

(12) STANDARD PATENT APPLICATION (11) Application No. **AU 2003203562 A1**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Assessing colorectal cancer

(51)⁷ International Patent Classification(s)
C12Q 001/68 C12N 015/12

(21) Application No: **2003203562** (22) Date of Filing: **2003.03.28**

(30) Priority Data

(31)	Number 60/368798	(32)	Date 2002.03.29	(33)	Country US
------	----------------------------	------	---------------------------	------	----------------------

(43) Publication Date: 2003.10.23

(43) Publication Journal Date: **2003.10.23**

(71) Applicant(s)
Ortho-Clinical Diagnostics, Inc

(72) Inventor(s)
Wang, Yixin

(74) Agent / Attorney
Freehills Carter Smith Beadle, Level 43 101 Collins Street, Melbourne, VIC, 3000

Abstract of the Disclosure

Assessing Colorectal Cancer

A method of assessing the presence or absence of colorectal cancer or the likely condition of a person believed to have colorectal cancer is conducted by analyzing the expression of a group of genes. Gene expression profiles in a variety of medium such as microarrays are included as are kits that contain them.

AUSTRALIA

Patents Act 1990

COMPLETE SPECIFICATION STANDARD PATENT

Invention Title: **Assessing colorectal cancer**

The following statement is a full description of this invention, including the best method of performing it known to us:

ASSESSING COLORECTAL CANCER

BACKGROUND

This application claims the benefit of U.S Provisional Application No.60/368,798 filed on March 29, 2002.

This invention relates to diagnostics and prognostics for colorectal cancer based on the gene expression profiles of biological samples.

Colorectal cancer is a heterogenous disease, consisting of tumors thought to emerge through three major molecular mechanisms: 1) mutations in the adenomatous polyposis coli (APC) gene, or the β -catenin gene, combined with chromosomal instability, 2) mutations in DNA mismatch repair genes, such as MLH1, MSH2, PMS1, PMS2 and MSH6, associated with microsatellite instability and mutations in genes containing short repeats, and 3) gene silencing induced by hypermethylation of the promoter regions of tumor suppressor genes. The genetic complement of individual colorectal cancers is likely to include different combinations of genetic instability, specific mutations, and gene silencing. Chromosomal instability (CIN) is a common feature of cancers in general. It implies an aneuploid phenotype, in which whole chromosomes or large parts of them are being lost or gained. Microsomal instability (MIN) is found in diploid tumors with an increased mutation rate in short repeats. Both forms of genetic instability are common in colorectal cancer.

Colorectal cancers thus have complex origins and involve a number of interactions in different biological pathways. Serum markers, histological, and cytological examinations historically used to assist in providing diagnostic, prognostic, or therapy monitoring decisions often do not have desired reliability. Likewise, while use of a single genetic marker (e.g., increased expression of a particular gene) may be beneficial, the diversity of the cancers make it more likely that a portfolio of genetic markers is the best approach.

SUMMARY OF THE INVENTION

The invention is a method of assessing the presence or absence of colorectal cancer or the likely condition of a person believed to have colorectal cancer. In the method, a gene expression

profile of a patient sample is analyzed to determine whether a patient has a colorectal cancer, whether a patient does not have colorectal cancer, whether a patient is likely to get colorectal cancer, or the response to treatment of a patient being treated for colorectal cancer.

Articles used in practicing the methods are also an aspect of the invention. Such articles include gene expression profiles or representations of them that are fixed in machine-readable media such as computer readable media.

Articles used to identify gene expression profiles can also include substrates or surfaces, such as microarrays, to capture and/or indicate the presence, absence, or degree of gene expression.

DETAILED DESCRIPTION

The mere presence or absence of particular nucleic acid sequences in a tissue sample has only rarely been found to have diagnostic or prognostic value. Information about the expression of various proteins, peptides or mRNA, on the other hand, is increasingly viewed as important. The mere presence of nucleic acid sequences having the potential to express proteins, peptides, or mRNA (such sequences referred to as “genes”) within the genome by itself is not determinative of whether a protein, peptide, or mRNA is expressed in a given cell. Whether or not a given gene capable of expressing proteins, peptides, or mRNA does so and to what extent such expression occurs, if at all, is determined by a variety of complex factors. Irrespective of difficulties in understanding and assessing these factors, assaying gene expression can provide useful information about the occurrence of important events such as tumorigenesis, metastasis, apoptosis, and other clinically relevant phenomena. Relative indications of the degree to which genes are active or inactive can be found in gene expression profiles. The gene expression profiles of this invention are used to diagnose and treat patients for colorectal cancer.

Sample preparation requires the collection of patient samples. Patient samples used in the inventive method are those that are suspected of containing diseased cells such as epithelial cells taken from a colon sample or from surgical margins. One useful technique for obtaining suspect samples is Laser Capture Microdissection (LCM). LCM technology provides a way to select the cells to be studied, minimizing variability caused by cell type heterogeneity. Consequently, moderate or small changes in gene expression between normal and cancerous

cells can be readily detected. In a preferred method, the samples comprise circulating epithelial cells extracted from peripheral blood. These can be obtained according to a number of methods but the most preferred method is the magnetic separation technique described in U.S. Patent 6,136,182 assigned to Immunivest Corp which is incorporated herein by reference. Once the sample containing the cells of interest has been obtained, RNA is extracted and amplified and a gene expression profile is obtained, preferably via micro-array, for genes in the appropriate portfolios.

Preferred methods for establishing gene expression profiles include determining the amount of RNA that is produced by a gene that can code for a protein or peptide. This is accomplished by reverse transcriptase PCR (RT-PCR), competitive RT-PCR, real time RT-PCR, differential display RT-PCR, Northern Blot analysis and other related tests. While it is possible to conduct these techniques using individual PCR reactions, it is best to amplify complimentary DNA (cDNA) or complimentary RNA (cRNA) produced from mRNA and analyze it via microarray. A number of different array configurations and methods for their production are known to those of skill in the art and are described in U.S. Patents such as: 5,445,934; 5,532,128; 5,556,752; 5,242,974; 5,384,261; 5,405,783; 5,412,087; 5,424,186; 5,429,807; 5,436,327; 5,472,672; 5,527,681; 5,529,756; 5,545,531; 5,554,501; 5,561,071; 5,571,639; 5,593,839; 5,599,695; 5,624,711; 5,658,734; and 5,700,637; the disclosures of which are incorporated herein by reference.

Microarray technology allows for the measurement of the steady-state mRNA level of thousands of genes simultaneously thereby presenting a powerful tool for identifying effects such as the onset, arrest, or modulation of uncontrolled cell proliferation. Two microarray technologies are currently in wide use. The first are cDNA arrays and the second are oligonucleotide arrays. Although differences exist in the construction of these chips, essentially all downstream data analysis and output are the same. The product of these analyses are typically measurements of the intensity of the signal received from a labeled probe used to detect a cDNA sequence from the sample that hybridizes to a nucleic acid sequence at a known location on the microarray. Typically, the intensity of the signal is proportional to the quantity of cDNA, and thus mRNA, expressed in the sample cells. A large number of such techniques are available and

useful. Preferred methods for determining gene expression can be found in US Patents 6,271,002 to Linsley, et al.; 6,218,122 to Friend, et al.; 6,218,114 to Peck, et al.; and 6,004,755 to Wang, et al., the disclosure of each of which is incorporated herein by reference.

Analysis of the expression levels is conducted by comparing such intensities. This is best done by generating a ratio matrix of the expression intensities of genes in a test sample versus those in a control sample. For instance, the gene expression intensities from a diseased tissue can be compared with the expression intensities generated from normal tissue of the same type (e.g., diseased colon tissue sample vs. normal colon tissue sample). A ratio of these expression intensities indicates the fold-change in gene expression between the test and control samples.

Gene expression profiles can also be displayed in a number of ways. The most common method is to arrange a raw fluorescence intensities or ratio matrix into a graphical dendrogram where columns indicate test samples and rows indicate genes. The data is arranged so genes that have similar expression profiles are proximal to each other. The expression ratio for each gene is visualized as a color. For example, a ratio less than one (indicating down-regulation) may appear in the blue portion of the spectrum while a ratio greater than one (indicating up-regulation) may appear as a color in the red portion of the spectrum. Commercially available computer software programs are available to display such data including "GENESPRING" from Silicon Genetics, Inc. and "DISCOVERY" and "INFER" software from Partek, Inc.

Modulated genes used in the methods of the invention are shown in Table 1. The genes that are differentially expressed are shown as being either up regulated or down regulated in diseased cells. Up regulation and down regulation are relative terms meaning that a detectable difference (beyond the contribution of noise in the system used to measure it) is found in the amount of expression of the genes relative to some baseline. In this case, the baseline is the measured gene expression of a normal cell. The genes of interest in the diseased cells are then either up regulated or down regulated relative to the baseline level using the same measurement method. Diseased, in this context, refers to an alteration of the state of a body that interrupts or disturbs, or has the potential to disturb, proper performance of bodily functions as occurs with the uncontrolled proliferation of cells. Someone is diagnosed with a disease when some aspect of that person's genotype or phenotype is consistent with the presence of the disease. However,

the act of conducting a diagnosis or prognosis includes the determination disease/status issues such as therapy monitoring. In therapy monitoring, clinical judgments are made regarding the effect of a given course of therapy by comparing the expression of genes over time to determine whether the gene expression profiles have changed or are changing to patterns more consistent with normal tissue.

Preferably, levels of up and down regulation are distinguished based on fold changes of the intensity measurements of hybridized microarray probes. A 2.0 fold difference is preferred for making such distinctions or a p-value less than .05. That is, before a gene is said to be differentially expressed in diseased versus normal cells, the diseased cell is found to yield at least 2 more, or 2 times less intensity than the normal cells. The greater the fold difference, the more preferred is use of the gene as a diagnostic. Genes selected for the gene expression profiles of the instant invention have expression levels that result in the generation of a signal that is distinguishable from those of the normal or non-modulated genes by an amount that exceeds background using clinical laboratory instrumentation.

Statistical values can be used to confidently distinguish modulated from non-modulated genes and noise. Statistical tests find the genes most significantly different between diverse groups of samples. The Student's t-test is an example of a robust statistical test that can be used to find significant differences between two groups. The lower the p-value, the more compelling the evidence that the gene is showing a difference between the different groups. Nevertheless, since microarrays measure more than one gene at a time, tens of thousands of statistical tests may be asked at one time. Because of this, there is likelihood to see small p-values just by chance and adjustments for this using a Sidak correction as well as a randomization/permutation experiment can be made. A p-value less than .05 by the t-test is evidence that the gene is significantly different. More compelling evidence is a p-value less than .05 after the Sidak correct is factored in. For a large number of samples in each group, a p-value less than 0.05 after the randomization/permutation test is the most compelling evidence of a significant difference.

Another parameter that can be used to select genes that generate a signal that is greater than that of the non-modulated gene or noise is the use of a measurement of absolute signal difference. Preferably, the signal generated by the modulated gene expression is at least 20%

different than those of the normal or non-modulated gene (on an absolute basis). It is even more preferred that such genes produce expression patterns that are at least 30% different than those of normal or non-modulated genes.

Genes can be grouped so that information obtained about the set of genes in the group provides a sound basis for making a clinically relevant judgment such as a diagnosis, prognosis, or treatment choice. These sets of genes make up the portfolios of the invention. In this case, the judgments supported by the portfolios involve colorectal cancer. Portfolios of gene expression profiles can be comprised of combinations of genes described in Example 3. As with most diagnostic markers, it is often desirable to use the fewest number of markers sufficient to make a correct medical judgment. This prevents a delay in treatment pending further analysis as well inappropriate use of time and resources. In this case, such a minimal portfolio can be comprised of a combination of genes from Example 4.

Preferably, portfolios are established such that the combination of genes in the portfolio exhibit improved sensitivity and specificity relative to individual genes or randomly selected combinations of genes. In the context of the instant invention, the sensitivity of the portfolio can be reflected in the fold differences exhibited by a gene's expression in the diseased state relative to the normal state. Specificity can be reflected in statistical measurements of the correlation of the signaling of gene expression with the condition of interest. For example, standard deviation can be used as such a measurement. In considering a group of genes for inclusion in a portfolio, a small standard deviation in expression measurements correlates with greater specificity. Other measurements of variation such as correlation coefficients can also be used in this capacity.

The most preferred method of establishing gene expression portfolios is through the use of optimization algorithms such as the mean variance algorithm widely used in establishing stock portfolios. This method is described in detail in the co-pending patent application entitled "Portfolio Selection" by Tim Jatkoe, et. al., of equal date hereto. Essentially, the method calls for the establishment of a set of inputs (stocks in financial applications, expression as measured by intensity here) that will optimize the return (e.g., signal that is generated) one receives for using it while minimizing the variability of the return. Many commercial software programs are

available to conduct such operations. “Wagner Associates Mean-Variance Optimization Application”, referred to as “Wagner Software” throughout this specification, is preferred. This software uses functions from the “Wagner Associates Mean-Variance Optimization Library” to determine an efficient frontier and optimal portfolios in the Markowitz sense.is preferred.

Use of this type of software requires that microarray data be transformed so that it can be treated as an input in the way stock return and risk measurements are used when the software is used for its intended financial analysis purposes. For example, when Wagner Software is employed in conjunction with microarray intensity measurements the following data transformation method is employed.

Genes are first pre-selected by identifying those genes whose expression shows at least some minimal level of differentiation. The preferred pre-selection process is conducted as follows. A baseline class is selected. Typically, this will comprise genes from a population that does not have the condition of interest. For example, if one were interested in selecting a portfolio of genes that are diagnostic for breast cancer, samples from patients without breast cancer can be used to make the baseline class. Once the baseline class is selected, the arithmetic mean and standard deviation is calculated for the indicator of gene expression of each gene for baseline class samples. This indicator is typically the fluorescent intensity of a microarray reading. The statistical data computed is then used to calculate a baseline value of $(X * \text{Standard Deviation} + \text{Mean})$ for each gene. This is the baseline reading for the gene from which all other samples will be compared. X is a stringency variable selected by the person formulating the portfolio. Higher values of X are more stringent than lower. Preferably, X is in the range of .5 to 3 with 2 to 3 being more preferred and 3 being most preferred.

Ratios between each experimental sample (those displaying the condition of interest) versus baseline readings are then calculated. The ratios are then transformed to base 10 logarithmic values for ease of data handling by the software. This enables down regulated genes to display negative values necessary for optimization according to the Markman mean-variance algorithm using the Wagner Software.

The preprocessed data comprising these transformed ratios are used as inputs in place of the asset return values that are normally used in the Wagner Software when it is used for

financial analysis purposes.

Once an efficient frontier is formulated, an optimized portfolio is selected for a given input level (return) or variance that corresponds to a point on the frontier. These inputs or variances are the predetermined standards set by the person formulating the portfolio. Stated differently, one seeking the optimum portfolio determines an acceptable input level (indicative of sensitivity) or a given level of variance (indicative of specificity) and selects the genes that lie along the efficient frontier that correspond to that input level or variance. The Wagner Software can select such genes when an input level or variance is selected. It can also assign a weight to each gene in the portfolio as it would for a stock in a stock portfolio.

Determining whether a sample has the condition for which the portfolio is diagnostic can be conducted by comparing the expression of the genes in the portfolio for the patient sample with calculated values of differentially expressed genes used to establish the portfolio. Preferably, a portfolio value is first generated by summing the multiples of the intensity value of each gene in the portfolio by the weight assigned to that gene in the portfolio selection process. A boundary value is then calculated by $(Y \times \text{standard deviation} + \text{mean of the portfolio value for baseline groups})$ where Y is a stringency value having the same meaning as X described above. A sample having a portfolio value greater than the portfolio value of the baseline class is then classified as having the condition. If desired, this process can be conducted iteratively in accordance with well known statistical methods for improving confidence levels. Optionally one can reiterate this process until best prediction accuracy is obtained.

The process of portfolio selection and characterization of an unknown is summarized as follows:

1. Choose baseline class
2. Calculate mean, and standard deviation of each gene for baseline class samples
3. Calculate $(X \times \text{Standard Deviation} + \text{Mean})$ for each gene. This is the baseline reading from which all other samples will be compared. X is a stringency variable with higher values of X being more stringent than lower.
4. Calculate ratio between each Experimental sample versus baseline reading calculated in step 3.
5. Transform ratios such that ratios less than 1 are negative (eg.using Log base 10). (Down

- regulated genes now correctly have negative values necessary for MV optimization).
6. These transformed ratios are used as inputs in place of the asset returns that are normally used in the software application.
 7. The software will plot the efficient frontier and return an optimized portfolio at any point along the efficient frontier.
 8. Choose a desired return or variance on the efficient frontier.
 9. Calculate the Portfolio's Value for each sample by summing the multiples of each gene's intensity value by the weight generated by the portfolio selection algorithm.
 10. Calculate a boundary value by adding the mean Portfolio Value for Baseline groups to the multiple of Y and the Standard Deviation of the Baseline's Portfolio Values. Values greater than this boundary value shall be classified as the Experimental Class.
 11. Optionally one can reiterate this process until best prediction accuracy is obtained.

Alternatively, genes can first be pre-selected by identifying those genes whose expression shows some minimal level of differentiation. The pre-selection in this alternative method is

$$1 \leq \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \quad \mu_t$$

preferably based on a threshold given by μ_t , where μ_t is the mean of the subset

known to possess the disease or condition, μ_n is the mean of the subset of normal samples, and $\sigma_t + \sigma_n$

represent the combined standard deviations. A signal to noise cutoff can also be used by

$$0.5 \leq \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)}$$

pre-selecting the data according to a relationship such as μ_t . This ensures that genes that are pre-selected based on their differential modulation are differentiated in a clinically significant way. That is, above the noise level of instrumentation appropriate to the task of measuring the diagnostic parameters. For each marker pre-selected according to these criteria, a matrix is established in which columns represents samples, rows represent markers and each element is a normalized intensity measurement for the expression of that marker according to the

$$\left| \frac{(\mu_i - I)}{\mu_i} \right|$$

relationship: where I is the intensity measurement.

It is also possible to set additional boundary conditions to define the optimal portfolios. For example, portfolio size can be limited to a fixed range or number of markers. This can be

$$.8 \leq \left| \frac{(\mu_i - MAX_n)}{(\sigma_i + \sigma_n)} \right|$$

done either by making data pre-selection criteria more stringent (e.g., instead

$$0.5 \leq \left| \frac{(\mu_i - MAX_n)}{(\sigma_i + \sigma_n)} \right|$$

of) or by using programming features such as restricting portfolio size. One could, for example, set the boundary condition that the efficient frontier is to be selected from among only the most optimal 10 genes. One could also use all of the genes pre-selected for determining the efficient frontier and then limit the number of genes selected (e.g., no more than 10).

The process of selecting a portfolio can also include the application of heuristic rules. Preferably, such rules are formulated based on biology and an understanding of the technology used to produce clinical results. More preferably, they are applied to output from the optimization method. For example, the mean variance method of portfolio selection can be applied to microarray data for a number of genes differentially expressed in subjects with breast cancer. Output from the method would be an optimized set of genes that could include some genes that are expressed in peripheral blood as well as in diseased breast tissue. If sample used in the testing method are obtained from peripheral blood and certain genes differentially expressed in instances of breast cancer could also be differentially expressed in peripheral blood, then a heuristic rule can be applied in which a portfolio is selected from the efficient frontier excluding those that are differentially expressed in peripheral blood. Of course, the rule can be applied prior to the formation of the efficient frontier by, for example, applying the rule during data pre-selection.

Other heuristic rules can be applied that are not necessarily related to the biology in question. For example, one can apply the rule that only a given percentage of the portfolio can

be represented by a particular gene or genes. Commercially available software such as the Wagner Software readily accommodates these types of heuristics. This can be useful, for example, when factors other than accuracy and precision (e.g., anticipated licensing fees) have an impact on the desirability of including one or more genes.

One method of the invention involves comparing gene expression profiles for various genes (or portfolios) to conduct diagnoses as described above. The gene expression profiles of each of the genes comprising the portfolio are fixed in a medium such as a computer readable medium. This can take a number of forms. For example, a table can be established into which the range of signals (e.g., intensity measurements) indicative of disease is input. Actual patient data can then be compared to the values in the table to determine whether the patient samples are normal or diseased. In a more sophisticated embodiment, patterns of the expression signals (e.g., fluorescent intensity) are recorded digitally or graphically. The gene expression patterns from the gene portfolios used in conjunction with patient samples are then compared to the expression patterns. Pattern comparison software can then be used to determine whether the patient samples have a pattern indicative of the disease in question. Of course, these comparisons can also be used to determine whether the patient results are normal. The expression profiles of the samples are then compared to the portfolio of a normal or control cell. If the sample expression patterns are consistent with the expression pattern for a colorectal cancer then (in the absence of countervailing medical considerations) the patient is diagnosed as positive for colorectal cancer. If the sample expression patterns are consistent with the expression pattern from the normal/control cell then the patient is diagnosed negative for colorectal cancer. Numerous well known methods of pattern recognition are available. The following references provide some examples:

Weighted Voting:

Golub, TR., Slonim, DK., Tamaya, P., Huard, C., Gaasenbeek, M., Mesirov, JP., Coller, H., Loh, L., Downing, JR., Caligiuri, MA., Bloomfield, CD., Lander, ES. *Molecular classification of cancer: class discovery and class prediction by gene expression monitoring*. Science 286:531-537, 1999

Support Vector Machines:

Su, AI., Welsh, JB., Sapinoso, LM., Kern, SG., Dimitrov, P., Lapp, H., Schultz, PG., Powell, SM., Moskaluk, CA., Frierson, HF. Jr., Hampton, GM. *Molecular classification of human carcinomas by use of gene expression signatures.* Cancer Research 61:7388-93, 2001

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, CH., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, JP., Poggio, T., Gerald, W., Loda, M., Lander, ES., Gould, TR. *Multiclass cancer diagnosis using tumor gene expression signatures* Proceedings of the National Academy of Sciences of the USA 98:15149-15154, 2001

K-nearest Neighbors:

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, CH., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, JP., Poggio, T., Gerald, W., Loda, M., Lander, ES., Gould, TR. *Multiclass cancer diagnosis using tumor gene expression signatures* Proceedings of the National Academy of Sciences of the USA 98:15149-15154, 2001

Correlation Coefficients:

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer.

Nature. 2002 Jan 31;415(6871):530-6.

