcctc atcaggagga aaagtaaatt gttcactacc atcctctagt atcctaactc 60
 ggtcttggtg ttggctaact tcagcagtta ctattctgtg actggtgtaa tattaaccaa 120
 ataaattact ggatttggtc tacaaatatt atgtcttaga ttggttcttt cctgtctctg 180
 aaaataaagt cttgcaatga aaataaatta ttttacaaca gttaattagc aatgtaaaat 240
 ttattgaaaa tgtatttgct ttttctgtaa atcatctgtg aatccagagg ggaaaaatat 300
 gacaaagaaa gctatataag atattatttt attttacaga gtaacagact agctagagac 360
 aatgaattaa gggaaaatga caaagaacag ctcaaagcaa tttctacacg agatcctctc 420

tctgaaatca ctgagcagga gaaagatttt ctatggagtc acaggtaagt gctaaaatgg 480
agattctctg tttctttttc tttattacag aaaaaataac tgaatttggc tgatctcagc 540
atgttttttac catacctatt ggaataaata aagcagaatt tacatgattt ttaaactata 600
aacattgcct ttttaaaaac aatggttgta aattgatatt tgtggaaaat catactacat 660
tggtagttag cacattaaat gcttttttct actctgaatt cctgatatga ctttcttttag 720
gattgtttta aatattctag tagtttttagg tcaatttaga tgtgatttag ttggtctaga 780
tattataatt tttaggggtt ccctttcatt tttctttttt cttacgtttc ttcaaatagt 840
ataat 845

<210> 9
<211> 497
<212> DNA
<213> Homo sapiens

<400> 9
gaaagcctct ctaattttgt gacatttgag caaagacctg aaggatttaa catcatttgc 60
tccaaactga ccaaactgtt cttattactt ataggtttca ggagatgtgt tacaaggctt 120
atctagctat tcgacagcat gccaatctct tcataaatct tttctcaatg atgcttggct 180
ctggaatgcc agaactacaa tcttttgatg acattgcata cattcgaaag accctagcct 240
tagataaaaac tgagcaagag gctttggagt atttcatgaa acaaataaat gatgcacatc 300
atggtggctg gacaacaaaa atggattgga tcttccacac aattaaacag catgcattga 360
actgaaaaga taactgagaa aatgaaagct cactctggat tccacactgc actgttaata 420
actctcagca ggcaaagacc gattgcatag gaattgcaca atccatgaac agcattagaa 480
tttacagcaa gaacaga 497

<210> 10
<211> 294
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (246)..(253)
<223> n is a, c, g, or t

<400> 10
ccagtgggtat agaaatcttc gattttttaa ttcttaattt taggttgcag tttcatcact 60
gtctgcggta atcaagtttt tagaactctt atcagatgat tccaactttg gacagtttga 120
actgactact tttgacttca gccagtatat gaaattggat attgcagcag tcagagccct 180
taaccttttt caggtaaaaa aaaaaaaaaa aaaaaaaaaa agggttaaaa atgttgaatg 240

gttaannnnn nnngacagat agtgaagaag gcttagaaaag gagctaaaag agtt 294

<210> 11
 <211> 994
 <212> DNA
 <213> Homo sapiens

<400> 11
 ggacaaagca gtaaaaccga acatatgtct tcaagcagtg agaatacgtc cacaccttca 60
 tctaattgcca agaggcagaa tcagctccat ccaagttctg cacagagtag aagtggctcag 120
 cctcaaaagg ctgccacttg caaagtttct tctattaacc aagaaacaat acagacttat 180
 tgtgtagaag atactccaat atgtttttca agatgtagtt cattatcatc tttgtcatca 240
 gctgaagatg aaataggatg taatcagacg acacaggaag cagattctgc taataccctg 300
 caaatagcag aaataaaaga aaagattgga actaggtcag ctgaagatcc tgtgagcgaa 360
 gttccagcag tgtcacagca ccctagaacc aaatccagca gactgcaggg ttctagttta 420
 tcttcagaat cagccaggca caaagctgtt gaattttctt caggagcgaa atctccctcc 480
 aaaagtgggtg ctcagacacc caaaagtcca cctgaacact atgttcagga gacccactc 540
 atgttttagca gatgtacttc tgtcagttca cttgatagtt ttgagagtcg ttcgattgcc 600
 agctccgttc agagtgaacc atgcagtgga atggtaagtg gcattataag cccagtgat 660
 cttccagata gccctggaca aaccatgcca ccaagcagaa gtaaaacacc tccaccacct 720
 cctcaaacag ctcaaaccaa gcgagaagta cctaaaaata aagcacctac tgctgaaaag 780
 agagagagtg gacctagca agctgcagta aatgctgcag ttcagagggt ccaggttctt 840
 ccagatgctg atactttatt acattttgcc acggaaagta ctccagatgg attttcttgt 900
 tcatccagcc tgagtgtctt gagcctcgat gagccattta tacagaaaga tgtggaatta 960
 agaataatgc ctccagttca ggaaaatgac aatg 994