The gene expression profiles of this invention can also be used in conjunction with other non-genetic diagnostic methods useful in cancer diagnosis, prognosis, or treatment monitoring. For example, in some circumstances it is beneficial to combine the diagnostic power of the gene

expression based methods described above with data from conventional markers such as serum protein markers (e.g., carcinoembryonic antigen). A range of such markers exists including such analytes as CA19-9, CA 125, CK-BB, and Guanylyl Cyclase C. In one such method, blood is periodically taken from a treated patient and then subjected to an enzyme immunoassay for one of the serum markers described above. When the concentration of the marker suggests the return of tumors or failure of therapy, a sample source amenable to gene expression analysis is taken. Where a suspicious mass exists, a fine needle aspirate is taken and gene expression profiles of cells taken from the mass are then analyzed as described above. Alternatively, tissue samples may be taken from areas adjacent to the tissue from which a tumor was previously removed. This approach can be particularly useful when other testing produces ambiguous.

Combining the use of genetic markers with other diagnostics is most preferred when the reliability of the other diagnostic is suspect. For example, it is known that serum levels of CEA can be substantially affected by factors having nothing to do with a patient's cancer status. It can be beneficial to conduct a combination gene expression/CEA assay when a patient being monitored following treatment for colon cancer shows heightened levels of routine CEA assays.

Articles of this invention include representations of the gene expression profiles useful for treating, diagnosing, prognosticating, and otherwise assessing diseases. These profile representations are reduced to a medium that can be automatically read by a machine such as computer readable media (magnetic, optical, and the like). The articles can also include instructions for assessing the gene expression profiles in such media. For example, the articles may comprise a CD ROM having computer instructions for comparing gene expression profiles of the portfolios of genes described above. The articles may also have gene expression profiles digitally recorded therein so that they may be compared with gene expression data from patient samples. Alternatively, the profiles can be recorded in different representational format. A graphical recordation is one such format. Clustering algorithms such as those incorporated in "GENESPRING" and "DISCOVER" computer programs mentioned above can best assist in the visualization of such data.

Different types of articles of manufacture according to the invention are media or formatted assays used to reveal gene expression profiles. These can comprise, for example,

microarrays in which sequence complements or probes are affixed to a matrix to which the sequences indicative of the genes of interest combine creating a readable determinant of their presence. Alternatively, articles according to the invention can be fashioned into reagent kits for conducting hybridization, amplification, and signal generation indicative of the level of expression of the genes of interest for detecting colorectal cancer.

Kits made according to the invention include formatted assays for determining the gene expression profiles. These can include all or some of the materials needed to conduct the assays such as reagents and instructions.

The invention is further illustrated by the following non-limiting examples.

Examples: Genes analyzed according to this invention are identified by reference to Gene ID Numbers in the GenBank database. These are typically related to full-length nucleic acid sequences that code for the production of a protein or peptide. One skilled in the art will recognize that identification of full-length sequences is not necessary from an analytical point of view. That is, portions of the sequences or ESTs can be selected according to well-known principles for which probes can be designed to assess gene expression for the corresponding gene.

Example 1- Sample Handling and LCM.

Twenty-seven fresh frozen tissue samples were collected from patients who had surgery for a colorectal tumor. Nineteen of the samples were colorectal malignancy specimens, and eight of the samples were of normal colon mucosa. The tissues were snap frozen in liquid nitrogen within 20-30 minutes of harvesting, and stored at -80°C thereafter. For laser capture, the samples were cut ($6\mu\text{m}$), and one section was mounted on a glass slide, and the second on film (P.A.L.M.), which had been fixed onto a glass slide (Micro Slides Colorfrost, VWR Scientific, Media, PA). The section mounted on a glass slide was after fixed in cold acetone, and stained with Mayer's Haematoxylin (Sigma, St. Louis, MO). A pathologist analyzed the samples for diagnosis and grade. The clinical stage was estimated from the accompanying surgical pathology and clinical reports, using the Dukes classification. The section mounted on film was after fixed for five minutes in 100% ethanol, counter stained for 1 minute in eosin/100% ethanol (100 μg of

Eosin in 100ml of dehydrated ethanol), quickly soaked once in 100% ethanol to remove the free stain, and air dried for 10 minutes.

Two of the colorectal adenocarcinomas were of grade 1, 10 of grade 2, and 5 of grade 3. One of the malignant samples was a carcinoid tumor of the caecum, and one a metastatic melanoma lesion. Two of the adenocarcinoma samples represented the mucinous subtype, and one the signet cell subtype. The Dukes staging of the adenocarcinomas divided them as follows: Dukes A: 2, Dukes B: 5, Dukes C: 7, Dukes D: 3. Six of the adenocarcinomas had been irradiated preoperatively.

Before use in LCM, the membrane (LPC-MEMBRANE PEN FOIL 1.35 μ m No 8100, P.A.L.M. GmbH Mikrolaser Technologie, Bernried, Germany) and slides were pretreated to abolish RNases, and to enhance the attachment of the tissue sample onto the film. Briefly, the slides were washed in DEP H₂O, and the film was washed in RNase AWAY (Molecular Bioproducts, Inc., San Diego, CA) and rinsed in DEP H₂O. After attaching the film onto the glass slides, the slides were baked at +120°C for 8 hours, treated with TI-SAD (Diagnostic Products Corporation, Los Angeles, CA, 1:50 in DEP H₂O, filtered through cotton wool), and incubated at +37°C for 30 minutes. Immediately before use, a 10 μ l aliquot of RNase inhibitor solution (Rnasin Inhibitor 2500U=33U/ μ l N211A, Promega GmbH, Mannheim, Germany, 0.5 μ l in 400 μ l of freezing solution, containing 0.15 mol NaCl, 10 mmol Tris pH 8.0, 0.25 mmol dithiothreitol) was spread onto the film, where the tissue sample was to be mounted.

The tissue sections mounted on film were used for LCM. Approximately 2000 epithelial cells/sample were captured using the PALM Robot-Microbeam technology (P.A.L.M. Mikrolaser Technologie, Carl Zeiss, Inc., Thornwood, NY), coupled into Zeiss Axiovert 135 microscope (Carl Zeiss Jena GmbH, Jena, Germany). The surrounding stroma in the normal mucosa, and the occasional intervening stromal components in cancer samples, were included. The captured cells were put in tubes in 100% ethanol and preserved at -80°C.

Example 2- RNA Extraction and Amplification.

Zymo-Spin Column (Zymo Research, Orange, CA 92867) was used to extract total RNA from the LCM captured samples. About 2 ng of total RNA was resuspended in 10 μ l of water

and 2 rounds of the T7 RNA polymerase based amplification were performed to yield about 50 ug of amplified RNA.

Example 3- cDNA Microarray Hybridization and Quantitation.

A set of cDNA microarrays consisting of approximately 20,000 human cDNA clones was used to test the samples. About 30 plant genes were also printed on the microarrays as a control for non-specific hybridization. Cy3-labeled cDNA probes were synthesized from 5 ug of aRNA of the LCM captured cells. The probes were purified with Qiagen's Nucleotide Removal Columns and then hybridized to the microarrays for 14-16 hours. The slides were washed and air-dried before scanning. cDNA microarrays were scanned for cy3 fluorescence and ImaGene software (Biodiscovery, Los Angeles, CA) was used for quantitation. For each cDNA clone, four measurements were obtained using duplicate spots and duplicate arrays and the intensities were averaged.

cDNAs were printed on amino silane-coated slides (Corning) with a Generation III Micro-array Spotter (Molecular Dynamics). The cDNAs were PCR amplified, purified (Qiagen PCR purification kit), and mixed 1:1 with 10 M NaSCN printing buffer. Prior to hybridization micro-arrays were incubated in isopropanol at room temperature for 10 min. The probes were incubated at 95°C for 2 min, at room temperature for 5 min, and then applied to three replicate slides. Cover slips were sealed onto the slides with DPX (Fluka) and incubated at 42°C overnight. Slides were then washed at 55°C for 5min in 1X SSC/0.2% SDS and 0.1X SSC/0.2% SDS, dipped in 0.1X SSC and dried before being scanned by a GenIII Array Scanner (Molecular Dynamics). The fluorescence intensity for each spot was analyzed with AUTOGENE software (Biodiscovery, Los Angeles).

Chip intensities were linearly normalized forcing the intensity reading at the 75th percentile equivalent to a value of 100 on each chip. Every gene on the chip was normalized to itself by dividing the intensity reading for that gene by the median of the gene's expression value readings over all the samples. Prior to clustering, genes that did not have an intensity reading of 100 or greater in at least one sample were filtered out in order to limit the background affect on the similarity metrics. A set of 6,225 genes was selected for clustering analysis. Hierarchical

clustering was performed using correlation as a measure of similarity, which groups together samples with genes that are showing positive changes at the same time without any consideration for negative changes (Silicon Genetics, Sunnyville, CA). Each of the major nodes in the dendrogram was then considered a subgroup of samples. Differentially expressed genes were identified by comparing each tumor subgroup to the normal group. The selection was based on a

$$1 \leq \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \quad \mu_t$$

signal to noise measurement threshold given by $\frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)}$, where μ_t is the mean of the tumor subset, μ_n is the mean of the subset of normal samples, and $\sigma_t + \sigma_n$ represent the combined standard deviations. The within-group coefficient of variation of the intensity readings of a gene had to be less than 0.33, for the gene to be included in the pair-wise comparisons. The median of the tumor group over the median of the normal group had to be greater than, or equal to 2 for up-regulation, and less than, or equal to 0.5 for down-regulation. If a gene met all the criteria, it was selected. The genes selected in all the comparisons were considered consistently dysregulated in colorectal cancer. The p-values for the statistical significance were calculated using a T-test assuming unequal variance. The gene set for clustering was also subjected to principal component analysis (PCA) using a software package (Partek, St Louis, MO). The data was then projected onto the reduced 3-dimensional space. The normal and tumor colorectal samples were represented by the projected expression levels.

A list of genes with large up-regulated differentials was created to distinguish between the tumor and normal samples. One-hundred and twenty-three genes were preselected by using

$$0.5 \leq \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)}$$

as a signal to noise cutoff. A ratio equal to, or greater than 1.5 was the

$$0.9 \leq \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)}$$

minimal criterion for up-regulation. Genes were also included if $\frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \geq 0.9$. A portfolio of four genes was established, each having at least a three fold expression differential between tumor and normal cells.

Differentially Expressed Genes in Colorectal Cancer. Thirty-nine genes were

differentially expressed in all tumor samples as compared to normal colon mucosa. Thirty-seven of them were significantly down-regulated in all the tumors, except for an outlier. Two of them were up-regulated. The identities of the genes were verified by sequencing the cDNA clones placed on the microarray. Results are shown in Table 1.

Table 1 Modulated Genes

ACCESSION	GENE DESCRIPTION	MEAN SIGNAL INTENSITY (NORMAL)	MEAN SIGNAL INTENSITY (TUMOR)	P-VALUE	
AF071569	CaM kinase II gene subtype delta 2.	93	39	4.64E-09	Seq. ID No. 1
AB014530	Homo sapiens mRNA for KIAA0630 protein	108	50	4.83E-07	Seq. ID No.2
AK000319	Human cDNA KIAA0630	236	69	7.84E-06	Seq. ID No.3
U81504	beta-3A-adaptin subunit of the AP-3 complex mRNA.	241	75	3.52E-05	Seq. ID No.4
AB011166	Human cDNA KIAA0594	116	55	3.53E-05	Seq. ID No.5
AB040914	Human cDNA KIAA1481	187	59	8.85E-05	Seq. ID No.6
AK025205	Human cDNA FLJ21552	322	97	0.00013	Seq. ID No.7
AJ278219	Fatty acid hydroxylase	143	53	0.00011	Seq. ID No.8
AB046854	Human cDNA KIAA1634	142	59	0.00020	Seq. ID No.9
R00585	Unknown	149	57	1.28E-09	Seq. ID No.10
S45844	Spi-B transcription factor	140	43	0.00043	Seq. ID No.11
X98311	Carcinoembryonic antigen family member 2 (CEM2)	6137	223	0.00044	Seq. ID No.12
BAA78050	NADPH oxidoreductase homolog	153	84	0.00048	Seq. ID No.40
N72128	Unknown	164	77	0.00068	Seq. ID No.13
AB040955	Human cDNA KIAA1552	334	120	0.00067	Seq. ID No.14
AF125101	HSPC040 protein	363	115	0.0011	Seq. ID No.15
AB023229	Human cDNA KIAA1012	263	88	0.00099	Seq. ID No.16
N95761	a-L-fucosidase gene	429	104	0.00047	Seq. ID No.17
AK025033	Human cDNA FLJ21380	180	85	0.0010	Seq. ID No.18
L10844	Human cellular growth regulating protein	206	101	0.0013	Seq. ID No.19
H96534	H.sapiens mRNA for gp25L2 protein.	147	58	0.0015	Seq. ID No.20
AK001521	Human cDNA FLJ10659	157	60	0.0019	Seq. ID No.21
AF151039	HSPC205 protein	117	60	0.0017	Seq. ID No.22
AF052059	SEL 1L protein	168	53	0.0016	Seq. ID No.23
N24597	Unknown	166	62	0.0016	Seq. ID No.24
AK001950	Inner centromere protein	148	64	0.0029	Seq. ID No.25
BAA02649	Macrophage scavenger receptor type I	118	44	0.0031	Seq. ID No.41
N75004	Unknown	98	48	0.0031	Seq. ID No.26
W16916	Human cDNA KIAA0260	162	61	0.0037	Seq. ID No.27
X52001	H.sapiens endothelin 3 mRNA.	89	33	0.0042	Seq. ID No.28
T50788	Unknown	364	102	0.0059	Seq. ID No.38
AJ005866	Putative Sqv-7 like protein	381	163	0.0049	Seq. ID No.29
AF113535	MAID protein	218	100	0.0053	Seq. ID No.39
AB037789	Human cDNA KIAA1368	164	62	0.0068	Seq. ID No.30
M33987	Carbonic anhydrase	652	46	0.0074	Seq. ID No.31
M77830	Desmoplakin 1 (DPI)	184	81	0.0092	Seq. ID No.32
H81220	EST domain transcription factor ELF1	113	55	0.017	Seq. ID No.33
AF000592	Human chromosome 21q11-q21 genomic clone	33	69	1.16E-05	Seq. ID No.35
AK021701	Human cDNA FLJ11639	31	63	0.00070	Seq. ID No.36

Example 4: Optimized Portfolio for Colorectal Tumors. The mean-variance optimization algorithm was used to generate a multiple gene-based signature, where the genes that are included can be used in combination to distinguish between the normal and tumor samples. Intensity measurements were processed using the samples and microarrays described in Examples 1-3. The data to be analyzed was first preselected based on a pre-specified 5-fold differential between tumor and normal cells. The expression data from genes preselected

according to this criteria were then used as follows. The mean and standard deviation of the intensity measurements for each gene were calculated using the non-metastatic samples as the baseline. A discriminating value of $X \times (\text{Standard Deviation} + \text{Mean})$ was then calculated for each baseline gene (X was assigned a value of 3). This value was used to ensure the resulting portfolio would be stringent. A ratio of the discriminating value to the baseline value was then calculated for each metastatic sample. This ratio was then converted to a common logarithm. This data was then imported into Wagner Software which produced an efficient frontier from which a portfolio of 4 genes was selected. The set included an unknown sequence, procollagen type I, large subunit of ribosomal protein L21, and fibronectin. These genes are identified as Seq. ID No. 42, Seq. ID No. 43, Seq. ID No. 44, and Seq. ID No. 45. Alternatively, a combination of genes used to make up the portfolio can be used to produce diagnostic information that is useful for making clinical decisions regarding colorectal cancer. This is particularly beneficial in the case when a combination of genes selected from the portfolio are combined with additional markers (genetic or not).

Optimized Gene Portfolio:

>gi|1264443|gb|N92134.1|N92134 za23f09.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:293417 5' similar to gb|M87908|HUMALNE32 Human carcinoma cell-derived Alu RNA transcript, (rRNA); gb:X57025_rna1 INSULIN-LIKE GROWTH FACTOR IA PRECURSOR (HUMAN)

>gi|2221047|gb|AA490172.1|AA490172 ab06b08.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:839991 3' similar to gb:J03464 PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR (HUMAN)

>gi|2188918|gb|AA464034.1|AA464034 zx86b09.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:810617 5' similar to SW:RL21_HUMAN P46778 60S RIBOSOMAL PROTEIN L21.

>gi|834491|gb|R62612.1|R62612 yi12d01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:139009 3' similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR (HUMAN);

Using a different set of criteria but the same method, a further four gene portfolio was selected by the software. These are Seq. ID no. 46, Seq. ID No. 47, Seq. ID No. 48 and Seq. ID No. 49. Two genes overlap with the first four-gene portfolio. The two optimized portfolios can also be combined to form a six-gene portfolio.

Optimized Gene Portfolio:

>gi|2114953|gb|AA431245.1|AA431245 zw78d06.r1 Soares_testis_NHT Homo sapiens
cDNA clone IMAGE:782315 5' similar to WP:F36H1.2 CE05814 ANKYRIN LIKE

>gi|2156172|gb|AA443497.1|AA443497 zw34d03.r1 Soares ovary tumor NbHOT Homo
sapiens cDNA clone IMAGE:771173

>gi|2221047|gb|2|AA490172 ab06b08.s1 Stratagene fetal retina 937202 Homo
sapiens cDNA clone IMAGE:839991 3' similar to gb:J03464 PROCOLLAGEN ALPHA 2(I)
CHAIN PRECURSOR (HUMAN)

>gi|1264443|gb|N92134.1|N92134 za23f09.r1 Soares fetal liver spleen 1NFLS Homo
sapiens cDNA clone IMAGE:293417 5' similar to gb|M87908|HUMALNE32 Human
carcinoma cell-derived Alu RNA transcript, (rRNA); gb:X57025_rnal INSULIN-LIKE
GROWTH FACTOR IA PRECURSOR (HUMAN);

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

I claim:

1. A method of assessing colorectal cancer status comprising identifying differential modulation of each gene (relative to the expression of the same genes in a normal population) in a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
2. The method of claim 1 wherein there is at least a 2 fold difference in the expression of the modulated genes.
3. The method of claim 1 wherein the p-value indicating differential modulation is less than .05.
4. The method of claim 1 further comprising employing a colorectal diagnostic that is not genetically based.
5. The method of claim wherein the cancer marker that is not genetically based is selected from the group consisting of carcinoembryonic antigen, CA19-9, CA 125, CK-BB, and Guanylyl Cyclase C.
6. A diagnostic portfolio comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
7. The diagnostic portfolio of claim 6 in a matrix suitable for identifying the differential expression of the genes contained therein.
8. The diagnostic portfolio of claim 7 wherein said matrix is employed in a microarray.
9. The diagnostic portfolio of claim 8 wherein said microarray is a cDNA microarray.

10. The diagnostic portfolio of claim 8 wherein said microarray is an oligonucleotide microarray.
13. A diagnostic portfolio comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
14. A kit for diagnosing colorectal cancer comprising isolated nucleic acid sequences, their compliments, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
15. The kit of claim 14 further comprising reagents for conducting a microarray analysis.
16. The kit of claim 14 further comprising a medium through which said nucleic acid sequences, their compliments, or portions thereof are assayed.
17. A method of assessing response to treatment for colorectal cancer comprising identifying differential modulation of each gene (relative to the expression of the same genes in a normal population) in a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
18. The method of claim 17 wherein the assessment of the response to therapy includes a determination of whether the patient is improving, not improving, relapsing, likely to improve, or likely to relapse.
19. Articles for assessing colorectal cancer statuts comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
20. Articles for assessing colorectal cancer status comprising representations of isolated nucleic

acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.

Ortho-Clinical Diagnostics, Inc.

By Freehills Carter Smith Beadle

Registered Patent Attorneys for the Applicant

28 March 2003

SEQUENCE LISTING

<110> WANG, YIXIN

<120> COLORECTAL CANCER DIAGNOSTICS

<130> CDS 267 US NP

<140> TBD

<141> 2003-03-21

<150> 60/368,798

<151> 2002-03-29

<160> 49

<170> PatentIn version 3.1

<210> 1

<211> 1500

<212> DNA

<213> human

<400> 1

```
atggcttga ccaccacctg caccagggtc acggacgagt atcagcttt cgaggagctt   60
ggaaaggggg cattctcagt ggtgagaaga tgtatgaaaa ttctactgg acaaggatat   120
gctgccaaaa ttatcaacac caaaaagctt tctgctaggg atcatcagaa actagaaaga   180
gaagctagaa tctgccgtct ttgaagcac cctaatttg tgcgactca ttagatcata   240
tcagaagagg gctttacta cttggtgtt gatttagtta ctggaggaga actgtttgaa   300
gacatagtgg caagagaata ctacagtga gctgatgcca gtcattgtat acagcagatt   360
ctagaaagtg ttaatcattg tcacctaat ggcatagttc acaggacct gaagcctgag   420
aattgcttt tagctagcaa atccaagga gcagctgtga aattggcaga ctttggtta   480
gccatagaag ttcaagggga ccagcaggcg tggtttggt ttgctggcac acctggatat   540
ctttctccag aagttttacg taaagatcct tatggaaagc cagtggatat gtgggcatgt   600
ggtgtcattc tctatattct acttgtggg tatccacct tctgggatga agaccaacac   660
agactctatc agcagatcaa ggctggagct tatgatttc catcaccaga atgggacacg   720
gtgactcctg aagccaaaga cctcatcaat aaaatgctta ctatcaacc tgccaaacgc   780
atcacagcct cagaggcact gaagcaccca tggatctgtc aacgttctac tgttgcttc   840
atgatgcaca gacaggagac ttagactgc ttgaagaaat ttaatgctag aagaaaacta   900
aagggtgcca tcttgacaac tatctggct acaaggaatt tctcagcagc caagagtttg   960
```

ttgaagaaac cagatggagt aaaggagtca actgagagtt caaatacaac aattgaggat 1020
gaagatgtga aagcacgaaa gcaagagatt atcaaagtca ctgaacaact gatcgaagct 1080
atcaacaatg gggactttga agcctacaca aaaatctgtg acccaggcct tactgctttt 1140
gaacctgaag cttgggtaa ttagtggaa gggatggatt ttaccgatt ctacttgaa 1200
aatgctttgt ccaaaagcaa taaaccaatc cacactatta ttctaaacc tcattgacat 1260
ctggtagggg atgatgccgc ctgcatagca tatattaggc tcacacagta catggatggc 1320
agtggaatgc caaagacaat gcagtcagaa gagactctg tgtggcaccg ccgggatgga 1380
aagtggcaga atgttcattt tcctgctcg gggcaccaa cagtacccat caagccaccc 1440
tgtattccaa atgggaaaga aaacttctca ggaggcacct cttgtggca aaacatctga 1500

<210> 2
<211> 5761
<212> DNA
<213> human

<400> 2
cacaccgag tatcggtgc ctttactct gagctgcgca gccggccggc cggcgctggt 60
tgaacagact gccgctgtac tggcgtggcc tggagggact cagcaaattc tctgccttc 120
aacttgcaa cagttgcctg gggtagctct acacaactct gtccagccca cagcaatgat 180
tccagaggcc atggggagtg gacagcagct agctgactgg aggaatgcc actctcatgg 240
caaccgtac agcactatca tgcagcagcc atccttctg actaaccatg tgacattggc 300
cactgctcag cctctgaatg ttggtgtgc ccatgtgtc agacaacaac aatccagttc 360
cctccctcg aagaagaata agcagtcagc tccagtctct tccaagtcct ctctagatgt 420
tctgcctcc caagtctatt ctctggttg gagcagtcct ctccgcacca catcttctta 480
taattcctg gtccctgtcc aagatcagca tcagcccatc atcattccag atactccag 540
ccctcctgtg agtgcatca ctatccgaag tgacactgat gaggaagagg acaacaata 600
caagcccagt agctctggac tgaagccaag gtctaattgc atcagttatg tcaactgtcaa 660
tgatttcca gactctgact ctctttgag cagcccttat tccactgata ccctgagtgc 720
tctccgaggc aatagtggat ccgttttga ggggcctggc agagttgtgg cagatggcac 780
tggcaccgc actatcattg tgctccact gaaaactcag cttggtgact gcactgtagc 840
aaccaggcc tcaggtctcc tgagcaataa gactaagcca gtcgctcag tgagtgggca 900

gtcatctgga tgctgtatca cccccacagg gtatcgagct caacgcgggg ggaccagtgc 960
agcacaacca ctcaatctta gccagaacca gcagtcacg gcggctcaa cctcacagga 1020
gagaagcagc aaccagccc ccgcaggca gcaggcggtt gtggccctc tctccaagc 1080
cccctacacc ttccagcatg gcagcccgt acactcgaca gggcacccac acctgcccc 1140
ggccccgtct cacctgccaa gccaggctca tctgtatag tatgtgccc cgacttctgc 1200
tgctgcactg ggctcaacca gtccattgc tcatctttc tccccacagg gticctcaag 1260
gcatgtgca gcctatacca ctaccctag cacttggtg caccaggtec ctgtcagtgt 1320
tgggccagc ctctcactt ctgccagct ggccccgtct cagtaccaac accagttgc 1380
cacccaatcc tacattgggt ctcccaggc ctcaacaatt tacactggat acccgctgag 1440
tctaccaag atcagccagt attctactt atagtgtg agcatgagg aggaggaatc 1500
atggctacct tctctggcc ctgcgttct aatattggc tatggagaga tctccttta 1560
ccctctgaa attcttagc cagcaactg ttctgcagg gccactgaa gcagaaggtt 1620
ttctctggg ggaacctgc tcagtgtga ctgcattgt gtagcttcc caaagttgc 1680
cctatttta aattcattt ttgtgaca gtaatttgg tacttgaag agttcagatg 1740
cccattctt gcagttacca aggaagagag attgtctga agttaccctc tgaaaaatat 1800
ttgtctctc tgacttgatt tctataaatg ctttaaaaa caagtgaagc ccctcttat 1860
ttcatttgt gttatttga ttgctgtca ggaaaaatgc tgatagaagg agtgaaatc 1920
tgatgacaaa aaaagaaaa ttacttttg ttgtttata aactcagact tgctatttt 1980
attttaaag cggcttacac aatctccct ttgtttatg gacatttaa ctacagagt 2040
ttcagtttg tttaatgct atattatact taatgggcaa ttgtatttt tgcaaaactg 2100
gttacgtatt actctgtgt actattgaga ttcttcaat tgctcctgt ttgttataa 2160
agtagtgtt aaaaggcagc tcaccattg ctggttaact aatgtgagag aatccatctc 2220
tgctgaaaa caccaagtat tctttttaa tgaagcacca tgaattctt ttaaattat 2280
ttttaaaag tcttctctc tctgattcag cttaaattt ttatcgaaa aagccattaa 2340
ggtggttatt attacatggg ggtggtggt ttattatg caaaatctc gtctattatg 2400
agatactggc attgatgagc ttgcctaaa gattagtatg aatttcagt aataccctc 2460
tgtttgcct atctctccct tctgtttat gtgatttgt tggggagaaa gctaaaaaaaa 2520
cctgaaacca gataagaaca ttctgtgt atagcttta tactcaaag tagcttctt 2580

tgatgccag cagcaaattg aatgctctct tattaagact tatataataa gtgcatgtag 2640
gaattgcaaa aaatatatta aaaatttati actgaattta aaaatatatt agaagttttg 2700
taatgggtgt gttttaatat ttacataat taaatatgta catattgatt agaaaaatat 2760
aacaagcaat ttttctgtct aacccaaaat gttatttga atcaaattgt tagtgattac 2820
actgaattg tgtacttagt gtgtatgtga tctccagtg ttatcccgga gatggattga 2880
tgtctccatt gtatttaaac caaaatgaac tgatactgt tggaatgtat gtgaactaat 2940
tgcaattata tttagagcata ttactgtagt gctgaatgag caggggcatt gcctgcaagg 3000
agaggagacc ctggaattg tttgcacag gtgtgtctgg tgaggagttt ttcagtgtgt 3060
gtctcttct tcccttctt cctctctcc ttattgtagt gccttatatg ataattgtat 3120
ggtaataga gtttacagt agcttgcctt aggatggacc agcaagcccc cgtggaccct 3180
aagttgtca ccgggattta tcagaacagg attagtagct gtattgtgta atgcattgtt 3240
ctcagtttc ctgccaacat tgaaaaataa aaacagcagc tttctcctt taccaccacc 3300
tctaccctt tccatttgg attctggct gagtctcac agaagcattt tcccatgtg 3360
gtctctcac tgtgcgtgc tacctgtct ctgtgagaat tcaggaagca ggtgagagga 3420
gtcaagccaa tattaaatat gcattcttt aaagtatgt caatcattt tagaatgaat 3480
ttttttcc tttccatg tggcagtcct tctgcacat agttgacatt cctagtaaaa 3540
tatttgctg ttgaaaaaaa catgtaaca gatgtgtta taccaaagag cctgtgtat 3600
tgcttaccat gtccccatac tatgaggaga agttttgtgg tgccgtgtg gacaaggaac 3660
tcacagaaag gttcttagc tggtaagaa tatagagaag gaaccaaagc ctgttgagtc 3720
attgaggctt ttgaggttc tttttaaca gctgtatag tottggggcc ctcaagctg 3780
tgaaattgtc ctgtacct cagctctgc atggatctgg gtcaagtaga aggtactggg 3840
gatggggaca ttctgccc taaaggattt ggggaaagaa gattaatcct aaaatacagg 3900
tgtgtccat ccgaattgaa aatgatata ttgagatata atttaggac tggttctgtg 3960
tagatagaga tgggtcaag gaggtcagg atggagatgg gagatttcat ggagcctggt 4020
cagccagctc tgtaccaggt tgaacaccga ggagctgtca aagtatttg agttcttca 4080
ttgaaggag taagggttc caagatgggg caggtagtcc gtacagcta ccaggaacat 4140
gtgtgtttt cttttttt taaaatcatt atattgagtt gtgttticag cactatattg 4200
gtcaagatag ccaagcagtt tgtataatt ctgtcactag tgcatacag tttctggtc 4260

aacatgtgtg atctttgtgt ctctttttg ccaagcacat tctgatttc ttgttgaac 4320
 acaggcttag ttctaaagg acaaatttt ttctcctgt ctttttctg taagggacaa 4380
 gatttgtgt tttgtaaga aatgagatgc aggaaagaaa accaaatccc attcctgcac 4440
 cccagtccaa taagcagata ccacttaaga taggagtcta aactccacag aaaaggataa 4500
 taccaagagc ttgtattgtt accttagtca ctgcctagc agtgtgtggc tttaaaaact 4560
 agagatttt cagtcttagt ctgcaaactg gcatttccga tttccagca taaaaatcca 4620
 cctgtgtctg ctgaatgtgt atgtatgtgc tctctgtggc tttagattct gtccctgggg 4680
 ttgccctgt tggccctgac aggaaggag gaagcctggt gaatttagtg agcagctggc 4740
 ctgggtcaca gtgacctgac ctcaaaccag cttaaggctt taagtctct ctcagaactt 4800
 ggcattcca actcttctt tccgggtga gagaagaagc ggagaagggt tcagtgtagc 4860
 cactctgggc tcatagggac acttggtcac tccagagttt ttaatagctc ccaggagggtg 4920
 atattatttt cagtgtcag ctgaaatacc aaccccagga ataagaactc cattcaaac 4980
 agttctggcc attctgagcc tgctttgtg atgtctatc cattgtctc cactagaggg 5040
 gctaagcttg actgccctta gccaggcaag cacagtaatg tgtgtttgt tcagcattat 5100
 tatgcaaaaa ttcactagt gagatggtt gttaggat aggaaatgaa atgcctctc 5160
 agtgacagga gtggcccgag cctgtctct attttgatt tttttttt taactgatag 5220
 atgggtcagc atgtctacat ggtgtttgt tgctaaactt tatataatgt gtggttcaa 5280
 ttcagcttga aaaataatct cactacatgt agcagtacat tatatgtaca ttatatgtaa 5340
 tgttagtatt tctgcttga atccttgata ttgcaatgga attcctactt tattaatatg 5400
 attgatatg ctagtattg ttgtcgattt aaacttttt tgcttctcc cttttttgg 5460
 ttgtgcgctt tctttacaa caagcctcta gaaacagata gttctgaga attactgagc 5520
 tatgtttga atgcagatgt acttagggag tatglaaaat aatcatttta acaaaagaaa 5580
 tagatattta aaatttaata ctaactatgg gaaaagggtc catttgttaa aacatagttt 5640
 atcttggat tcaatgtttg tcttgggtt tacaaagtag ctgtatttt cagtatttc 5700
 tacataatat ggtaaaatgt agagcaattg caatgcatca ataaatggg taaattttct 5760

g

5761

<210> 3

<211> 2129

<212> DNA
<213> human

<400> 3

ctgtattgag acaaaggaag ggatctgtca gaaagcaaca ctgttatct tgggcttggc 60
agcaaggaag aggacaggtg gtggagatcc tgcaatctga aaagcagact gaaaggtgac 120
aaagaagctg aagatgggtg gtggagagag gtataacatt ccagcccctc aatctagaaa 180
tgttagtaag aaccaacaac agcttaacag acagaagacc aaggaacaga attcccagat 240
gaagattgtt cataagaaaa aagaagagg acatggttat aactcatcag cagctgcctg 300
gcaggccatg caaaatgggg ggaagaacaa aaattttcca aataatcaaa gttgaattc 360
tagcttatca ggtcccaggt tactttttaa atcicaagct aatcagaact atgctggtgc 420
caaatttagt gagccgcat caccaaggtg tcttccaaa ccaccaagcc actgggtccc 480
tgtttcctt aatccttcag ataaggaaat aatgacatt caacttaaaa ccttactaa 540
agtacaggtg taaaataaga caaatgttta aatttagtta tgtcacggg tagttgtcaa 600
ttggtctgaa acaaatttc tagggaatct atttgtag aactaattaa tgtaaaaaa 660
atagaccatc tctgtgtg tgacactgtga tataatggta gtatcagtgc aactaaact 720
aatgattgta ctgtatatta agtgttctca actgagtaac tttaagtgg aaaccaagtt 780
tagattggg gagtggtaaa ggaatcagct tttctattg ttaggggaag acagtaattt 840
atcattcatg gaccagtaga ttgtgaaag ttggtgaatc ggattataag ctctagcta 900
acacaaggat tcagaattag gtaaacaatc gaaggtttag tatattagaa acacccaaac 960
cagtaatatg ctaacctgat gcactgctga aagaaaatgt gaatttticg taataattgc 1020
atttagtga attgtacagt gggtggaag ggcatttga gctcattaga atgagacata 1080
gtacacccca atggccctgt ttattaaatg tagtggatta agtgtctgtc acaaataca 1140
ccaaaacat ttttataga aacagtattt aatggtcact caatagcttt caaaatacat 1200
tttgtatta cagcactgca caagctattc taatagtgt ctggcctcat cattcctgca 1260
aagcttgctt tggggagttg gataatgtga aaattttaag tacctagggg agaaagagcc 1320
atgtaaatat ctgtaataaa ctgtagcat atgtaaagtt ttctggcct ttatcttaca 1380
aaaatggagt atttagtat gaatttctg aatgtaagac cgtggactgt ttttataat 1440
atggccta at ttaaaggtc caaaataact tgttttaaa gttgcccct gtgctaaagt 1500
gccagtgtat gtatgttata ctgatttgg ttgtaaacta tattcaaag taaaccctag 1560

tgtaataagt ttataacta aaaagggtta agctgctaaa actatttta agagatgtga 1620
 aatgcagtat gggactatct tttttcctc ctctaagccc aaagattaac tagagtcctt 1680
 ccaaccttat agattgttgg ctttcacaat cttataacct aggatacagg tagtttcgag 1740
 tatggtgccca gtgatgtttt gttttgttt ggtcaagggg taggtgcaac ccaatggacc 1800
 acttatgcaa aagatgtaaa ctctgcata atacattgat aacatgtttt gccaacitaa 1860
 aatgcttaaa cataagcgaa accagtagca agtatgtggg tcagcttaaa aattttgatt 1920
 gtaatgccc tattttctaa ttggcacct cttgatgcct aagcaggtaa gcagatgcct 1980
 aagctgtatt tctccaaata aatcaagatg aagtactgcc caagttaaatt atgatagcc 2040
 taaagacaag ttatgtagt acttaatgta catgatatga agcataaaat taaataaaat 2100
 tttcccat tgaaaaaaaa aaaaaaaaaa 2129

<210> 4
 <211> 3950
 <212> DNA
 <213> human

<400> 4
 cgagaactag tttgttccg tgccctctgg actggaacct ttggagaga acccccggca 60
 ggaccaaccc cgcacccgcc agcaccgcgg caatgtccag caatagtttt cttacaatg 120
 agcagtccgg aggaggggag gcgacggagc tgggtcagga ggcgacctca accattccc 180
 cctcgggggc cttcggcctc tttagcagcg atttgaagaa gaatgaagat ctaaagcaaa 240
 tgttagagag caacaaagat tctgctaaac tggatgctat gaagcggatt gttgggatga 300
 ttgcaaaagg gaaaaatgca tctgaactgt ttctgtctgt tgtgaagaat gtggccagta 360
 aaaatattga gatcaagaag ttggtatatg ttacctggt tcgatatgct gaagaacagc 420
 aggatcttgc actcctgtcc ataagcactt tcagcagagc tctgaaggac ccaaaccaac 480
 taattcgtgc aagcgcttg agagtctgt caagtattag agtgccaatt attgtaccta 540
 tcatgatgct tgctattaag gaagcttctg ctgacttacc accatatgtt aggaagaatg 600
 cagcccatgc aatacaaaaa ttatacagcc ttgatccaga gcagaaggaa atgttaattg 660
 aagtaattga aaaacttctg aaagataaaa gcacattggg agctggcagt gttgtgatgg 720
 cttttgaaga agtatccccg gacagaatag atctgattca taaaattac cgcaagctat 780
 gtaacttact agtggatgtt gaagagtggg ggcaggttgt cataatccac atgctaactc 840

gatatgctcg gacacagttt gtcagccctt ggaaagaggg tgatgaatta gaagacaatg 900
gaaagaattt ctacgaatct gatgatgatc agaaggaaaa gactgacaaa aagaagaagc 960
cgtatactat ggatccagat catagactct taattagaaa taaaagcct ttgcttcaga 1020
gcaggaatgc tgcggtggtt atggcagttg ctacagctga ttggcacata tcacaaaaat 1080
ctgaagctgg cataatttct aaatcactag tgcgtttact tctagcaat agggaggtgc 1140
agtatattgt cctacaaaat atagcaacta tgtcaattca aagaaagggg atgttgaac 1200
cttatctgaa gagtttctat gtaggtcaa ctgatccaac tatgatcaag aactgaagc 1260
ttgaaatttt gacaaacttg gcaaatgaag ccaacatatc aactcttct cgagaatttc 1320
agacctatgt gaaaagccag gataaacaat ttgcagcagc cactattcag actataggca 1380
gatgtgcaac caacatcttg gaagtcactg acacgtgcct caatggcttg gtctgtctgc 1440
tgtccaacag ggatgaaata gttgttgctg aaagtgtggt tgtataaag aaattactgc 1500
aaatgcaacc tgcacaacat ggtgaaatta ttaacatat ggccaaactc ctggacagta 1560
tcaactgttc tgttgctaga gcaagtattc ttggctaata tggagaaaac tgtgaacgag 1620
ttcctaaaat tgcctctgat gtttgagga agatggctaa aagcttact agtgaagatg 1680
atctggtaaa actgcagata ttaaatctgg gagcaaaatt gtatttaacc aactccaaac 1740
agacaaaatt gcttaccag tacatattaa atctcgcaa gtatgatcaa aactacgaca 1800
tcagagaccg tacaagattt attaggcagc ttattgttc gaatgaaaag agtggagctt 1860
taagtaaata tgccaaaaaa atattcctag cacaaaagcc tgcaccactg ctgagtcctc 1920
cttttaaaga tagagatcat ttccagcttg gcaccttctc tcatactctc aacattaaag 1980
ctactgggta cctggaatta tctaattggc cagaggtggc gcccgacca tcagttcgaa 2040
atgtagaagt aatagagttg gcaaaagaat ggaccccgagc aggaaaagca aagcaagaga 2100
attctgctaa gaagttttat tctgaatctg aggaagagga ggactcttct gatagtagca 2160
gtgacagtga gagtgaatct ggaagtgaaa gtggagaaca aggcgaaagt ggggaggaag 2220
gagacagcaa tgaggacagc agtgaggact cctccagtga gcaggacagt gagagtggac 2280
gggagtcagg cctagaaaac aaaagaacag ccaagaggaa ctcaaaagcc aaaggaaaaa 2340
gtgattctga agatggggag aaggaaaatg aaaaatctaa aacttcagat tctcaaag 2400
acgaatctag ttcaatagaa gacagttctt ccgattctga atcagagtca gaacctgaaa 2460
gtgaatctga atccagaaga gtcactaagg agaaagaaaa gaaaacaaag caagatagaa 2520

ctctcttac caaagatgtt tcactctag atctggatga tttaaccca gtaaccactc 2580
 cagttgcact tcccacacca gctctttctc caagtttgat ggctgatctt gaaggtttac 2640
 actgtcaac ttctcttca gtcacagtg tcagtactcc tgcatttga ccaacgaaaa 2700
 ctacagtgt gcttcacga atgagtggaa aaggactagc tgcccattat ttcttccaa 2760
 gacagcctg catttttgg gataagatgg tctctataca aataacactg aataacacta 2820
 ctgatcgaaa gatagaaaat atccacatag gggaaaaaaa acttctata ggcataaaaa 2880
 tgcattgttt taatccaata gactctctg agcctgaggg atccattaca gttcaatgg 2940
 gatttgactt ttgtattct actcagactg ccagtttcca gttgtgtacc aaggatgatt 3000
 gctcaatgt taatattcag ccacctgtg gagaactgct ttacctgtg gccatgtcag 3060

 agaaagatt taagaaagag caaggagtgc taacaggaat gaatgaaact tctgctgtaa 3120
 tcattgctgc accacagaat ttactccct ctgtgatctt tcagaagggt gtaaattgtag 3180
 ccaatgtagg tgcagtccct tcggccagg ataataaca caggtttgca gctaaaactg 3240
 tgcacagtgg gtcattgatg ctagtacag tggaaactgaa ggaaggctct acagcccagc 3300
 ttatcataaa cactgagaaa actgtgattg gctctgttct gctgcgggaa ctgaagcctg 3360
 tctgtctca ggggtaacct gcttacctt ggactttaga atctggcaca caacaaaagt 3420
 gcctggcacc cactactgct gccttcatt tataataata gcccttccat ctggcagtgg 3480
 gggtagaata cactcttgac attctgtct cctgctttag aatgctagtg tgtatctatc 3540
 atgtatgcaa tactttcccc cttttgctt tgctaaccga agagcatata tttaactgtc 3600
 agttgtctca actcttgaat ccatgtggcg ttttctgt cctgctgctt ctttggcct 3660
 cctcgtttc ctctctttt tcgacaatgg tagacatgaa tgagatattt aaagttcatt 3720
 ggaaatcttc ttccctacag cagtaagcaa aaattagcaa agagatagtc taaatggcct 3780
 ctacagctgg tatgtgaaaa tgagatcaca tactttttaa atccaaatac aaaagcatag 3840
 tctctgcaag attttgtct ttgaattct tgatattga attgattatt gataactgtc 3900
 atcatgaaat tatctctcaa taataagata aataaactag catatgaatc 3950