<210> 12
 <211> 284
 <212> DNA
 <213> Homo sapiens

<400> 12
 gtactgggtg agtatttgat agtgatttaa ccttatgtgt gacatgttct aatatagtca 60
 catttttcatt atttttatta taaggcctgc tgaaaatgac tgaatataaa cttgtggtag 120
 ttggagctgg tggcgtaggc aagagtgcct tgacgatata gctaattcag aatcattttg 180
 tggacgaata tgatccaaca atagaggtaa atcttgtttt aatatgcata ttactgggtg 240
 aggaccattc tttgatacag ataaagggtt ctctgaccat tttc 284

<210> 13
<211> 1648
<212> DNA
<213> Homo sapiens

<400> 13
gccctgactt tcaactctgt ctccttcctc ttcctacagt actcccctgc cctcaacaag 60
atgttttgcc aactggccaa gacctgccct gtgcagctgt gggttgattc cacacccccg 120
cccggcaccc gcgtccgcgc catggccatc tacaagcagt cacagcacat gacggagggtt 180
gtgaggcgct gccccacca tgagcgctgc tcagatagcg atggtgagca gctggggctg 240
gagagacgac agggctgggt gcccagggtc ccaggcctc tgattcctca ctgattgctc 300
ttaggtctgg cccctcctca gcatcttata cgagtggaag gaaatttgcg tgtggagtat 360
ttggatgaca gaaacacttt tcgacatagt gtgggtgtgc cctatgagcc gcctgagggtc 420
tggtttgcaa ctgggggtctc tgggaggagg ggttaagggg ggttgtcagt ggccctccgg 480
gtgagcagta ggggggcttt ctcctgctgc ttatttgacc tccctataac cccatgagat 540
gtgcaaagta aatgggttta actattgcac agttgaaaaa actgaagctt acgaggctaa 600
gggcctcccc tgcttggtg ggcgcagtgg ctcatgcctg taatcccagc actttgggag 660
gccaaggcag gcgatcacg aggttgggag atcgagacca tcctggctaa cggtgaaacc 720
ccgtctctac tgaaaaatac aaaaaaaaaat tagccgggcg tggtgctggg cacctgtagt 780
cccagctact cgggaggctg aggaaggaga atggcgtaga cctgggcggt ggagcttgca 840
gtgagctgag atcacgccac tgcactccag cctgggcgac agagcgagat tccatctcaa 900
aaaaaaaaaa aaaaggcctc ccctgcttgc cacaggctc cccaaggcgc actggcctca 960
tottgggcct gtgttatctc ctaggttggc tctgactgta ccaccatcca ctacaactac 1020
atgtgtaaca gttcctgcat gggcggcata aaccggaggc ccctcctcac catcatcaca 1080
ctggaagact ccaggtcagg agccacttgc caccctgcac actggcctgc tgtgccccag 1140
cctctgcttg ccgctgacct ctggggccac ctcttacoga tttcttccat actactacct 1200
atccacctct catcacatth ccggcgggaa tctccttact gctccactc agtttctctt 1260
tctctggctt tgggacctct taacctgtgg ctctcctcc cactcctgg agctggagct 1320
taggctccag aaaggacaag ggtggttggg agtagatgga gcctggtttt taaatggga 1380
caggtaggac ctgatttctt tactgcctct tgcttctctt ttctatcct gagtagtggg 1440
aatctactgg gacggaacag ctttgagggt cgtgtttgtg cctgtcctgg gagagaccgg 1500
cgcacagagg aagagaatct ccgcaagaaa ggggagcctc accacgagct gccccaggg 1560
agcactaagc gaggtaagca agcaggacaa gaagcgggtg aggagaccaa ggtgagctt 1620
atgcctcaga ttcactttta tcaccttt 1648

<210> 14
<211> 623
<212> DNA
<213> Homo sapiens

<400> 14
taagacatat atttttggtt gaaatacact gaaactgggt tcaaaatatt cgttttaagg 60
gtaaagaaaa aagttaaaaa atctatttac ataaaaaata agaacactga tttttgtgaa 120
tactgggaac tatgaaaata ctatagttga gaccttcaat gactttctag taactcagca 180
gcatctcagg gccaaaaatt taatcagtgg aaaaatagcc tcaattctta ccatccacaa 240
aatggatcca gacaactggt caaactgatg ggacccactc catcgagatt tcaactgtagc 300
tagaccaaaa tcacctatit ttactgtgag gtcttcatga agaaatatat ctgaggtgta 360
gtaagtaaag gaaaacagta gatctcattt tcctatcaga gcaagcatta tgaagagttt 420
aggtaagaga tctaatttct ataattctgt aatataatat tctttaaaac atagtacttc 480
atctttcctc ttagagtcaa taagtatgtc taaaacaatg attagttcta tttagcctat 540
ataacctgct tttaagattt ttggggcttg aaatgtgtta ggatgaggtg agatgctttc 600
ctaagtttat aggagaacct aaa 623