<210> 5
 <211> 5191
 <212> DNA
 <213> human

<400> 5

gagaaagaaa aacagctcga gacctcatgc aaagagaaaa ctgagtatct acagaaaatg 60
gttcagagga atgaaagata taaacaagat gtggagaggt tctatgaacg gaagcgacat 120
ttagatttaa ttgagatgct tgaagcaaaa aggccatggg tggaatatga aaatgttcgt 180
caggaatatg aagaagtaaa actagttcgt gaccgagtga aggaagaggt cagaaaactt 240
aaagaagggc agattcctat aacatgtcga attgaagaaa tggaaaacga gcgtcacaat 300
ttggaggctc gaatcaaaga aaaggcaaca gatattaagg aggcattcga aaaatgcaaa 360
cagaagcaag atgttataga aaggaaagat aaacatattg aggaacttca gcaggcttta 420
atagtaaagc aaaatgaaga gcttgaccga cagaggagaa taggtaatac ccgcaaaatg 480
atagaggatt tgcaaaatga actaaagacc acggaaaact gcgagaatct tcagccccag 540
attgatgcca ttacaaatga tctgagacgg attcaggatg aaaaggcatt atgtgaaggc 600
gaaataattg ataagcgaag agagagggaa actctagaga aggagaaaaa gagtgtggac 660
gatcatattg tacgtttga caatcttatg aatcagaagg aagataagct aagacagaga 720
ttccgtgaca cgtatgatgc tgtttatgg ctaagaaata acagagacaa atttaacaa 780
agagtctgtg agcccataat gctcacgac aatatgaaag ataataaaaa tgccaaatat 840
attgaaaatc atattccatc aatgactta agagccttg tatttgaaag tcaagaagat 900
atggagggtt tctcaaaga ggttcgtgac aataaaaaat taagagtaaa tgctgttatt 960
gtcccaaga gtcatatgc agacaaagca ccttcaagat ctttgaatga acttaacaa 1020
tacggatttt tctctattt gagagaatta ttgatgcac ctgacctgt aatgagttac 1080
cttgctgtc agtatcatat tcatgaagt cctgtaggaa ctgaaaagac cagagaaaga 1140
attgaacggg taatacaaga aacccgatta aaacagattt atacagcaga agaaaagtat 1200
gtggtgaaaa cttctttta ttcaaacaaa gttatttcta gtaacacatc tctaaaagta 1260
gcgcagtttc tcaactgtac tgtggaccta gagcagagaa gacacttaga agaacagcta 1320
aaggaaattc atagaaaatt gcaagcagt gattcagggt tgattgcctt acgtgaaaca 1380
agcaaacatc tggagcacia agacaatgaa cttagacaaa agaagaagga gcttcttgag 1440
agaaaaacca agaaaagaca actggaacaa aaaatcagtt ccaaactagg aagtttaaag 1500
ctgatggaac aggatacttg caatcttgaa gaggaagagc gaaaagcaag taccaaaatc 1560
aaagaaataa atgttcaaaa agcgaaactt gttaccgaat taacaaacct aataaagatt 1620
tgacttctt tgcatataca aaaagtagat ttaattctcc aaaatactac agtgatctct 1680

gagaagaaca aattagaatc agattatatg gccgcatctt cacaactccg tcttacagag 1740
caacatttca tgaattgga tgaaaataga cagagattat tgcagaaatg caaggaactt 1800
atgaaaagag ctaggcaagt atgtaacctg ggtgcagagc agactcttc tcaagaatac 1860
cagacacaag taccaccat tccaaatgga cacaactcct cactcccat gggttccaa 1920
gaccttcaa acacattgga tgaaattgat gcttlatata ctgaagaaag atcaagagct 1980
tctgcttca cgggactgaa tctacaatt gttcaggaat atacaaaaag agaagaagaa 2040
atagaacagt taactgagga actaaaggga aagaaagtg aactagatca atacagggaa 2100
aacattcac aggtaaaaga aaggtggctt aatccttaa aagagctggt agaaaaaatt 2160
aatgaaaaat tcagcaattt tttagtcc atgcagtgtg ctggtgaagt tgatccat 2220
acagaaaatg aggaagatta tgataatat ggaattcga ttagagtaa attcgaagt 2280
agtactcaac tgcataat aactcctcat catcaaagt gaggtgaaag aagtttct 2340
accaattat actgatggc actcaggag ctaaatagat gtccattcag agtagtgat 2400
gaaatcaatc aggaatgga cccaatcaat gaacggagag tgttgaaat ggtgtaaat 2460
actgcctgta aagaaaatac atcacaatc ttttcataa caccaaagct cctgcaaaat 2520
cttcttatt ctgaaaagat gacagtttg ttgtctaca atggccctca tatgtggaa 2580
ccaaacacat ggaatttaa ggcttccaa aggcggcggc gccgtattac attcactca 2640
ccttctaataaaaagtaaag agagggaact tgggaattt tttgttaa tctgttat 2700
aagtatggct caactgaata aaaggagatt cactaaaacg aaaagcagt attttgaa 2760
acctgcttt aaatacaaat aggtgataa tggaactat aatgacctt ccaaaatagc 2820
agctgtagt aaaagtaa tcttctcag tctggtga actgagtc ttggactct 2880
gacctgagt cattcagtc tcatgttaa atgtactaa tattacaaat caaaggtaca 2940
gtggaagaag ggtaatcac aagaagttac ttatatgga gccctgagct ttaattgag 3000
agtaacttta attacttta gagcctaaag atgacttag agcctaagtc ctagttctc 3060
ccaatgtat atttaattt aaaaaattga tatgaaatg tctaattat agtaataat 3120
tatgacagat ctagtcattt ctctatta aaaaagatta ccttatctc agtaggaaat 3180
ggaatttat gggccttaa aagaaagtt tatgaaact gatgtataa tttattgt 3240
attcaaggg gaaaaagca ctgggttca aaatggtag cagaactgct ttgaaatgct 3300
gcaagggtgc cactagatga tgcaaaatc aacaaaaga ttgactgaga ataaaattag 3360

gtgacaaggg ttttaaaga ataaccttt aaagtgtggg ggcaggggtt gctttttt 3420
atattatta aagtcaatta tattttacat cttacatttc taaaagcatt ttataattat 3480
tttagtaag atttttctta aaatttcata tactggtttc tacaatttat atttgaaatt 3540
tctcagtgtt atgtaaagag tgatggaaaa gcattgattt ctttaaaacc gtaatgttt 3600
tagaacttaa gcctataggg cctttcttac aatgttgatg taccattat cttagaaaaat 3660
ctagtttaa cigtttctt tcaccgcaa agaattaaat gggaaaatca ttgtttatc 3720
tctaagttat actaattagt agaaccaaac aaattatctt ctttaaaaa ataatctta 3780
taggaaaata gacagtccaa agtcatgtct ttgaacagtg gatggatct gtgccagtaa 3840
tgacaaaatt atttttga ctgcttgcc tgaataaatt gaagaatgc ttcagtttg 3900
ggtttgtat attcttaagt agccattgaa atttatctt ttaactaggt caaaaaataa 3960
tgagccataa gtttatgtcc tctacttag acattttctc tttaaaaagg tattttcttc 4020
ttataaaca ttttaaaga gcctccctt cttaaactaa ctccagtga tgaagtgtga 4080
aaatatttta aaatgacatt ttactaata tgagcaagtc atgtaaacat tgaagaactt 4140
ggtaacatat tagtaaatgg atattacaa atgtttcat cgtaattac ttgcgttcc 4200
acaaaaatat cttfactaaa atgtgcttgg ttagtttgt ttattgtcta aattagtacc 4260
agtcatctta ttctgcaa atgagiatca atgtgaaaa gacacgtgaa gattaagcat 4320
gttgaaaaat aaaatggtca attacatttc aattacata ggccaacaac tgttcatac 4380
ttgtttgta aacatttaat ttctctactg gacaaaatta atattggct ttacattgaa 4440
tttgagctg tgaagaataa attatgtatc atttagcat attaaacagt agtaagtcta 4500
gcacatagtc tcagccactt aaaacaaaag tttttgtt tgtttgttg ttgttttt 4560
tgagatggag tctactctg ttgccaggc tggagtgcag tggcgtgatc tcggcttact 4620
gcaacctccg cctccgggt tcaagcgatt ctctgcctc agcctccaa gtaactggga 4680
caacaggcgc gtcccaccac acccagctaa tttttatc tttagtaga gatggggtt 4740
cagcatattg gccaggctgg tctgaactc ctgacctgt galccaccg cctcggcctc 4800
ccaaagtgtt gggattatg gcgtgagccc ctgcaccgg ccaaagtgtt attttaatt 4860
acataaaaat cgtaaaaact tctagtaaaa acttgatttg gtgaatacag ttatattta 4920
aaaccttaag gtgacaagca ttctatgc ctaaatcttc attggttgc ctggaaagag 4980
tctctgttaa aagatttcc atattcaaag taaaaggaaa gatttctgc ttctaattg 5040

ttctttggac acatgcctat ttctttgag gtataaacct ttagatgtga aaaatgtaat 5100

ttcattctgc tattgtgtgt gcttgtgtgt gtgtaattga aaaaactggg aaatcctgct 5160

ttgttggtaa taaatcaata ttttatatt c 5191

<210> 6

<211> 4755

<212> DNA

<213> human

<400> 6

aagagatctt ccaggctctc agagccctgg gagggcgatt tccaggaaga ccacaatgcc 60

aacctctgga ggaggctgga gagagaaggc ctaggccaga gcctgtcagg caactttggc 120

aagaccaagt cagccttctc atctctccag aacattcctg agagtctgag aagacacagc 180

agcctggagc taggccgggg aaccaggag ggttaccctg ggggcaggcc cacctgtgca 240

gtcaacacca aggcagaaga ccctgggagg aaagccgctc ctgacctcgg gagccatctg 300

gaccggcagg ttctctacc cgggcccag gggaggaccg gtgcctcggc ttcttcaac 360

agcacagacc caagtccga agagccgctt gcccctcgc acccgcacac atccagtctg 420

ggccggaggg ggccggccc aggcagcgc tcggctcttc agggcttca gtacgggaag 480

ccccactgct cgggtctgga gaaggtctcc aaattcgagc agcgagagca agggagccag 540

agaccgagtg tgggcggctc tggtttggc cataactata ggccccacag gaccgttca 600

acttccagta ctctgggaa tgacttcgag gagacaaaag cacacattcg ttctctgag 660

tcagctgaac ccctaggcaa cggggagcag cactcaaaa acggggagct gaagttgaa 720

gaggcttccc ggcagccctg cggtcagcag ctgagcggag gagcgtcgga cagcggccgt 780

ggccccaga ggccggacgc tcggctctc cgtagccaga gcaccttcca gctctcagc 840

gagccagaga gggagcccga gtggcgggac agggccggct cggccgaatc gcccctgctg 900

gatgccccct tcagccgcgc ctaccggaac agcatcaagg acgcacagtc ccgtgtcttg 960

ggggccacct ccttcgacg tcgagacctg gagctggggg cgcccgctgc gtcgaggtcc 1020

tggcggccac ggccttctc ggcccacgtg gggctgcgga gcccagaggc gtcggcctcc 1080

gcctccccgc acacgccccg ggagcggcac agcgtgacct ctgctgaggg cgacctggcc 1140

aggcccgctg ccctgccgc ccggagaggt gctcgccggc gcctgactcc cgagcagaag 1200

aagcgtcctt actcggagcc cgagaagatg aacgaggtgg ggaicgtgga ggaggccgaa 1260

ccggcacccc tgggcccgcg gagaaatggg atgcgtttcc cggagagcag cgtggccgac 1320
cggcgcgcgtc tcttcgagcg cgatggcaag gcctgctcca cgctcagcct gtcggggccc 1380
gagctgaagc agttccagca gagcgccctg gcggactaca tccagcgcaa gaccggcaag 1440
cggcctacct ccgcccgcgg ctgcagcctc caggagcccc ggccactgcg tgagcgcgcc 1500
cagagtgcct acctccagcc cggccccgcg gcgctcgaag gctccggcct cgcctcggcc 1560
tccagcttga gctcactgcg ggagcccagc ctgcagcccc gcagggaggc cacgctcctg 1620
ccggccacag ttgcagaaac ccagcaggct ccccgagatc gcagcagctc cttcgcgggt 1680
ggccgcgcgc tcggggaacg gcgacgcggg gacctgtta gcggagcaaa cggtggaaca 1740
aggggcaccc agagagggga tgagaccccc agggagccat cctcctgggg ggccagggcc 1800
gggaagtcca tgcggccga ggacctgctg gaacgctcg acgtccttgc gggccctgtc 1860
catgtgaggt ccaggtcatc tcccgccacc gcagacaagc gccaggatgt gctttgggg 1920
caagacagtg gctttggtct tgtgaaggat ccatgttatt tggctggtcc tggatctagg 1980
tcactcagtt gttcagaaag aggccaagaa gagatgctgc tgctcttcca ccatctcacc 2040
cctcgttggg gtggttcagg ctgcaaagcc attggtgatt cctccgttcc tagtgaatgt 2100
cctggaaccc tggaccatca gaggcaagcc agtaggacac cctgccccag gccaccactg 2160
gcaggaacgc aagggtgtgt cacagacacc aggggtgcac cctgacccc aattggcacc 2220
cctctgcctt cagccattcc ctctggctac tgcacacagg acggtcagac agggcgacag 2280
cctctccgc cctacacccc tgccatgatg cacagaagca atggtcacac cctgaccag 2340
cctccgggtc caagaggctg tgagggcgat ggcccagagc atggggtaga agagggaacg 2400
aggaagaggg tctcgtgcc tcagtggcca cctccttctc gagcaaagtg ggcccacgca 2460
gccagagagg acagccttcc tgaggaatcc tcagcccctg atttgcaaa cctgaagcac 2520
tatcaaaaac agcagagtct tccaagtta tgcagcactt ctgaccaga cacacctt 2580
ggggccccga gcactccagg gaggatctcc ctccgaatat ctgagtctgt cctgcgggac 2640
tccccgccac ctcatgagga ttatgaagac gaagtgttg tgagggatcc gcacccaag 2700
gccacgtcca gcccacatt tgaacctt ccccccccc cacctctcc accgagttag 2760
gaaaccccgg tgtatagcat ggatgacttc cctccacctc ctccccacac tgtatgtgag 2820
gcgcagctgg acagtgagga tcccaggggg ccacgcccc gcttaacaa actttctaaa 2880
gtgacaattg caagggaaag gcacatgcct ggtgcagccc atgtgtagg tagtcagaca 2940

ctggcttcca gactccaaac ttctatcaag gggtcagagg ctgagtcac accaccctcc 3000
ttcatgagcg ttacgcccc acttgctggg tctcttggtg ggcagccagc acccatccag 3060
actcaaagcc tcagccatga tccagtcagt ggaactcagg gtttagaaaa gaaagtcagt 3120
cctgatcctc agaagagttc agaagacatc agaacagagg ctttgccaa ggaaattgtc 3180
caccaagaca aatctctagc agacatttg gatccagact ccaggctgaa gacaacaatg 3240
gacctgatgg aaggtttgtt tccccgagat gtgaactgc tgaaggaaaa cagtgtaaag 3300
aggaaggcca tacagagaac tgtcagctct tcaggatgtg aaggcaagag gaatgaagac 3360
aaggaagcag tgagcatgtt ggtaactgc cctgcctact acagtgtgc tgcctccaag 3420
gctgagctac tgaacaaaat caaagagatg ccagcagaag tgaatgagga agaggaacag 3480
gcagatgtca atgaaaagaa ggctgagctc attggaagtc taccacaaa gctggagacc 3540
ctccaggagg cgaaggggag cctgctcacg gacatcaagc tcaacaacgc cctgggagaa 3600
gaggtggagg ctctgatcag cgagctctgc aagcccaatg agtttgaaa gtataggatg 3660
ttcatagggg atttgacaa ggtggtaac ctgctgctct cctctcggg gcgtctagcc 3720
cgtgtgaga atgtccttag cggccttggt gaagatgcca gtaatgaaga aaggagctct 3780
ctttacgaga aaaggaagat cctggctggt cagcatgagg atgccggga gctgaaggag 3840
aacctggatc gcagggagcg agtagtctg ggcattctgg ccaattacct ttcagaggag 3900
cagctccagg actaccagca ctctgtgaaa atgaagtcca cgctcctcat tgagcaacgg 3960
aagctggatg acaagatcaa gctgggccag gacaggtca agtgtctgct ggagagcctg 4020
ccctcagatt tcattccaa ggctggggcc ctggctctgc ccccaaacct cagagtgag 4080
cccattcctg ctgggggctg tacttcagt ggtatttcc caacattaac ctctccactt 4140
taacctcttc taaaataccc aacaaaaaga tcaactgttc tctaacact attaatctg 4200
aaaaatgtt cagtacaaac cactgttga actatctggg ttattggtg ttgtcctga 4260
tgaaaggaaa aaaattctct ccaggaggaa gccttttcc ttctgccct tctgattga 4320
tctctgaga gctcgaatgc tgctggacac gtacccttt ctattattac ttgtagtag 4380
aaagaaagtt aatgaaactg agaactgatt ggagggtgtt tgatcattta gttttaaca 4440
ggctgaggca acatggatca gtgtgtgcc cctcaggaa tgtatccaca gtgccttcc 4500
ttgctggtg gcagtgtatc ctgatggcag ggtacaagta ccattaatga aggtctgca 4560
acataaagcc taaaaagac acacactaag aaaactgtaa aacctgaac attgtattt 4620

atattttta aaatggaaaa gatcactatg ttgtgtgc taaccactta ttgattctg 4680
tttgtggtg gacatagatg attacgttg agcttgtat ttgtgaaaa ccttaatgaa 4740
atgaattcca aagat 4755

<210> 7
<211> 2045
<212> DNA
<213> human

<400> 7
gaaactgac cccggctcat cctgtctctg gctgtggccc ggcaaagcac tgaaaacccc 60
tctggtctca gagacagtag gggcagtgcc actttctaca acctgccaac ccacacactg 120
gagtaattct gaaaaaaatt attcctaaac tctctaagtg tggacggaga atgagcaagc 180
cccagaagta tttaacaacc agagtgggta atgaggaggg ggcttactgg aatcgtcata 240
tctctgaata ttgaaaacaa caactaaaaa agtggacctt ctcagaaaaa aagggcagca 300
aatgaccaag ggcgcccctt ctggccgtgc ttggcttgag taactgtctc tcttcccca 360
ccccatcac agggcttca gtttggcaaa ggaaaagcag ataaaaacag aacattccat 420
atgtttctt ctccatcggc caaaaacatt ttgacacaat gtttgtgaaa cacctttgga 480
gaggtgcact tctgaatgct gcctctgccg taaatcctgg ggcaagggat cagcctcttc 540
ccaggaacca tcgcttcta taaaccgtga actcaagcag gcatttttt ttcttaccg 600
aaaggctgct atttgcaag ggcacataat gggctctgtg ctcttattgg ctccaaatg 660
tgcatggcaa agagagagat gtgggcctag agcagatata ttcagcaagg tgacagcttc 720
ccataacaat ttaacactt ctatcttat gtgagaataa aatatttaag ggtgaacct 780
tatttgcca aatgtatctt ttctgctttt gaattgggca gaagatttta gcaactatat 840
tctacaaatg ttactataa cacacacaca cacatctgaa atatatgccg aaaattgacg 900
tcttgacct caggagagagc acctgtccag gtctgcctaa aggaaatggc tccagtgggt 960
ctaaacaacc acatcctatc catgatagg tctagtcata acactttaga gagaatgtca 1020
gagcaggagg gaggcaagcc gcctctctc ggccatcaac tgcagatgat gaaagagcgg 1080
gattcaactt tgtttctt ttctgtggcc ccagtgaac ctctgccct cctgcacgt 1140
ctgtgtcttc atttclaaaa tgggggtgat gcttcatat tgacctcacc ccatactacc 1200
tcacagatgt gttgtgagga ttaataaaat tatgtctatg gtatttcag ttctggaga 1260

aaaatactta tagacagttt aactattaca tagatatata agtgatctca gtttctgtt 1320
 tgctgtgata ctaatgtgtt gttttaactt attccataaa atgacagttg tgcctagcc 1380
 acatcagaca gctatctaag ctctggacta ccccttgtg cagctgaatc actgcagggt 1440
 cgaccatgcc tgggtccaca gccatggtt ccatttctag atgaaaggat ggcctaggac 1500
 ataggctca aagactcttg gatcagaatc aggagattag ggaaaacagg atggatacct 1560
 gagcactaac agcagtagac gtagacctct gtccttacc atctgaggtc ttctggattc 1620
 ttgtgggtt taatttcat ttgatgtcat ctgttgccc tcatcttgc ttgcaagtgt 1680
 gcatggttca atccctcaca tccaggaaat gaatttga attgggccag atgctaatt 1740
 gcacgtgat tcacctctt tgccttaag cctttttt cttttttt tttttggca 1800
 aatgaatga ccatttcaac ttgatttta atagtctag ttgatattg taataatgt 1860
 aaccaagaga tcaatgccag attttctct tgggtaagt tagctgaagt cattaaaga 1920
 tggaaagggt ggaaaattct ttgatattg atgtcattgt atccacatt gtgtaagac 1980
 atattgata ccaattataa ttatatcaat taaagttgat aaaagctca aaaaaaaaaa 2040
 aaaaa 2045

<210> 8
 <211> 2096
 <212> DNA
 <213> human

<400> 8
 atggagaacg agcctgtagc ccttgaggaa actcagaaga cagatcctgc tatggaacca 60
 cgggtcaaag tgggtgattg ggacaaggac ctggtggact ggcgaaagcc tctcctgtgg 120
 caggtgggcc acttgggaga gaaglacgat gagtgggtc accagccggt gaccaggccc 180
 atccgcctct tccactcaga cctcattgag ggccctctca agactgtctg gtacagtgtc 240
 cccatcatct ggggtcccct ggtgctgtat ctacgtggt cctactaccg aaccttgcc 300
 cagggcaacg tccgactctt cagtcattt acaacagagt acacgggtgc agtgcceaag 360
 tccatgtcc ccgggctctt catgctgggg acattcctct ggagcctcat cgagtacctc 420
 atccaccgtt tctgttcca catgaagccc cccagcgaca gctattacct catcatgctg 480
 cactctgca tgcacggcca gcaccacaag gcaccctcg acggctccc cctggtcttc 540
 ccccctgtgc cagcctccct ggtgatcggc gtcttctact tgtcatgca gctcatcctg 600
 cccgaggcag tagggggcac tgtgttgcg gggggcctcc tgggtacgt cctctatgac 660

atgacccatt actacctgca cttggctcg ccgcacaagg gctcctacct gtacagcctg 720
 aagggcccacc acgtcaagca ccactttgca catcagaagt caggatttgg taccagcact 780
 aaattgtggg attactgttt ccacaccctc actccagaga aacccacact gaagacgcag 840
 tgacaactcc caccacctcc gtcctgccct cagcccgccc ctggccctt cccgaccccc 900
 acccgccatt cagaccccat taagaagggt ggcttgcca ggcaggatgg gctgtgtccg 960
 gccctgcagc ctagtgaag gtgctgagg ggccctgagg caggaccgcc ctctgaccc 1020
 ctggtaggag ggtcacatcc acttggtgca ggtggccctt ggtgaccac ttctcttg 1080
 agcgctccctg cctagagctc agccacagg actgctcag gccgtggcca caggtagcag 1140
 ccgcaagggg aaatgaagaa aactgagccc tcgtggccac ctgtgcacc cttgtgcctt 1200
 agcctcatgg gctgcctagg agctgcctgc acggcacagc tcgcttcac agtcagaagt 1260
 gggctgtgg gatctgtgt cctgtcctc cctgtgtcc ctctgggga ggcttgggtg 1320
 gctctgaggt ggacaaagag ctctgcaag aagagacagc gtatgcctc ccacagtcca 1380
 cccagaccc tggggcagcc cctctggccc tgccagctgc ctgcgtcgtt gggcccaggg 1440
 ttgctggcag gactccagc tgcctgctt aggaacctggc agctttctt gccgtccctc 1500
 cctgcctcc agaatcacag cctctctcc caaggaggc tgaggaggct tctccaccag 1560
 tggcagcccc acccgtccc tggcattct tggcctccac ccgctcagg cccctactcg 1620
 ggcgtccca gaaggagcca cctcagtg cctcacctcc cctgcctcc cagcctccgc 1680
 agatgaggtt cctgcccctt cctcctcgt accaaaacc tactgtctcc caggacggtc 1740
 ttattataa accagataca tgtcttagt ctggtcccag accaaggagc tggcagacg 1800
 gcccttcta atctacatg ttgagcttat gtaaaaaatg ttgttcctc ctgttttg 1860
 ttctttctt accacaaac cattactact tgaaactaa aaaactgcc aagtgtaaag 1920
 gctaaagaga agcagttga cggacctgt gattgtact gttgtgcg gagctattta 1980
 aagatttgg aataaatata caaaactacg gttgtgaaat aaaaactaa attgtatatt 2040
 ttgaaaaata aaacactgaa aagaaccaa caaaaaaaaa aaaaaaaaaa aaaaaa 2096

<210> 9
 <211> 5640
 <212> DNA
 <213> human
 <400> 9

ggaacgcag aaaacagaga gaggcattct gagtcactg actggatgaa gactgtcca 60
agttacaacc aaacaaatag ctccatggac ttagaaatt ata'tgatgag agatgagact 120
ctggaaccac tgcccaaaaa ctgggaaatg gcctacactg acacagggat gatctactc 180
attgaccaca ataccaagac aaccacctgg ttggatcctc gtctttgtaa gaaagccaaa 240
gcccctgaag actgtgaaga tggagagctt ccttatggct gggagaaaat agaggaccct 300
cagtatggga catactatgt tgatcacctt aaccagaaaa ccagtttga aaatccagt 360
gaggaagcca aaaggaaaaa gcagttagga caggttgaaa ttgggtcttc aaaaccagat 420
atggaaaaat cacactcac aagagatcca tccagctta aagggtcctt tgttcgagca 480
tactgaaaa aaagcacaat gggatttggg ttactatta ttggtggaga tagacctgat 540
gagttcctac aagtgaaaaa tgtctgaaa gatgtcccg cagctcagga tgggaaaatt 600
gcaccaggcg atgtattgt agacatcaat ggcaactgtg tctcgggtca cactcatgca 660
gatgtgtcc agatgttca attggtacct gtcaatcagt atgtaaacct cactttatgt 720
cgtggtatc cacttctga tgacagtga gatcctgtg ttgacattgt tctgtctacc 780
cctgtcatca atggacagtc attaaccaag ggagagactt gcataaatcc tcaggattt 840
aagccaggag caatgttct ggagcagaat ggaaatcgg gacacactt gactggtgat 900
ggctcaatg gaccatcaga tgcaagtga cagagagtat ccatggcatc gtcaggcagc 960
tccagcctg aactagtga tatcccttg attaagggcc ctaaagggtt tgggttgca 1020
attgtgaca gccctactgg acagaagggt aaaatgatac tggatagtca gtggtgtcaa 1080
ggccttcaga aaggagatat aattaaggaa atataccatc aaaatgtgca gaatttaaca 1140
catctcaag tggtagaggt gctaaagcag ttccagtag gtgtgatgt accattgctt 1200
atcttaagag gaggtcctc ttaccaacc aaaactgcc aatgaaaac agataaaaag 1260
gaaaatgcag gaagtttga gccataaat gagcctatc ctacgctat gcctttcca 1320
ccgagcatta tcaggtcagg atcccaaaa ttgatcctt ctgaggicta cctgaaatct 1380
aagacttat atgaagataa accaccaaac accaaagatt tggatgttt tctcgaaaa 1440
caagagtcag ggttggctt cagggtgcta ggaggagatg gacctgacca gtctatatat 1500
attggggcta ttatccctt gggagcagct gagaaagatg gtcggctccg cgcagctgat 1560
gaactaatgt gcattgatg aattcctgt aaagggaat cacacaaaca agtcttgac 1620
ctcatgacaa ctgtgctcg aaatggccat gtgtactaa ctgtcagacg gaagatctc 1680

tatggagaaa aacaacccga ggacgacagc tctcaggcct tcatttcaac acagaatgga 1740
tctccccgcc tgaaccgggc agagggtcca gccaggcctg caccacagga gccctatgat 1800
gttgtcttgc aacgaaaaga aaatgaagga ttggctttg tcatcctcac ctcaaaaaac 1860
aaaccacctc caggagtlat tcttcataaa attggccgag tcatagaagg aagtccggct 1920
gaccgctgtg gaaaactgaa agttggagat catatctctg cagtgaatgg gcagtcatt 1980
gttgaactgt ctcagtataa cattgttcag ctgatcaaag atgctggtgt caccgtcaca 2040
ctaacggta ttgctgaaga agagcatcat ggtccacat caggaacaaa ctgagccagg 2100
caaagcccag cctgcagca caggcccatg ggacagtcac aggccaaacca catacctggg 2160
gacagaagt cctagaagg tgaaattgga aaagatgtct ccactctta cagacattct 2220
tggtcagacc acaagcacct tgcacagcct gacaccgcag taatttcagt ttaggcagt 2280
cggcacaatc agaacctgg ttgtatcca gtagagctgg agagaggccc ccggggcttt 2340
ggattcagcc tccgaggggg gaaggagtac aacatggggc tgttcacct tegtcttgc 2400
gaagatggc ctgccatcaa agatggcaga attcatgtg gtgaccagat tgtgaaatc 2460
aatggggaac ctacacaagg aatcacacat actcgagcaa ttgagctcat tcaggctggt 2520
ggaaataaag ttctctctt ttgaggcca ggaactggct tgatacctga ccatggttg 2580
gtccttccg gtctgtctc ctacgtgaaa cccgagcaac attaggctt tcagggttt 2640
tcttggtct tcttaaaaa gacttggtga ttggatatt aataatcct cgtctcaaa 2700
tgtgattat gatgaacagt caccattacc cccatctca catttgctt ccatattga 2760
agagtctcac gtgccagtaa ttgaagaatc ttgagagtt cagatatgtg aaaaggcaga 2820
agaattaaag gacattgtc ctgaaaagaa aagcacttta aatgaaaatc agcctgagat 2880
aaagcatcag tctctctcc agaaaaatgt gagtaagagg gatccacca gcagtcattg 2940
gcacagtaac aagaaaaatc tattaagt agaaaatgt gtacacgaa gaggtagatc 3000
ggtagtccc aaaaagccag ccagtcaaca ttcagaggaa catttgata agattctag 3060
tctctaaaa aataacccca aaagaagacc cagagatcaa tccctcagcc ccagcaaagg 3120
ggaaaataaa agttgtcagg tcagcaccag ggcaggctct ggacaagatc agtcagaaa 3180
aagcagaggt cgtcggcca gcccaaaaa gcagcaaaaa attgaaggaa gcaaagctcc 3240
atcaaatgt gaggccaaat tattagagg taagagtcga agaatagcag gctatcggg 3300
cagtaatgt gagcagatcc cagatgggaa ggaaaaatca gacgtcatca ggaaagatgc 3360

aaagcagaat cagttggaaa aaagcagaac aaggtctcca gagaaaaaaa tcaaaagaat 3420
ggttgagaaa tctctccat ccaaatgac taataagact acaagtaaag aagtatctga 3480
aatgaaaaa ggaaagaaag taaccacagg agaacaagt tctagtaacg ataaaatagg 3540
agaaaatgtc cagctatcag aaaagaggct gaagcaagaa cctgaagaga aggtagtctc 3600
aaacaaaaca gaagatcaca aagggaaga actagaggca gctgacaaaa acaaagagac 3660
tggaaggctc aaaccggaaa gcagttctcc agttaagaaa aactgataa ctccagggcc 3720
ctggaagggt ccaagtggaa ataaagtcac aggcactatt ggtatggctg agaaacggca 3780
gtaaccttta gtataaaca aagaaaaaca agttgtaac tttcttaca gcagcatttt 3840
tccagaaaaa gcctttttt tttttcaga tattctgaaa cagataagta catgttaatg 3900
tgagcctcaa gttacctagg ctgcatgaag ggcctttagg attgctaaga accaactgtc 3960
cccctggccg gctgccctcc ctgcctcca ggaaggagct gcaccacat gctcatctga 4020
cccgcctgc tcaggctgcc cagctcgtct tcatgagtgt ctgaacaaat gacatatgtt 4080
gatattaaca atgtggtcac aactcacttt gtattgtgc caagttatct actgtatcat 4140
gtctgtttt atccttttg ttcagctgtt tccacagtaa tgaaaaagit aggtttggct 4200
tggaagtga tgatctcaat agcatgtgc atgtttacag agagaaatat gtgagtcctt 4260
gcagaagaag agactgtta ctcatcgta aagatggccg ttgtctctc taacagctac 4320
tgatgatgc ccacttaaa aataaaacc ccaacatca ctacttaag gaaaaaaaaa 4380
atgtagtcca atattgatgc ttcttatgg cttttattt taatttggtt ggataagttg 4440
ttcaaataa ctgttaaaga tattacttac aattgaatgt ttgaaataag aaagtacttt 4500
aagcaataga gttcatctcc tgcgtgttta tccaacctcg atgtatactt acagcatctc 4560
aggtcacctt tttatttca gttattaat tatgaaacca taaagaagca tgtggaaata 4620
gtgtttattg ctctttgaag aaaaaccacc aactatttct ggatattttg gctgtaccta 4680
ctactaaagt cattagtctt taatacataa tacatatgtt aaaagtaaac atattatata 4740
gattatgtga gggacttaac catgaaacca gttcacagt ccaagtacca actcttctgg 4800
tagcagggtc acaagcttgg gtgtttaaaa acaacctgtg tagggtatgc ccagcaaatg 4860
aggacaaatg tglagacagt acttactgga tcttatttaa ctttagcta cattaactaa 4920
ctttctatt taaaaacaag aaaggagac taaacatctg cttacttgt acacattttc 4980
agaattcttt taaaagtct agttaagat gttcttaga agttggagac tgtaacaac 5040

ttccataaaa tagatccagg ttttcagtt ccctgaagca gcattcagta gcattctatat 5100
aaataaaggc accttctgag aataaaaacta tttatggag tgtgtgaaca cactgttct 5160
gtcacctggg ttcattctgt tgtgaagcac attaggtcca ggtcctccc tctgggagtc 5220
tgactgtgaa actctttaac ccaacaactc aattagcccc ttagataag acatgctcc 5280
cagagtgaga ttttgaaat cccctttca tccagaacta tattaccca cctattgaa 5340
ctattcaaat agagcaaaa taggaggctt gataaatact aagaatttag taccacagaa 5400
attatttatt atttccctg tagtccacaa ttagtgataa cgaatcctat tttgttaac 5460
tgtgacataa cttgatgtc atatgtgtc ctatgtggtt ctctctaagt aaactctga 5520
ctgattatat actgacttag caatgtggcc ttggaatgct gagcaaaatg tggatgtact 5580
ggttgtaaat gtttatatat tgtacagtac ctttatatat acacttgagg ttctgattag 5640

<210> 10
<211> 457
<212> DNA
<213> human

<220>
<221> misc_feature
<222> (242)..(242)
<223> any kind of base

<220>
<221> misc_feature
<222> (369)..(369)
<223> any kind of base

<220>
<221> misc_feature
<222> (394)..(394)
<223> any kind of base

<220>
<221> misc_feature
<222> (406)..(406)
<223> any kind of base

<220>
<221> misc_feature
<222> (457)..(457)
<223> any kind of base

<400> 10
 tcagtcactc ttcaccctg ccaaagctc actgtctac tgattgaatt gtatgtgaga 60
 aataaaatgt catcatatta agccactggg attgtatgt ttatctgta tagcagcaag 120
 tcttaatta cctaatacac acattgtgac agatgttctt aatgtccac cccatattgt 180
 tacatgtcca gctttgagga tccctggcat gtgggggtag gagttctgg gcatgctgga 240
 tncaattccc actttaagg catctgtggc ctctgtggcc tctgtggcct tcaactgtat 300
 ggaagggatt tatctggggc accataggaa actttacat ggcacagtgg acaacctagg 360
 agggggtgng gaggaggggc ctcaggccc aacngggggg accagngttc gtggggtag 420
 ggtggttg ggggtttcc ctctaccgc tgggggn 457

<210> 11
 <211> 1493
 <212> DNA
 <213> human

<400> 11
 aatagggtg gcggctgcag cgggcggcaa acagccgcc cggcaccacc atgctcgccc 60
 tggaggctgc acagctgcac gggccacact tcagctgtct gtaccagat ggcgtctct 120
 atgacctgga cagctgcaag cattccagct accctgattc agagggggct cctgactccc 180
 tgtgggactg gactgtggcc ccacctgtcc cagccacccc ctatgaagcc ttcgaccgg 240
 cagcagccgc tttagccac cccaggtcg cccagctctg ctacgaacct cccacctaca 300
 gccctgcagg gaacctcgaa ctggccccc gctggaggc cccggggcct ggctccccg 360
 cataccccac ggagaacttc gctagccaga ccctggttcc cccggcatat gccccgtacc 420
 ccagccctgt gctatcagag gaggaagact taccgttga cagccctgcc ctggaggct 480
 cggacagcga gtcggatgag gccctcgtgg ctggccccga ggggaaggga tccgaggcag 540
 ggactcgcaa gaagctgcgc ctgtaccagt tctgtctgg gctactgacg cgcggggaca 600
 tgcgtgagtg cgtgtggtg gtggagccag gcgccggcgt ctccagttc tctccaagc 660
 acaaggaact cctggcgcgc cgctggggcc agcagaaggg gaaccgcaag cgcacacat 720
 accagaagct ggcgcgcgcc ctccgaaact acgccaagac cggcgagatc cgcaaggta 780
 agcgcaagct cacctaccag ttcacagcg cgctgctgcc tgcagtccgc cgggctgag 840
 cacacccgag gctccacct gcggagccgc tgggggacct cactgccag ccaggatccc 900
 cctggaagaa aaaggcgct cccacactct aggtgatagg acttacgcat cccacacatt 960

tggggaagg ggagtgctgc cctgccataa tcccaagcc cagcccgggc ctgtctggga 1020
 tccccactt gtgcctgggg tccctctggg atttcttgt catgtacaga ctccctggga 1080
 tcctcatgtt tgggtgaca ggacctatgg accactatac tcggggaggc agggtagcag 1140
 tgcttcaga gtccaagag ctctctggg atttcttgt gatctctgat tccccagtga 1200
 ggctgggac cttttaaga tcgtgtgtg tctgtaaacc ctgaatctca tctggggtgg 1260
 gggccctgct ggcaacctg agccctgtcc aagggtccct ctgtcagat ctgagatttc 1320
 ctagttatgt ctggggccct ctgggagctg ttatcatctc agatctcttc gcccatctat 1380
 ggctgtgttg tcacatctgt cccctcattt ttgagatccc ccaattctct ggaactattc 1440
 tgctgcccct tttatgtgt ctggagttcc ccaatcacat ctagggtccc tcc 1493

<210> 12
 <211> 2292
 <212> DNA
 <213> human

<400> 12
 ccatgggttc cccttcagcc tgccataca gagtgtgcat tccctggcag gggctcctgc 60
 tcacagcctc gctttaacc ttctggaacc tgccaaacag tgcccagacc aatattgatg 120
 tcgtgccgtt caatgtcgca gaagggaagg aggtccttct agtagtccat aatgagtcac 180
 agaattctta tggctacaac tggtaaaag gggaaagggg gcatgccaac tatgaatta 240
 taggatattg aaaaaatata agtcaagaaa atgccccagg gcccgcacac aacggtcgag 300
 agacaatata cccaatgga accctgtgta tccagaacgt taccacaat gacgcaggat 360
 tctataccct acacgttata aaagaaaac ttgtgaatga agaagtaacc agacaattct 420
 acgtattctc ggagccaccc aagccctcca tcaccagcaa caactcaat ccggtggaga 480
 acaaagatat tgtggtttta acctgtcaac ctgagactca gaacacaacc tacctgtggt 540
 gggtaaacaa tcagagcctc ctggtcagtc ccaggctgct gctctccact gacaacagga 600
 ccctcgttct actcagcgcc acaaagaatg acataggacc ctatgaatgt gaaatacaga 660
 acccagtggtg tgccagccgc agtgaccag tcacctgaa tgcgcgtat gactcagtac 720
 aagcaagttc acctgacctc tcagctggga ccgctgtcag calcatgatt ggagtactgg 780
 ctgggatggc tctgatatag cagccttgtt gtagttctg catttcggga agagtgtttt 840
 tattatccac ctgcagactg gactggattc ttctagctcc tcaatccca tttctcctg 900

tggcatcact aagtataaga cctgctctct tctgaagac ctataagctg gaggtggaca 960
 actcaatgta aattcaagg aaaaaccctc atgctgaga tgtgggccac tcagagctaa 1020
 ccaaaatgtt caacaccata actagagaca ctcaaattgc caaccaggac aagaagtga 1080
 tgactcatg ctgtggacag ttttccaa gatgtccaa gcctcatcgt gacgaggctc 1140
 ttatccact ccattttcc ctgctcatg ctgcctctt aattggtaa gataatgctg 1200
 taactagaat ttacaaatca ggccttggtg caggcaattt gacagagtgt tggatgtgc 1260
 atgtcatcat gtcaaaccac aatatttgac ctaagggatc ctttattctg cccagtggct 1320
 aactttaaca acatccctaa tacaactgtt tattcaaatg cacggtggtc cctgtagag 1380
 ttagacctct agactcaact gtctcacgc cctgttttaa ttaaccag ctatgggatg 1440
 ccagataaca gaattgctgc ctacgagctg aacagggagg agttgtgca gttgctgaca 1500
 ctcttggtg cacataaata aatacagtgg gtactataga gactcagtg caaaaattaa 1560
 caaatatgct gcttgattaa aatgggtagg ctctcatgt ggctcattct ttaatctatt 1620
 ctctttatt tggttgggt catggggtct ctgcctatgg atcatactc aaactctgg 1680
 tgtgatcctc ctgattgca caatattagt taccctgggt tgcgtattc tctaaacct 1740
 ttaaagtgtt gcatgcagcc attcgtcaaa tgcataatat tctctcttg gctggaatga 1800
 caaaaactca aataaatgta tgattaggag gacatcataa cctatgaatg atggaagtcc 1860
 aaaatgatgg taactgacag tagtgtaat gccttatgtt tagtcaaact ctcathtag 1920
 tgacagcctg gtgactccag aatggagcca gtcagtctaa atgcatata ctacactga 1980
 aacatgagga agcaggtaga tcccagaaca gacaaaactt tctaaaaac atgagagtcc 2040
 aggcgtctg agtcagcaca gtaagaaagt ccttctgct ttaactctta gaaaaaagta 2100
 atatgaagta ttctgaaatt aaccaatcag ttattttaa tcaatttatt tatattctc 2160
 tgttctgga ttccatttt acaaaaccac ctgttctact gttgtattgc ccagtaggag 2220
 ctatcactat atttgcaga atggaaactg cctgactct tgaatcaca ataaaagcca 2280
 attgtatctg tt

2292

<210> 13
 <211> 519
 <212> DNA
 <213> human
 <220>

<221> misc_feature
<222> (212)..(212)
<223> any kind of base

<220>
<221> misc_feature
<222> (451)..(451)
<223> any kind of base

<400> 13
gaaacaacaa cagtgtaatc ttaacaggg atgttaaagg taagaagta ggaagataaa 60
ccaaaatgat tgagtatgat aaagaatttt gcatggcgat taaaatagaa aacctataaa 120
tgtagaaaaa gcaggtctgg acttagcaaa gaaacaatat agtttgaga aggcattgaaa 180
taagttcttt tcatgttcac tgctggtcac ancataacag agagtgtatg ggagagcttt 240
gggaagggtt cacgttgagt tacatcagtg gtcaacaatg gagcaacaag actccgtaga 300
ggatgccacc ctgggagaat tgcaagggaa aggaggctga agcacaactg gtaatagcct 360
tcagatatat aatggatatg caaataaagc tctgattaat tgtatttca cttattatat 420
atcatcttgg gacctttcta aaagtgggac nctagaaaag atatactgaa actccaaaag 480
aatacttcag ctcgagtga atggattcaa gatgttgtt 519

<210> 14
<211> 5294
<212> DNA
<213> human

<400> 14
ggctcgcac cccatagtc tgggttacag tgaaggtag ccccgcgctc tgctctggag 60
aggcaggggtg gcatagggaa cgtctcgagt ggcgcccga gtcattgttg tttcgttg 120
ccgcccctc ccggcgctcc tagggctgtt taagaagaag ggctctgcca aggcctgagaa 180
tgacaaacat ctaagttag ggcttgcca ggggccaggg tctgcagtgg atgagcacca 240
ggacaacgtc ttctttcca gtggcgacc ccccccactg gaagagctgc acactcaggc 300
ccaggagggg ctccgctccc tacaacacca agagaaacag aaactgaaca aggggtggctg 360
ggaccatgga gacaccaga gtatccagtc ctccgggacg gggccggatg aagacaacat 420
ctcctctgc agtcagacca catctactgt ggctgagagc tccacagcag aggacgcgt 480
ctccatccgc tcggagatga tccagcgcaa aggtccacc ttccgacccc atgactcatt 540

tcccaaactt ggaaagtcag ggcggcgctg gcgggagcgg cggagcactg tgctgggact 600
cccgagcat gtgcagaagg agcttggcct gaggaatgag cgtgaggcac caggcacgcc 660
ccgggctcct ggtgcacggg atgccgtacg catccccaca gtggacggcc gcccccgagg 720
cacctcaggg atggggggccc ggggtgccct gcaggcgctg gaggcggagg cagaggctgg 780
cgctgagaca gaggccatgc tgcagcgcca cattgaccgt gtctaccggg atgacacctt 840
tgttggccgg tccacgggta cccggggccc accattgacc cggcccatgt ccctagcagt 900
gcctggattg acaggagggg cagggcctgc agagccccctg agcccgccca tgtccatctc 960
ccccaggcc acctacctgt cgaagtgat tccacatgct gtgtgccgc ctacagtga 1020
cgtggtggcc ctaggccgt gcagcctgcg cacactaagc cgctgcagcc tgcactggc 1080
cagcccagcc tcagtccgt cgctggggcg ctctctctc gtctccagcc cacagccccg 1140
cagccggcac ccatcctct ccagtgcac ctggagccac tctcaatct ccgacacat 1200
tgtgtctgac ggtccaccc tctctctaa ggggtgctct gagggccagc cggagagctc 1260
tacggctagc aatagcgtg taccctctc ccagggaggc agtgggaggg gctctcccag 1320
tgggggcagc actgtgagg cctcagacac actcagcatt cggagcagtg ggcagttgtc 1380
tggccggagt gtgtccctgc gtaagctgaa gcggcctcca cccctcccc gccggacca 1440
ctccctcat cagcggggct tagcagtgc tgatgggcca ttaggtgtc cccctaagcc 1500
tgagcgtgag cagcagcccc agctgcctc gccaccacc actggtggct cagaaggggc 1560
gggggcagca ccctgtccac ccaaccagc caacagctgg gtacctggct tgtctccggg 1620
tggttccgg cgccccccac ggtccccaga acggacactt tcgacctcca gtggatactc 1680
gagccaaagt ggtactcca cctcctctc caagggcctg gcaggctccc ctgttcccc 1740
aggcaaggcc cagcccccta aaccagagcg gtgcacgtct ctctctccc ctggggcctc 1800
cgtctctct tccctcacgt ctttatgtc ctctctctc gaccagccc cctcagaccg 1860
ctctgggcca cagatattga cccccctgg tgacaggtt gtcatactc ctaccccaa 1920
ggtgcctgcc ccctctccc cactcctc caagcccagg agccctaacc cagctgcccc 1980
tgctctagcc gcccctgctg tggttcctg gcctgttct accactgacg ccagtcctca 2040
gtccccctc actccccaga caacctgac tccactgcag gactctctg tcactccaa 2100
agaccagtca cccccactt cccaccccc atcttatcat ccacccccac caccactaa 2160
gaagccagag gtggtgtgg aggcaccatc tgcctcagag actgtgagg agcccccca 2220

agatcccaac tggccccctc ccccccacccc tgcccctgag gagcaggacc tgtccatggc 2280
tgactcccc ccaccagagg aggcttttt ctctgtggcc agccctgagc ctgcaggccc 2340
ttcaggctcc ccagagcttg tcagctcccc ggctgcttcg tctcctcag ctactgctt 2400
gcagattcag cccccgggta gcccagacc tctccagct ccgccagccc cagctcctgc 2460
tagttccgcc ccagggcatg tggccaagct cctcagaag gaaccggtgg gctgtagcaa 2520
gggtggtggg cctcccaggg aggacgtagg tgcgcccctg gtcacgccct cgctcctgca 2580
gatggtgctg ctgcgtccg tgggtgctcc aggaggggct cccaccccag cactggggcc 2640
atcgcccccc cagaaaccac tgcgaagggc cctgtcaggg cgggccagcc cagtgcctgc 2700
ccccctca gggctccatg ctgcgtccg actcaaggcc tgcagcctgg ccgccagtga 2760
aggccttca agtgtcagc ccaacggacc gcctgaggca gagccacggc ctcccagtc 2820
ccctgcctca acggccagtt tcattctc caagggtctt aggaagctgc agctggagcg 2880
gcccgtgtcc cctgagacc aggctgacct ccagcggaat ctggtggcag aactccggag 2940
catctcagag cagcgccac cccagggccc aaagaagtca cctaaggctc cccacctgt 3000
ggcccgcaag ccgtctgtgg gactcccccc accgcctcc cccagttacc ctgagctga 3060
gcccctact gctctccca ccaatgggt cctcacacc caggacagga ctaagaggga 3120
gttggcggag aatggagggt tctgcagct ggtgggcca gaggagaaga tggcctccc 3180
gggtcagac tcacagaaag agctggcctg accaccaggc acctactgg cactgtgac 3240
ccatcccaga aacacaatct cagggacctg agcagctcca aggacgagag gatacagcag 3300
acacaacctc atagagaggg cgctgcagc cttacctcc acggccttcg atactatgc 3360
aagcctgggt ttgtctgt cctcagatc atctgcgt catgccttt cccgaatggg 3420
ttaccttg gcagttgcc ctctagctt ggccttagcc tcatttgaa gtgggtagct 3480
ggcgggagag ggtggctgcg cccctgctg gccctgaggc tgcagagtg ggagcaggac 3540
acctcacctg agttcattt ttttcatgt ccaaaccatg cacatactat agtcagaat 3600
caaagcactt ttgaaaagt gctgcatggc catctccag ggcccaggaa gttgcattcc 3660
aagggcctgt ttacatggca gcagaatcca tccccggcag tcagccata gcttgggacc 3720
agtctgtgcc ctctgccc gtcagttta ctctcttg ttctgaagg tggccaagtc 3780
attgtgtcc cacaggctc tctaggctgg gggcagggtg ggggctgtg aattccaaag 3840
cacaaaagg gcagagggga ttggcctcc tgtgcctcaa ctaccaacc acctcctgc 3900

cticcagttc tgccaggtgc tccatgctgg ggacaagtag gagactgccca gggcccaaag 3960
 aaatgggtga gcagtagagt catctcgggg cacttggcag tgtcaagcac ctgccccttg 4020
 cctccttgac cactctgggg tgggtgggcc ccagcactt cagaggcagg agcctttggg 4080
 ctgagcaagc actgaggagg tggatggaag ggagcatctg gaggggggga gcttcctga 4140
 gcagtgggcc caggcctggc cctccacact tcattctctg acctttctct ctctcattt 4200
 cgggtcatgt ctttctgca gctgccttc agcacagggt gtccactgg gggcagctaa 4260
 cgctgagtga caaggatggg aagccacagg tgcattttac tcaagtcttc tctagtcaat 4320
 gaggggcacc cagtgtctct agggcaggct ggggtggtgt cccctaggta tcagcctctc 4380
 ttactgtact ctccgggaat gtaaaccttt ctatttcag cctgtgccac ctgtctaggc 4440
 aagctggctt cccattggc cctgtgggt ccacagcagc gtggtgccc ccaggggcca 4500
 ccgtctctt ctgatccct ttccttaac agtgacttg gcttgagtct ggcaaggaac 4560
 ctgtcttta gctcaccac caaggagaga ggtgacatg acctccccgc cccctacca 4620
 aggtctggaa cagaggggat gtgtgagag ccaggttct ctggccctct ccagggtgtt 4680
 ttccactagt cactactgc ttctcctgt agctaataa tcaatattct tccctgcct 4740
 gtgggcagtg gagagtctg ctgggtgtac gctgcacctg cccactgagt tggggaaaga 4800
 ggataatcag tgagcactgt tctgtcaga gctcctgac taccacccc cctaggatcc 4860
 aggactgggt caaagctgca tgaaccagg ccctggcagc aacctgggaa tggtggagg 4920
 tgggagagaa cctgactct cttccctct cctcctcca acattactgg aactctatcc 4980
 tgttaggatc ttctgagctt gttccctgc tgggtgggac agaggacaaa ggagaaggga 5040
 gggctagaa gaggcagccc ttcttgtcc tctgggtaa atgagctga cctagagtaa 5100
 atggagagac caaaagcctc tgatttttaa ttccataaa atgttagaag tatatatata 5160
 catatatata ttctttaa ttttgagtc ttgatatgt ctaaaaatcc attcctctg 5220
 ccctgaagcc tgagtgagac acatgaagaa aactgtgtt cattaaaga tgtaattaa 5280
 atgattgaaa ctg 5294

<210> 15
 <211> 988
 <212> DNA
 <213> human
 <400> 15

gtcgtgaggc gggccttcgg gctggctcgc cgtcggctgc cgggggggtg gcctgggtgt 60
 cattggctct gggaaagcggc agcagaggca gggaccactc ggggtctggt gtcggcacag 120
 ccatggcggg cgcgttggtg cggaaagcgg cggactatgt ccgaagcaag gattccggg 180
 actacctcat gactacgcac ttctggggcc cagtagccaa ctgggggtctt cccattgctg 240
 ccatcaatga tatgaaaaag tctccagaga ttatcagtgg gcggaatgaca ttgccctct 300
 gttgctattc ttgacattc atgagattg cctacaaggt acagcctcgg aactggcttc 360
 tgtttgatg ccacgcaaca aatgaagtag cccagctcat ccaggagggg cggcttaica 420
 aacacgagat gactaaaacg gcatctgcat aacaatggga aaaggaagaa caaggtcttg 480
 aagggaacgc attgccagct gctgctgagt cacagattc attataaata gccicccata 540
 ggaaaataca ctgaatgcta ttttactaa ccattctatt ttatagaaa tagctgagag 600
 ttctaaacc aactctctgc tgccttaaca gtattaaata tttacttct ttccataaag 660
 agtagctcaa aatatgcaat taatttaata atttctgatg atgtttatc tgcagtaata 720
 tgtatatcat ctattagaat ttacttaatg aaaaactgaa gagaacaaaa ttgtaacca 780
 ctagcactta agtactcctg attcttaaca ttgtctttaa tgaccacaag acaaccaaca 840
 gctggccacg tacttaaaat ttgtcccca ctgtttaaaa atgttacctg tgtattcca 900
 tgcaagtgtat atattgagat gctgtaactt aatggcaata aatgatttaa atattgtta 960
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 988

<210> 16
 <211> 4908
 <212> DNA
 <213> human

<400> 16
 ggataacctc gcagggtggg ccggagggcg ggcgccgccc ctgcctgtgc tgcggcgatg 60
 gccagtggtg tacaatcagt gcaggagcta atcccggaact ccttcgtccc ctgtgtcgt 120
 gcgctgtgca gcgacgaagc cgagcggctc actcgtctca atcacctcag cttcgcggag 180
 ctgcttaagc ccttctcccg cctcacttcc gaggttcaca tgagagatcc taataatcaa 240
 cttcacgtaa ttaaaaattt gaagatagca gtaagcaaca ttgtcaccca gccacctcag 300
 cctggagcca tccggaagct ttgaaatgat gttgtttctg gcagtcagcc tgcagaagga 360
 ttagtagcta atgtgattac agcaggagat tatgacctta acatcagtcg cactactcca 420

tggttgagt cttacagaga aacctttctt cagtcgatgc cagcatcgga tcatgaatt 480
ctgaaccact atttagcatg tatgttgga ggcgtcatcga gtgaagctga acctgtggaa 540
cagttttcaa agttgtcaca agaacagcat cgaattcagc acaacagtga ttattctac 600
cccaagtggg ttataccaaa tacacttaaa tactatgtac tttacatga tgtaagtga 660
ggagatgaac agagagctga atcaatttat gaagaaatga aacagaaata tggaactcag 720
gggtgctatt tacttaaaat taattctga acatctaac gagcatcaga tgaacagata 780
ccagatcctt ggagtcagta tctccagaaa aatagtattc aaaaccagga atcatatga 840
gatggccctt gtactataac ttcaaataag aattctgata ataactgct tcattggat 900
ggattagata acgaagtcaa agatggctta ccaaataact ttgagctca ccacttcag 960
ttggagcaat ccagtgacc ttctaacagt attgatggcc cagatcatc aagatctgct 1020
tcatcgctac atgaaacaaa gaaaggaaat actggaataa ttcatgggtc atgtttaaca 1080
cttactgatc atgatagaat tcgacagttt atacaagagt tcacattcg gggcctttg 1140
ccacatatag agaaaacaat taggcaatta aacgatcagc taatatcaag aaaaggttg 1200
agtcgaicgc tatttctgc aactaaaaa tggtttagtg gcagtaaagt tccagaaaag 1260
agcattaatg acctgaaaaa tacatctggc ttgctgtatc ctccggaagc accagaactt 1320
caaatcagga aaatggctga cttatgttt ttggtgcagc attatgatt ggcttacagt 1380
tgctatcata ctgcaaagaa agattttctt aatgatcaag caatgctta tgcagctgg 1440
gccttggaat tggcagcagt gtctgtttt ctcaaccag gagcacctag gccatatcct 1500
gctcattaca tggatacagc aattcagaca tacagagata tctgcaagaa tatggtgtg 1560
gctgaaagat gtgtgtgtc tagtgctgaa ctttaaaaa gccaaagcaa atattcagag 1620
gctgcagctc tctaatacg gttgaccagt gaggattctg atcttcaag tgcacttct 1680
ttggaacagg cagcacattg cttataaac atgaaaagtc ccatggtag aaaatatgca 1740
ttcatatga tattggcagg ccatcgattt agtaaagcag ggcagaaaaa gcatgctta 1800
cgctgttatt gtcaagccat gcaagtttac aaaggaaaag gctggtctct tgcagaggat 1860
cacattaatt tcactattgg ggcgcagtc tatactctta gacagctgga taatgctgtg 1920
tctgtctta ggcatactt aattaatgaa agtaaacaat ctgctgtca acagggggct 1980
ttctcagag aatatctta tgtttacaag aatgtaagtc agctgtcacc agatggctct 2040
ttgccacagc ttctttacc gtatattaac agtcagcaa cacgggttt tttggccat 2100

gacagacgac cagcggatgg tgaaaaacaa gcagctactc atgtaagtct tgatcaagaa 2160
tatgattctg aatccictca gcagtggcga gaactgagg aacaagttgt ttctgtggtt 2220
aacaaggag taattccatc caatttcat cccacacaat actgttgaa cagtactca 2280
gataattcaa gattccact tgcagtgtga gaagaaccaa ttacagtga agtggcttt 2340
agaaccctt tgaaagtct actttgtg actgattgt cattgcttg gaagttcat 2400
cctaaagatt tcagtggaaa ggataatgaa gaagttaac aactagttac aagtgaacct 2460
gaaatgattg gagctgaagt tatttcagag ttcttaatta atggcgaaga atcaaaagt 2520
gcaagactaa agctcttcc ccatcacata ggggagctgc atattctggg agttgttat 2580
aatctggca ctattcagg ctctatgaca gtagatggca ttgtgtct tcccgatgt 2640
cacacaggaa aatattcctt gagtatgtca gtccgagga agcaggatt agaaattcaa 2700
ggctctcgac ttaacaacac aaaagaagag aaaacatctg taaatatgg cctgatcga 2760
cgtttagatc ccataatcac agaagaaatg ccaactgttg aggtgtctt tatacattt 2820
cctacagggc ttctgttg agaaatccga aaagcatatg tagaattgt caatgtcagc 2880
aatgtccac ttactggatt gaaggtgtt tctaaacgtc cagagttctt tacttccgt 2940
ggtaatactg ctgttctaac accactaagt ccctcagctt ctgagaattg tagtgcttac 3000
aagactgttg tgacagatgc taccctgtg tgtacagcac tcatatcatc agcttctt 3060
gtagacttg gcattggcac aggaagtcaa ccagaggtga ttctgttcc cctctcgac 3120
actgttctt taccggagc ctcagtgcag ctgccaatgt ggttacgtg gctgatgaa 3180
gaaggtgtcc atgaaattaa cttttgtt tactatgaaa gtgtcaaaaa gcagcaaaa 3240
atacggcaca gaatattaag acacactgca attattgta ccagtcggtc ttaaattgta 3300
cgggccactg tctgcagaag taattctt gaaaatgaag aaggcagagg aggcaatat 3360
ctagtcttg tggatgtgga aaataccaat actagtgaag caggcgtaa ggaattccac 3420
atagtcaag tatcaagtag tagcaaacac tggaagttac agaaatctgt aaatcttct 3480
gaaaacaaag atgcaaact tgccagtagg gagaaggga agtttgctt taaggcaata 3540
agatgtgaga aagaagaagc ggccacacag tctctgaaa aatatacctt tgcagatc 3600
atcttggaa atgaacagat aataagtca gcaagccat gtgcagactt ctttatcga 3660
agtttctt ctgaattgaa aaaaccacaa gtcactgc ctgtgcatac agaaaaacag 3720
tcaacagagg atgtgtgag atgattcaa aaatgcagt aggtagattt gaatattgtc 3780

atattatgga aggcatacgt tgggaagac agtaaacagc ttatttggga aggtcaacat 3840
 catgttattc ttgcactat aggaaaagaa gccttttcat atcctcagaa acaggagcca 3900
 ccagaaatgg aactattgaa attttcagg ccagaaaaca ttacagttc ctcaaggcca 3960
 tcagtagagc agcttctag tctcattaa acgagtcctc actaccaga atcattaat 4020
 catccattc atcaaaaaag ccttggtta gtaccagtca ctttttact ttcaattgt 4080
 tctaaggctg atgtagatgt catagtgtat ctcggcata aaacaacaag tccagaagca 4140
 ctggaaatcc atggatcatt cacatggctt ggacaaacac agtataaact tcaactaaa 4200
 agccaggaga ttacagctc gcagctgaaa gcagcttg ttacacagg tgttataac 4260
 ctggaactc ctagggtatt tgccaagta tcggaccaag ttacagtgt tgaacaagt 4320
 cagcagaatt ccatgcctgc cctgatcctc atcagtaatg tgtgacaact tggaaattg 4380
 tactgaaatc cacaataatc agtttttctt ggtgggtt tacagcagta ttgatatac 4440
 ctaacttgtt atggaggtg atgtatatc gatccctgca aaatacttg actgtcatt 4500
 ttgtgatga tgcaaagcac gttggactga gaatactaa cattctttt ctgtattct 4560
 ttaaaccctg agaataatt acatgctcat aatacaggat atcagcatat ttgtcacct 4620
 tattaagccc catctaaga aaacacaaag tctaagctg ctgtacaac ttgtcaatgg 4680
 tatacgaata ttaggatg atctcagaa aggaaggcc ttgtggcag tactcctgt 4740
 aagccattg tctctaaat ccagctttac tgtgaagtc tatagagtgt taaatacaaa 4800
 tttctctg ttcttcaca cagttccta aaatcagtt tgaacttgg tcatagagtc 4860
 ttcatattc agtatttgt ggtccctatg actatacat aacttgt 4908

<210> 17
 <211> 435
 <212> DNA
 <213> human

<220>
 <221> misc_feature
 <222> (30)..(30)
 <223> any kind of base

<220>
 <221> misc_feature
 <222> (49)..(49)
 <223> any kind of base

<220>
<221> misc_feature
<222> (75)..(75)
<223> any kind of base

<220>
<221> misc_feature
<222> (76)..(76)
<223> any kind of base

<220>
<221> misc_feature
<222> (78)..(78)
<223> any kind of base

<220>
<221> misc_feature
<222> (79)..(79)
<223> any kind of base

<220>
<221> misc_feature
<222> (109)..(109)
<223> any kind of base

<220>
<221> misc_feature
<222> (136)..(136)
<223> any kind of base

<220>
<221> misc_feature
<222> (137)..(137)
<223> any kind of base

<220>
<221> misc_feature
<222> (149)..(149)
<223> any kind of base

<220>
<221> misc_feature
<222> (227)..(227)
<223> any kind of base

<220>

<221> misc_feature
<222> (236)..(236)
<223> any kind of base

<220>
<221> misc_feature
<222> (246)..(246)
<223> any kind of base

<220>
<221> misc_feature
<222> (342)..(342)
<223> any kind of base

<220>
<221> misc_feature
<222> (363)..(363)
<223> any kind of base

<220>
<221> misc_feature
<222> (389)..(389)
<223> any kind of base

<220>
<221> misc_feature
<222> (426)..(426)
<223> any kind of base

<400> 17
ggtagaatg attgtgatg acaaatttn tatttgatc atacttaana agacagagca 60
gactcacatt cattnncna atagtatcac tgtacacata gcgaattnt ggcgcttta 120
gattgctcg aaaatnctg aagagtgnc catagcagcc tggtaacct ttctcttcc 180
cccaaagctc tctgccctt tgcagaaaga ctgttggtga caactgntgc taactnaata 240
gcatngggtt gaacttcgcc aaaatcctc cacctctcc catagggcaa caggggtgac 300
ttgggcttaa agggcattga gtaagcaagt aggttatcag anaacagagg gaagattcca 360
ttntagataa ttccaaata ttacaattng tggaactcag agttcaactg ctcagttcct 420
tcttngctg accct 435

<210> 18
<211> 2224

<212> DNA
<213> human

<400> 18

ctttagatct gtgcagcctt tgcgtgccaa acttgtgaaa ttccctttac ctttttga 60
gtacttgcta taaagccacc tgtaacaaa cccccattat gtacagaata ggacctatcc 120
agtagccagg ccagtaggca gtggggaag gtggaagga tccagcgagg cccctgagcc 180
tgcacctgga cagggtgacg tctgcacca tcacctcag caccaggcca ccctgcagtc 240
cacttactgt actgtgtgt ggaaggatat gctaagtgat gaaagttgcg agcagtcctca 300
ctggtcgtgt aaacttttt ttttttgg aaattgaagc ttagagtgct tgcccgaat 360
ctctaggaag ttggtggcaa gggacagcac tcacactctt ctggtcatga tctctgatct 420
ccacctcaaa tgacaataaa aaactggtcc aacgaagaca ctgctcagca cttcagccat 480
caggactaat ccatcgatga ctggaaaaga ggctagcttt gagggaaaaca gcctgggctc 540
ttgggagcag agtccagtgg gtgtgaggct gacttgccga cggtcggcag gtaatggctc 600
tcagccggcg aggcgggtccc acagctctcc tcccagggca gcctgaggag gaggaggccg 660
ggtcctgtt tgggtgcagc ttcagcctag ggatacctga agctgttgag caacaccttt 720
atgaaatgtt gccagagcag caacacttcc ctgtgggcac agccccggga aatccggtag 780
caagttagca aggtggcagg acccacccaa gcctgatacg catctgggcc cgccgggctc 840
agcaggggag gctgctacgg ctgccactt cccagcaccg tctgtcaggc ttgaaccctt 900
ctgtgctgtt ccttctctgg ctaataggga gacccttcgc aggcaccac tgtttcaact 960
tgacctccc acccctgct actctctcc acacaccct ccgttcgct agcctaccct 1020
gtcagcctt caataaaagt tatgcacaaa tgtgaacacc tgagatggag ctgaacattt 1080
cttcaattg tctttttct gaagtcaaac tctatcaaa tgccctaaaa ttattaccac 1140
ccaagagaaa caggaaaaag gttacatgtt ttgtttact gagagtaaga tcacctgcat 1200
ctggaagacg ggctggtaaa ttggttggc tacagaacag aaagaaaaca aaaacaaacc 1260
tcgtaaggga agtatgcac tcagacacca ccacttccta gagccaaatg agcaatccca 1320
aactgcaagt gccgtaagt ggctgtgac gtcacaccgc ccggcccgag gtatgcatg 1380
tgcgggggag gccacacta cagctgtcct ctgtctaga aggcaccacc tcgtttcat 1440
gtcccgtgtg ttttgaaaa gcagtatgtt gtgtcatgtc tagcggcgaa cacttcctc 1500
cctctgtcct tgaggttgta atataaaac tgtgttctg tacgtgtggg tggaattct 1560

ctgacgggtgc tcgttcatac cacaagctta cgctgagttc tgaactgtcg ttcacagctg 1620
 cgtgtctgca tgggtgcga tctgtgtac ctttggggaa aatttgtatg taaatgtaca 1680
 gaaataaaaa cgttgcccca ttaacagatt tcctctggaa tgtctccct acctcacctg 1740
 atggtatcca ccgaaggga tttactacc attaatgggt agtaataaaa tcctccgtgt 1800
 tcattcagac ctactgcgt cactacttg aacgcctctg taagctgtgt cttcacccgc 1860
 cccgagggtg gtggaggag gcctctact ctgctcgag tcctggtctt aaaggtagtc 1920
 agaggcagag gctggattaa acacacactg ttaccaagt gccactctca gaccacctga 1980
 gagacggggg gccatcagta aaattaagag gaatttttt ccctgttcg tgtatgtct 2040
 gctgatccgt ggctgaagg ttctagaga cgtaagaaa tgaatatctt acactgtgat 2100
 tctgtgagga aagactggta acccaaaact ctctctcta atgtatttt taacgaaaat 2160
 gacaatattt cttaataaa gtatttatac caaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2220
 aaaa 2224

<210> 19
 <211> 2244
 <212> DNA
 <213> human

<400> 19
 gttgtttaa agcaaggcat gctgtggat gactctgtaa cagactaatt ggaattgtg 60
 aagctgctcc ctggtccac tctggagagt aatctgggac atcttagtgt ttgtttgt 120
 ttttccct cctctttt tgggggggag tgtgtgtgg gttgtttt tagtctgtt 180
 ttttaattc attaacagt ggtagcctt aaggggagga ggacggattg attccacatt 240
 ccacttcta gatctagtt agaaaacatg ttcccatct ggtgcttta ggaaggagta 300
 tagtaaatgc ctatttaat aacatactc ttttgaaag ttgcctttc tctccacct 360
 tgagtagatc cagtattga tgaaactcat gaaagtgggt ggagcccatc tccccctcc 420
 tctttctag gacgcactat atgtgactgt gactttaagg acattgttt gccatttgct 480
 gattttttg ggaagttaat tttaacttc ttactgat aatgaagaa aagtattgca 540
 ccttgaaat gcaccaaag aattgagtt gtaattaaaa aaattttt tcccttcag 600
 tcattgtctt atatgcttag catagattg cagctcagta gtatatgtgt tcctagaatg 660
 cagctgaaga cctgttatgt agaggaaata cgaggggtgg tgctagaaga cagacatctg 720

tggaatgatt cacatcctct caagttagga ggaatggaggc ctgcttcatt aagaagctgg 780
 gggtaggggtg ggggtgggga gaacacttaa caacatgggg accagtcagg ggaatcccct 840
 tattctgtt ttgcatatga ggaaccctag agcagccagg tgaggctctc tagttaata 900
 aaaatcatgg aaagactctt aatgcagact cttcttaagt gtaatatagg atttttcag 960
 ctatttttg ttgcagtttc caatttttaa aaatgttgag gtaatcttc ccacctccc 1020
 aaacctaatt ctgtagatg cattagtgt gaaccaatgc ttctcatgtc tcaatctgt 1080
 atatcatctt ttcatatga ttaacaaaca aaaccttaaa aagagtagat gaattgcaa 1140
 acacaattcc taccaataat aaatcgatca actctatcta ttcaggaaag caggaagcat 1200
 ttggaccaca gtgcatgaaa actcaacat tctgttatta gataatgaat caaccaaag 1260
 aacaatccag agaaaagaaa attgcaataa taaaaggtaa attaacagaa agataatata 1320
 agcaagatag taatagtga ccattctgaa aagcttataa catcactcat catccagcat 1380
 cctttctgaa aacaaaggat tttaaatca cttatgcac atatacaaca taggaggttg 1440
 gcaaaataat gcactatttc ttaacagcca tgtctctgtt agaactcaa gtaatctac 1500
 aaatgaccat tgttcttaa tttagattat gaataccaca ttagtcaggt attgcacta 1560
 acccttaata gtatatacag ttctatgga aaattcagtg gtccaaaaat ttccgtagaa 1620
 ttgagagga cgttggtggg ctgaagatag ctcttgagg gtcactgatg taggctgcaa 1680
 tgggggttca caaggccctg acaccgtatt tatagtctaa ccttttatg aaaatctgac 1740
 tacagctatt taaggagtag tcttaatagc tgaaaatgaa gatagagaaa gacaccaaga 1800
 atatgacaca gttacattc tagtgaggga cacaacaaaa tcaaatttaa aaaagagtgt 1860
 aatagatgct gataaatact gtagataaag cacataagaa aatagaaata aaggctgtca 1920
 atggagaagt catgattttt attttattta ttatttatt tatttgagac agagtcaggc 1980
 tctgtgcagg ctggagtga atggtgtgat ctgcctcact acaacctctg ctctggctc 2040
 aagctatcct cccacctcag ctctcaagta gctgggatca caggctcgtg ctaccatgcc 2100
 cggctaattt ttgtagaga tgaggttttg ccatgttgc caggctggtc tcgaactcct 2160
 ggactcaact gacccacct cggcctctca aagtgtgag attataggcg tgcagccggc 2220
 agctggccat tgttatgtt ctgc 2244

<210> 20

<211> 351

<212> DNA
<213> human

<220>
<221> misc_feature
<222> (62)..(62)
<223> any kind of base

<220>
<221> misc_feature
<222> (121)..(121)
<223> any kind of base

<220>
<221> misc_feature
<222> (207)..(207)
<223> any kind of base

<220>
<221> misc_feature
<222> (220)..(220)
<223> any kind of base

<220>
<221> misc_feature
<222> (276)..(276)
<223> any kind of base

<220>
<221> misc_feature
<222> (300)..(300)
<223> any kind of base

<220>
<221> misc_feature
<222> (315)..(315)
<223> any kind of base

<220>
<221> misc_feature
<222> (336)..(336)
<223> any kind of base

<400> 20
tctactcca catcgcgag accgagaagc gctgttcat cgaggaaatc cccgacgaga 60

cnatggatcat cggcaactat cgtaccacaga tgtgggataa gcagaaggag gtcttctgc 120

nctcgacccc tggcctgggc atgcacgtgg aagtgaagga ccccgacggc aaggtggtgc 180
tgtcctggca gtacggctcg gagggcnctt tcacgttcan ctcccacacg cccggtgacc 240
atcaaatctg tctgcactcc aattcttacc aggatngctc tcttgcgtg gtgggcaaan 300
tgcgtgttc atctngacat ccaggtttgg gggagnatgc caacaaatta c 351

<210> 21
<211> 2631
<212> DNA
<213> human

<400> 21
accttccaac ccagccctcg gctgagccgc gccgcacat gcccgccgtg gacaagctcc 60
tgctagagga ggcgttcag gacagcccc agactcgctc ttactgagc gtgtttgaag 120
aagatgctgg caccctcaca gactatacca accagctgct ccaggcaatg cagcgcgtct 180
atggagccca gaatgagatg tgcctggcca cacaacagct ttctaagcaa ctgctggcat 240
atgaaaaaca gaactttgct cttggcaaag gtgatgaaga agtaattca acactccact 300
atttttcaa agtgggtgat gagcttaac ttctccatac agagctggct aaacagttgg 360
cagacacaat ggttctacct atcatacaat tccgagaaaa ggaatcaca gaagtaagca 420
ctttaaagga tctatttga ctcgctagca atgagcatga ccttcaatg gcaaaatata 480
gcaggctgcc taagaaaaag gagaatgaga aggtgaagac cgaagtcgga aaagaggtgg 540
ccgcggcccc gcggaagcag cacctctcct cccttcagta ctactgtgcc ctcaacgcgc 600
tgcagtacag aaagcaaatg gccatgatgg agcccatgat aggccttgcc catggacaga 660
ttaactttt taagaaggga gcagagatgt ttccaaacg tatggacagc ttttatact 720
ccgttgacga catggttcaa agcattcagg tagaactgga accgaggcgg aaaagatgcg 780
ggtgtcccag caagaattac ttctgttga tgaatctgt tacactccag actctgatgt 840
ggccgcacca cagatcaaca ggaacctcat ccagaaggct ggttacctta atcttagaaa 900
caaaacaggg ctggtcaccg ccacctggga gaggccttat ttctcacc caggcgggaa 960
tctcatgtgt cagcccaggg gagccgtggc tggaggttg atccaggacc tggacaactg 1020
ctcagtgatg gccgtggatt gcgaagaccg gcgctactgc ttccagatca ccacgccccaa 1080
tggaaaatcg ggaataatcc tccaggctga gagcagaaag gaaatgaag agtggatatg 1140
tgcaataaac aacatctcca gacagatcta cctgaccgac aaccctgagg cagtgcgat 1200

caagtgaat cagaccgctc tgcaagcagt gactcctatt acaagtttg gaaaaaaca 1260
 agaaagctca tgcccagcc agaacctgaa aaatcagag atggaaaatg aaaatgacaa 1320
 gattgtccc aaagcaacag ccagtctacc tgaagcagag gagctgatcg cgcttgagc 1380
 gccgattcaa ttgatattg tgcttctgc tacagaattc ctgatcaga acagaggag 1440
 caggcgtacc aaccttttg gtgaaactga ggatgaatca ttccagaag cagaagattc 1500
 tctttgcag cagatgtta tagtcgggt ttgggaatca atggcagta aaacagacag 1560
 cactactgaa gtgattatg aagcgatgag acaagtattg gctgctggg ctattcataa 1620
 catctccgc atgacagaat cccatctgat ggtcaccagc caatcttga ggtgataga 1680
 tccacagact caagatcaa gggccaatt tgaactacc agtgcacac aattgctgc 1740
 tcatcaagaa aacaagagac tggttggtt tgcacccgt gtctcgaat ccactggaga 1800
 agaatctctg agtacatata ttttgaaag caactcagaa ggcgaaaaga tatgttatgc 1860
 tattaattg ggaaaagaaa ttattgaggt tcagaaggat ccagaagcac tggctcaatt 1920
 aatgctgtcc ataccactaa ccaatgatgg aaaatatgta ctgttaaagc atcaaccaga 1980
 tgacgatgat ggaaatccaa atgaacatag aggcgcagaa tccgaagcat aactcattg 2040
 cgctgtggg ggaagagcga acaggaagga gagctacctc ctaagggtt taactctct 2100
 gacatacagg cacactgacc tgattccga aggctgacaa tcgttggtg aatgtaatt 2160
 tgatgcctg atactgagac ttgggagga aactaagaaa tggtagacag cgtcccacc 2220
 catctacaat gttatttag gtgcttctg glaagtctt ttcttagat tgcgctaaaa 2280
 ttcttagat tgtcagcgc tcagaacaaa agttgaaaa atgcattgt catatgaatg 2340
 tcactcttt tcagttcca gtaaccttt taaaaaatg caaaagccta gattacaat 2400
 ttgatgaaca ctaaataatt ctattaata taatctatt ttgtattta ottaatgagc 2460
 ttaagtgcc tgcgttctg aaaattgtg atttataatt cagcttatct cataattgga 2520
 cctaatagca ttcttttg cagttagggt atgagcactg cttgaggcc caagcactag 2580
 tagagatgcg cgatacaggt ctagttcgg taactgttc agacatcaag c 2631

<210> 22
 <211> 2851
 <212> DNA
 <213> human

<400> 22

agcatctcag gccatcatcc tgaacttg cagccttcgt ggagtataag gacagcatta 60
ttagccatca ttgggtttac tgccaacaaa aggagaggga gccatagggt ctctagatta 120
cactoctgag gaaagaagag cacttgccaa aaaatcacia gatttctgt gtgaaggatg 180
tggctctgcc atgaaggatg tctgttgcc ttataatct ggaagcgatt caagccaagc 240
tgaccaagaa gccaagaac tggctaggca aataagctt aaggcagaag tcaattcatc 300
tggaaagact atctctgagt cagactaaa ccactcttt tcaactaactg atttacaaga 360
tgatatactt acaacattcc aggtgtctac ggccagtaca tctacggac tccagaattc 420
ctcagcagca tctttcatc aacctacca acctgtagct agaataacct ccatgagccc 480
tcgacagcgc cgggccagc agcagagtca gagaagggtg tctacttac cagatgtaat 540
ccagggccac cagccaagag acaaccacac tgatcatggt gggtcagctg tactgattgt 600
catctgact ttggcattgg cagctcttat attccgacga atatatctgg caaacgaata 660
catattgac ttgagttat aatatggtt tgtgacttat gagctgtgac tcaactgctt 720
cattaaacat tctgcattgg gtataatcta agaattgtt acaaaaagat tattttgtat 780
ttaccttca ttcttttt tgatcttgt aagtttagta taaatatatc tagacattca 840
gactgtgtct agcagttacg tctgtctaa aggactaga agtcaaagtt cctgtctca 900
ctattgac tgcttgacg ggaataact tgttttct catgttcat ctctttta 960
tgtaaattg taatacttc ctatattgcc cttgaaatt ttggataaa agatgatgtt 1020
ttaagtcca atgagtatta ctgttactc aataccactt attgagtact ctgttctac 1080
gtatgtagaa tgtataggga tagaagagt gaaaaggga agcaaaacti ctaagtggc 1140
ttcttaaaa tgcattcat aggagatgta ctggaattgc tcatctgtg actttattg 1200
tgtcctaaac atttctcagt gaaaataatt ttattcagt caaacattt tgaggaaatg 1260
agatcacatc ttgtcactg gatgctact gaagaggag tactttgta ccacttgat 1320
atgtgttat caccacccc tgcctctgc tgccataatc acacaaatt aaaaagaaag 1380
aaaacagtct tccatagatt ttaaggaag aaagggccca agtcaggaga tgccttggt 1440
ttctccaga agttaaattg ggggatctga agattgaat gtcggtctg cttgaaatg 1500
tatgtcttt gggaatgtat tatatgccta gctttataat caggtataaa atttaatta 1560
ttccaggaa tatgcataat attgaatatt tcatgtcta tttaataga aaacctcagg 1620
gccaagtaa ccagtgatag aagtagaaa aaccttta cttagaattg tccacctagt 1680

cagagcccaa gaaagaattt tcagtggaaa aatcaatata taacttagtg ctagctagcg 1740
 ccacagactc tagtagataa tattatcatc ataatggctg gtgaaacatc ataatacacag 1800
 aaaaacattg ccttcagcat gttcagttcg cagcactgag ggcactcttg aggggtgtgt 1860
 taatgaagat ttaatttta aatacaggtg gttccaagct tcaaatagg ttatgctcca 1920
 aaagtgttat tigtgaagta attttttac aagtcaaaca atgttgaag tggatttag 1980
 gttctagatc ggtccacgaa agttagccca tatgtatc ttgaatagta taggggaggg 2040
 tattcataaa gtccttatgt ggttttaact aagtgaatt atggacaaga gaaataattg 2100
 taaaatcgtc ttaaaggcaa attaatatt taccctgtt tatgggacat tegtctatt 2160
 aactgtcaga cacaattct gtttcatct gagagccagg ttctttat ttctacatc 2220
 aaaataagaa catattgtac actattatat aatacagaat tgtcttacac ttaataaat 2280
 tcgcatttta aagggttta caggattatt tttatatc ttagctgaat ttgttaaagt 2340
 ctaaaaagct caaggactt atgaagatc cattatatga ggaaaatcat aggttaccat 2400
 ttataactc tattgccata agaaaataca ctctaaaac ttgattgaa acatattaga 2460
 aacctgatt cagtgcicag tggctccta gtaagaagtc accgacggta gcgtcatatg 2520
 agaagaaaga aatccccacc acctcaacct ctgtgagat tgtgtgtag gaacagcctt 2580
 cctccgttt cccctcagtc aaactgagc cagcctctgg atcgatgta tctattgca 2640
 tgttccatg ggggttacct atacttaag ccaatcctgc tgcattcact gctaagtaa 2700
 ataaaaagcc aagaagaaaa aaaaaattt gcactgtgca gatccttgc tatctgactt 2760
 gcatctctc cccacctgt cagctagcca cctgctgtt tgtgtggga tatttttag 2820
 cacctgaagc accatctgaa aggggcacca t 2851

<210> 23
 <211> 3473
 <212> DNA
 <213> human

<400> 23
 aagagcagcg gcgaggcggc ggtggtggct gagtccgtgg tggcagaggc gaaggcgaca 60
 gctctagggg ttggcaccgg ccccgagagg aggatgcggg tccggatagg gctgacgctg 120
 ctgctgtgtg cgggtcgtct gagctggcc tcggcgtcct cggatgaaga aggcagccag 180
 gatgaatcct tagattcaa gactacttg acatcagatg agtcagtaaa ggaccacact 240

actgcaggca gagtagtgc tggtaaata ttcttgatt cagaagaatc tgaattagaa 300
 tcctctattc aagaagagga agacagcctc aagagccaag agggggagag tgtcacagaa 360
 gatatacgt ttctagatc tccaaatcca gaaaacaagg actatgaaga gccaaagaaa 420
 gtacggaaac cagctttgac cgccattgaa ggcacagcac atggggagcc ctgccacttc 480
 cctttcttt tcctagataa ggagtatgat gaatgtacat cagatgggag ggaagatggc 540
 agactgtggt gtgtacaac ctatgactac aaagcagatg aaaagtgggg cttttgtgaa 600
 actgaagaag aggctgctaa gagacggcag atgcaggaag cagaaatggt gtatcaaact 660
 ggaatgaaaa tcctaatgg aagcaataag aaaagccaaa aaagagaagc atatcggtat 720
 ctccaaaagg cagcaagcat gaaccatacc aaagccctgg agagagtgc atatgctctt 780
 ttatttggtg attactgcc acagaatac caggcagcga gagagatgt tgagaagctg 840
 actgaggaag gctctccaa gggacagact gctctggct ttctgtatgc ctctggactt 900
 ggtgttaatt caagtcaggc aaaggctctt gtatattata cattggagc tctgggggc 960
 aatctaatag ccacatggt ttgggttac agatactggg ctggcatcg cgctctccag 1020
 agttgtgaat ctgccctgac tcactatgt ctgttgcca atcatgtgc tagtgatc 1080
 tcgtaacag gaggtcagt agtacagaga atacggctgc ctgatgaagt ggaaaatcca 1140
 ggaatgaaca gtggaatgt agaagaagat ttgattcaat attaccagtt ctagctgaa 1200
 aaaggtagt tacaagcaca ggttggtctt ggacaactgc acctgcacgg agggcgtgga 1260
 gtagaacaga atcatcagag agcattgac tactcaatt tagcagcaaa tgctggcaat 1320
 tcacatgcca tggcctttt gggaaagatg tattcggaag gaagtacat tgtacctcag 1380
 agtaatgaga cagctctcca ctacttaag aaagctgctg acatgggcaa ccagttgga 1440
 cagagtgggc ttggaatggc ctacctat gggagaggag ttcaagttaa ttatgatc 1500
 gcccttaagt atttcagaa agctgctgaa caaggctggg tggatgggca gctacagctt 1560
 ggttccatgt actataatgg cattggagtc aagagagatt ataacaggc ctggaagtat 1620
 tttaattag ctctcaggg aggccatc ttggcttct ataacctagc tcagatgcat 1680
 gccagtggca ccggcgtgat gcatcatgt cactgcag tggagtgt taagaatga 1740
 tgtaacgag gccgtggc tgaaggctt atgactgcct ataacagcta taaagatggc 1800
 gattacaatg ctgcagtgt ccagtacct ctctggctg aacagggcta tgaagtggca 1860
 caaagcaatg cagccttat tctgtatcag agagaagcaa gcattgtagg tgagaatgaa 1920

acttatccca gagctttgct acattggaac agggccgcct ctcaaggcta tactgtggct 1980
 agaattaagc tcggagacta ccatttctat gggtttgga ccgatgtaga ttatgaaact 2040
 gcatttattc attaccgtct ggctctgag cagcaacaca gtgcacaagc tatgtttaat 2100
 ctgggatata tgcattgagaa aggactgggc attaaacagg atattcacct tgcgaaacgt 2160
 tttatgaca tggcagctga agccagccca gatgcacaag ttccagtctt cctagccctc 2220
 tgcaaattgg gcgtcgtcta ttcttgacg tacatacggg aaacaaacat tcgagatatg 2280
 ttacccaac ttgatatgga ccagctttg ggacctgagt gggacctta cctcatgacc 2340
 atcattgcgc tgctgttggg aacagtcata gcttacagc aaaggcagca ccaagacatg 2400
 cctgcaccca ggctccagg gccacggcca gctccacccc agcaggaggg gccaccagag 2460
 cagcagccac cacagtaata ggcactgggt ccagccttga tcagtacag cgaaggaagt 2520
 tatctgttg gaacacttgc atttgattta ggaccttga tcagtgtca cctccagaa 2580
 gaggcacggc acaaggaagc attgaattcc taaagctgct tagaatctga tgcctttatt 2640
 ttcagggata agtaactctt acctaaactg agctgaatgt ttgttcagt gccatatgga 2700
 ataacaactt tcagtggctt ttttttct ttctggaaa catatgtgag acactcagag 2760
 taatgtctac tgtatccagc tatcttctt ggatccttt ggtcattatt tcagtgtgca 2820
 taagtctta atgtcaacca tcttaaggt attgtgcatc gacactaaa actgatcagt 2880
 glaaaaagga aaaccaggt gcaagttta acgtgttca aagtctgaaa atagaactg 2940
 cctttaagt taaaaaaaa aaaagctatc ttgaaaatgt ttggaactg cgataactga 3000
 gaaactctta ccagtccaca tgcaattaga catattcagc atattgtta tttaaaagg 3060
 gaggggtggg aggtttctta ttggtgattg tcacacggta taccatactc ctctcctca 3120
 aagaatgaaa ggccttgta aggagtttt tgtgagctt acttcttgg aatggaatat 3180
 acttatgcaa aaccttgta actgactcct tgcactaacg cgagtttggc ccacctactc 3240
 tgtaattgc ttgttgtt tgaatataca gagccttgat ccagaagcca gaggatggac 3300
 taagtgggag aaattagaaa acaaacgaa ctctggttg ggtactacga tcacagacac 3360
 agacatactt ttctaaagt tgaagcattt gtcccagga ttattttac ttgcattc 3420
 ctttgcaca aagaacacat caccatttcc ttgtcaca agaacacac acc 3473

<210> 24

<211> 401

<212> DNA
<213> human

<220>
<221> misc_feature
<222> (252)..(252)
<223> any kind of base

<220>
<221> misc_feature
<222> (303)..(303)
<223> any kind of base

<220>
<221> misc_feature
<222> (390)..(390)
<223> any kind of base

<400> 24
ttagattatt ttcaatttat taticagaat aaatatact ttttcttta acttctcaa 60
tagttattga attgtattgg tttaaattaa atgcgtcatg tgtatataic agtattaatt 120
caagagatac aaaaggaaat tgagtgaataa ataagtctgc ctcttccca tcaactcat 180
gtctctacct agaggcaatt attgtcaaca gttttgatg tgtcttcaa aaaatagtc 240
attaagcctg gngtactaga tctcttttaa aagttacaa cctgttacag aatatata 300
aangttcaat tactagtaac acctattac agatacagat tacaacttag gaaatatatt 360
ttcatggacc attgatgtca ttggattcn ccctacaat c 401

<210> 25
<211> 1820
<212> DNA
<213> human

<400> 25
aatgtcttag aaaaaggctt tctaaaagaa aaagagcaag aggccatttc tttaagat 60
agatacaaag aacttcagga aaaacataaa caagaattgg aagacatgag gaaagctggt 120
cacgaagccc tcagcattat tgttgatgaa tataaggcac tactgcagtc tcagttaag 180
caacaagtag aagctattga aaaacagtac atttctgcaa ttgagaaaca ggcacacaag 240
tgtgaggagt tgctaaatgc tcagcatcag aggcctcttg aaatgctaga tacagagaag 300
gaactgttaa aagaaaaaat aaaggaagct ttgattcagc aatctcaaga acagaaggaa 360

atattggaaa agtgtttgga ggaagaaagg caaagaaata aagaggcatt agtatccgct 420
 gcaaagcttg aaaaagaagc agtgaaggat gcagtttaa aagtcgtaga agaagaaaga 480
 aaaaatttag aaaaagcgca tgctgaagaa agggaattat ggaagacaga acatgcaaaa 540
 gatcaagaaa aagtatctca ggaaattcaa aaagctatac aagaacaaag aaaaataagt 600
 caggaaactg ttaaggcagc aataatagaa gagcagaaac gaagtgaaaa ggctgtggaa 660
 gaggcagtga aaagaacaag agatgaattg atagagtata taaaagaaca gaaaaggctc 720
 gatcaagtca tccgccaaag aagcctgtcc agtttgaac tgttcctctc ctgtgcacag 780
 aaacagttaa gtgctttaat agctacggaa ccagttgaca ttgaataaaa agaacatgac 840
 aaacccacac tggcattgga taaatcatat tacaccttca aaatacacac tctgaattat 900
 aaagatgtgt tigtittctt tccaaatcat gtagaattga ttccagttc aaggataaac 960
 caaaacaata tttagaacta tcaagtgatc taatttattt tctttggtt tcttcttac 1020
 atttactgtt attttattat tattagtagt agcagcaaca gagtatgata tgacccaaaa 1080
 gccattgtaa agtgccacat taccaaaatt aattaagtaa actttatagc ctgtgggagt 1140
 ctattatata ttatttgcga aaagtagtaa atatattatt gttcatgat gactcttgat 1200
 gagatgctag aatgaacca tacatttate ttatttgag gatagaaata gcatggattt 1260
 caacatcact tatttatctg tataattgga aataaaacac cgatatgata gagaatcatt 1320
 ccggcattac ctaacctctt ctgcagttgg atctatgtat ttcatgtgt ctactgaaaa 1380
 cgaacaatac aattaaagc actaaagatt attatattaa ttcaacttgg atctgatata 1440
 tcacttaaac taaaggggtg tgtgtggtgt atgcttggtt cctatttctg ctctttaaag 1500
 atacttgaa tcaataaaac cattagtcta caaatcaaat tgtgaactta atctctagaa 1560
 agagaatata actcagccat ttataggaat ttaggttcaa gtacaggata tatgaaatct 1620
 ttcccagta ttccagaatg tacttaattc acaggcagga tgcttcaatg caaaatcatg 1680
 aatattttta attcaaaact aaaatgtcat taatatgtat gtatgcaaat gttttatctt 1740
 atttctgaa atgcatctac ttcatgggc ttgtacgtt tctgagattt ctcagtgtaa 1800
 taaaaagagc tcccaaaact 1820

<210> 26
 <211> 280
 <212> DNA
 <213> human

<220>
<221> misc_feature
<222> (261)..(261)
<223> any kind of base

<220>
<221> misc_feature
<222> (237)..(237)
<223> any kind of base

<400> 26
tcaagtcata agataaagtt taatcatttg atcatgttaa aagacacaaa acacagccaa 60
tctaaccaaa tttaggcat gcatttcat aaatatatta aattaagaaa agaaattgta 120
cacttaaacg tcttttcac ctagaatca ttaaatccac agatcaacaa taaaaccaat 180
tctctgcatt taccactca agatacaatt gtctatatt aaagataaca caaactncac 240
tagtctggtt aggaatttat ntgcattata catatattat 280

<210> 27
<211> 392
<212> DNA
<213> human

<400> 27
ttggttgaa atggcacccc aggacttgg gcctgcctta ctgtagacc tcgttcagt 60
agcaaagact tagtgagcag ctctgtatg ccaagtatt tgctaagctc tggaaaaaag 120
ataaacaaga catggttct gcttcaagg agtgtgtaat tcttagcca gatatggaaa 180
cctggaccct gagtgggaga aaggagacag atgaaaggag tccgtgatt tgaaccaag 240
agctgcctgc atggttatga gtatcactga ttttagggac gccacagag ctaaagcatt 300
ttttaatcc gagaagactt ttgtaactca tattagtaa tctttagct ctgagatagc 360
aacacagctc ttagaattct gtaagtaagc tt 392

<210> 28
<211> 2299
<212> DNA
<213> human

<400> 28
cgaaccccca cagctggagg gcgaggccag ctgtaccgg cccagtgcc ctttcgggc 60
cacaagcggc cgtcctctg gtccggtgct ccggcgctg atctagggtc atggagccgg 120

ggctgtggct cctttcggg ctacacgtga cctccgccgc aggattcgtg ccttgctccc 180
agtcctgggga tgcctggcagg cgcggcgtgt cccaggcccc cactgcagcc agatctgagg 240
gggactgtga agagactgtg gctggccctg gcgaggagac tgtggctggc cctggcgagg 300
ggactgtggc cccgacagca ctgcagggtc caagccctgg aagccctggg caggagcagg 360
cggccgaggg ggcccctgag caccaccgat ccaggcgctg cacgtgcttc acctacaagg 420
acaaggagtg tgcctactat tgccacctgg acatcattg gatcaacact cccgaacaga 480
cgggtgcccta tggactgtcc aactacagag gaagcttccg gggcaagagg tctgcggggc 540
cactccagg gaatctgcag ctctcacatc ggccacactt gcgctgcgt tgtgtggga 600
gatatgaaa ggctgcctg cacttttga cccaaactct ggacgtcagc agacagggtg 660
aagtcaagga ccaacaaagc aagcaggctt tagacctcca ccatcaaag ctcatgcccg 720
gcagtggact cgcctcgtc ccatctacct gcccccgctg cctcttcag gaaggagccc 780
cttaggagga caggcctgca gcctcctgt ctggggaggc ttctgtcatt gctcacacac 840
agttcagatt tccacctct tatagacaag aagtgaattt gcctggggca gaacaccac 900
ccaaagagtc cccactaac aatacccccc cccacggca agaatgcca aatccgaatg 960
accccagttt tctaatag taaaatgac ccagatgtc cccagagcat gacgcctgca 1020
gtcccggtt catgcaggaa attggtttg gagagtttg gcaagtggg aagccacta 1080
ctggctttg acatgacttc tcttgagaa taagtggact ccaagctaac tcttgcaa 1140
tgtaaacaca tgcctactt gtaataaatg caaatgccc gtgcagcaga agcatgcgac 1200
ttcatalcc ttgcctagaa taggtgcat ggtgtatgtc agtgagggcc acgaggcgtc 1260
ggcttagac acagatcata gctctacagg agtttatgaa ttgaagct atgggattt 1320
ggcagagaaa ttctcagctg tgcctgatac ccacaaaag aatgtatctc gaaagaatga 1380
aggaagaaga aaaaaggatc ctgatgtt gtgacaagaa aatgagaaag ttagtatctg 1440
caatacagag ctgttcctg ttcagtgaat gacctctgt attctgata gacaccaggc 1500
cgatacacag tggagtccc aggcctgtt tgcaggaagc cgactgtaaa gacagcccca 1560
gctcaaggct attaggtga atattgctt tcatgagtaa atgtggatct ttggggaatg 1620
gctcaaaaat aagtcacgaa cacaaattct ttgtaatta tgtaaatcc tgtttatata 1680
aattggcaac aactatacc gctgcagat tcaaaatctc tticagctgc gctcttcca 1740
ccgagccgag ctactgtga gtgtggagat gttatccac catgtaaagt cgctgcgca 1800

ggggagggct gcccatctcc ccaaccagc cacagagaga taggaaacgg catttgagtg 1860
 ggtgtccagg gcccctaga gagacattta agatggtgta tgacagagca ttggccttga 1920
 ccaaatgtta aatcctctgt gtgtatttca taagttatta caggataaaa agtcatgacc 1980
 tatcatgagg aaatgaaagt ggctgatttg ctggtaggat ttgtacagt ttagagaagc 2040
 gattatttat tgtgaaactg ttctccactc caactcctt atgtggatct gtcaaagta 2100
 gtactgtat atacgtatag agaggttagt aggtaggtag attttaaatt gcattctgaa 2160
 taaaactca tactccttag agctgaatt acattttta aatgcatatg tgctgtttgg 2220
 caccgtggca agatggatc agagagaaac ccatcaattg ctcaaatact cagaaagtac 2280
 tgtcaaaagc ctaataaaa 2299

<210> 29
 <211> 1339
 <212> DNA
 <213> human

<400> 29
 ctaaacaaaa tcattcactt ccctgattt galaagaaaa ttctgtaaa gctgttcct 60
 ctgcctctcc tctacgttgg aaaccacata agtggattat caagcacaag taaattaagc 120
 ctaccgatgt tcaccgtgct caggaaattc accattccac ttacctact tctgaaacc 180
 atcatacttg ggaagcagta ttactcaac atcatcctca gtgtctttgc cattattctc 240
 ggggctttca tagcagctgg gtctgacctt gcttttaact tagaaggcta tatttttga 300
 ttctgaatg atatcttcac atcagcaaat ggagtttata ccaaacagaa aatggaccca 360
 aaggagctag ggaaatacgg agtactttc tacaatgcct gcttcatgat tatcccaact 420
 cttattatta gtgtctccac tggagacctc caacaggcta ctgaattcaa ccaatggaag 480
 aatgttgtgt ttatctaca gtttctctt tctgtttt tggggttct gctgatgtac 540
 tccacggttc tgtgcagcta ttacaattca gccctgacga cagcagtggg ttgagccatc 600
 aagaatgtat ccgttccta cattgggata ttaatcggg gagactacat ttctctttg 660
 ttaaacttg tagggttaaa tattgcatg gcagggggct tgagatattc cttttaaca 720
 ctgagcagcc agttaaacc taaacctgtg ggtgaagaaa acatctgtt ggatttgaag 780
 agctaaagag tctgcagcag gattggagac tgactgtga ctgcgggctg ggggggcatt 840
 ccagtagga atgtgaagcc agaggtttcg gattcgtgac atccaccccc tgggcaagt 900

agagcatctg caaaatgcaa agagaactac ctcatatgca ggatgagcca atggcagtct 960
 caagaaatgt actcgggcga caccttacct gtggaaagca aatctttca aaataagcca 1020
 ctgggactcg gtaggtggag ccccagctgc tcttctaggg acctatgggg ccttcgtggc 1080
 atctctgtgc tgtgtctgg ggaggaggt gatgtaatgg tgactcttt ctgatcagca 1140
 ccttgccgt gaticccaag gtcccagcca aagcaaagg ccagttgtt cagtttaaac 1200
 agacatgtct ttagtctaataaaaattagtt aactgccagt aaagttatt gttagctttg 1260
 atgaaagcta tgttggtatc ttccctaata catcaaagta aataaaaaat catttctatg 1320
 taaaaaaaaa aaaaaaaaaa 1339

<210> 30
 <211> 4250
 <212> DNA
 <213> human

<400> 30
 gaacacatcg cggttgcac ccagaaagta gtcgccgga ctattcccc caaagagaca 60
 agcacacatg taggaatgac aaaggcttc gaaggagaga ggcagcccg cggcccgag 120
 agatccctc gataatggat tactaaatgg gatacacgct gtaccagtc gtcgagcc 180
 ccggccgct gtcgctgat gcaccgaaa ggtgaagta gagaaataaa gtctcccg 240
 tgaactacta tgaggtcaga agccttgctg ctatattca cactgctaca cttgctggg 300
 gctggttcc cagaagattc tgagccaatc agtatttcg atggcaacta tacaaaacag 360
 tatccggtgt ttgtgggcca caagccagga cggaacacca cacagaggca caggctggac 420
 atccagatga ttatgatcat gaacggaacc ctctacattg ctgctagga ccatattat 480
 actgttgata tagacacatc acacacggaa gaaatttatt gtagcaaaaa actgacatgg 540
 aaatctagac aggccgatgt agacacatgc agaataagg gaaaacataa ggatgagtgc 600
 cacaacttta ttaaagtct tctaaagaaa aacgatgatg cattgttgt ctgtggaact 660
 aatgcctta acccttctg cagaaactat aagatggata cattggaacc attcggggat 720
 gaattcagcg gaatggccag atgcccata gatgccaac atgccaacgt tgcactgtt 780
 gcagatggaa aactatactc agccacagt actgacttc ttgcatga cgcagtcatt 840
 taccggagtc ttggagaaag cctaccctg cggaccgtca agcacgattc aaaatggtg 900
 aaagaacat actttgtca agccgtggat tacggagatt atatctact ctcttcagg 960

gaaatagcag tggagtataa caccatggga aaggtagtt toccaagagt ggctcaggt 1020
tgtaagaatg atatgggagg atctcaaaga gtcttgaga aacagtggac gtgttcctg 1080
aaggcgcgct tgaactgctc agttcctgga gactctcatt ttatttcaa cattctccag 1140
gcagttacag atgtgattcg tatcaacggg cgtgatgtg tcctggcaac gttttctaca 1200
ccttataaca gcatccctgg gtctgcagtc tgtgcctatg acatgcttga cattgccagt 1260
gtttttactg ggagattcaa ggaacagaag tctcctgatt ccacctggac accagttcct 1320
gatgaacgag ttctaagcc caggccaggt tgctgtgctg gctcatcctc cttagaaaga 1380
tatgcaacct ccaatgagtt cctgatgat accctgaact tcatcaagac gcacccgctc 1440
atggatgagg cagtgcctc catcttcaac aggccatggt tcctgagaac aatggtcaga 1500
taccgcctta ccaaaattgc agtggacaca gctgctgggc catatcagaa tcacactgtg 1560
gttttctgg gatcagagaa gggaatcctc ttgaagttt tggccagaat aggaaatagt 1620
ggtttctaa atgacagcct ttcttgagg gagatgagtg ttacaactc tgaaaaatgc 1680
agctatgatg gagtcgaaga caaaaggatc atgggcatgc agctggacag agcaagcagc 1740
tctctgatg ttgcgttctc tacctgtgtg ataaagggtc cccttgccg gtgtgaacga 1800
catgggaagt gtaaaaaaac ctgtattgcc tccagagacc catattgttg atggataaag 1860
gaaggtggtg cctgcagcca ttatcaccc aacagcagac tgactttga gcaggacata 1920
gagcgtggca atacagatgg tctgggggac tgtcacaatt ccttgtggc actgaatgac 1980
atttcaactc cttaccaga taatgaaatg tottacaaca cagtgtatgg gcattccagt 2040
tccctctgc ccagcacaac cacatcagat tgcacggctc aagaggggta tgagtctagg 2100
ggaggaatgc tggactggaa gcatctgctt gactcacctg acagcacaga cccttgggg 2160
gcagtgtctt cccataatca ccaagacaag aaggagtgta ttcgggaaag ttacctcaa 2220
ggccacgacc agctgggtcc cgtcaccctc ttggccattg cagtcacctt ggcttctgc 2280
atgggggccc tcttctcggg catcacctc tactgcgtct gtgatcatc gcgcaaagac 2340
gtggctgtgg tgcagcgcaa ggagaaggag ctcacccact cgcgccgggg ctccatgagc 2400
agcgtacca agctcagcg cctcttggg gacactcaat ccaaagaccc aaagccggag 2460
gccatctca cgccactcat gcacaacggc aagctcgcca ctcccgcaa cacggccaag 2520
atgctcatta aagcagacca gcaccacctg gacctgacgg ccctccccac cccagagtca 2580
acccaacgc tgcagcagaa gcggaagccc agccgcggca gccgcgagt ggagaggaac 2640

cagaacctca tcaatgcctg cacaaaggac atgcccccca tgggctcccc tgtgattccc 2700
acggacctgc cctgcgggc ctccccagc cacatcccca gcgtgggtgt cctgcccatac 2760
acgcagcagg gctaccagca tgagtacgtg gaccagccca aaatgagcga ggtggcccag 2820
atggcgctgg aggaccaggc cgccacactg gagtataaga ccatcaagga acatctcagc 2880
agcaagagtc ccaacctatg ggtgaacctt gtggagaacc tggacagcct gcccccaaa 2940
gttcacagc gggaggcctc cctgggtccc ccgggagcct cctgtctca gaccggtcta 3000
agcaagcggc tggaaatgca ccactcctct tctacgggg ttgactataa gaggagctac 3060
cccagaaact cgctcacgag aagccaccag gccaccactc taaaagaaa caacactaac 3120
tctccaatt cctctacct ctccagaaac cagagcttg gcaggggaga caaccgccc 3180
ccgccccgc agagggtgga ctccatccag gtgcacagct ccagccatac tggccaggcc 3240
gtgactgtct cgaggcagcc cagcctcaac gcctacaact cactgacaag gtcggggctg 3300
aagcgtacgc cctcgtaaa gccggacgta cccccaaac catccttgc tccccttcc 3360
acatccatga agcccaatga tgcgtgtaca taatcccagg gggaggggtg cagggtcga 3420
accagcaggc aaggcgaggt gcccgctcag ctacgaagg ttctcaactg cctcgagtac 3480
ccaccagacc aagaaggcct gcggcagagc cgaggacgct ggttctctct ctctgggaca 3540
caggggtact cagcaaaact gggccgcgtg gtttggtgaa ggtttgcaac ggcggggact 3600
caccttcatt ctctctctc acttcccc acacctaca acaggtcgga cccacaaaag 3660
acttcagtta tcatcacaaa catgagccaa aagcacatac ctacccatac ccccccccc 3720
acacacacac acacatgcac acaacacata cacacacag cacagaggtg aacagaaact 3780
gaaacatttt gtccacaact tcacgggacg tggccagact gggtttgcgt tccaacctgc 3840
aaaacacaaa tacattttt aaaatcaaga aaatttaaaa agacaaaaaa aaaagaattc 3900
attgataatt ctaactcaga cttaacaat ggcagaagtt tactatgcgc aaatactgtg 3960
aatgccccgc cagtgttaca gcttctgtt gcagcagata aatgccatgt tgggcaacta 4020
tgtcatagat ttctgtcct cctctctttt aatgaaataa cgtgaccgtt aacgcaagta 4080
actctttatt tatgttcac cctttttc ctaaggaaa ggactcttc aaatatcacc 4140
ctatgaacag ctctcagaa agcccattga aagttaaact atttaacgtg aaatccatta 4200
actggaataa ttgatttct ttatttttac aataaattca ctgagtaaat 4250

<210> 31
<211> 2785
<212> DNA
<213> human

<400> 31
ctttagccca acagtcaaaa ataattgatg ctaccctaca aatgtccaaa actctagtat 60
atcatatttc taagttacag caaatattag tcttgctaaa ccaggagct ttggcaaaaa 120
tgtttttga cagtaaattt gtcttgatt atatattaac tagtcaaaga ggtgtttgta 180
acattattag agcttctgt ttaggtggg ttaacaccac caatcaagag gtcattctaa 240
cagaaagcct ggtacagaaa accatcaccc taaaaaaca tgccttcat atttaacaca 300
ctctgaaatc cagtcaaaa atgactaaag gcccttgcca tgactgatgt attctctgg 360
ccaacgccaa acaaattggga gcctgggtac gagtcagcct tcagggacti gtcacatttc 420
tacttggtt ctctctgtt attgcataa taaaatgtt tctatgctgt ttagtgcaac 480
ttaggcccta ttctgtagaa gtctcctcta ctattcaggc cactcaaaca ccccaaataa 540
ttgagttcaa aatcgacatc aagatataaa ggaatcagtg actaaatata ttcatatat 600
ggfattttta ttgattattg tgctgtcttg acctagtatg gaggccttgg ctgagggtg 660
gtcagtttcc tctctgagc agctgattaa atccacaccc caaccacttc ccttatcagg 720
ttctcacact ctggggccac tatgtaccca ctctaatac cacagggccca gacatcagac 780
aattaaggac agcgcccatg ccccaaagcc cgccaaaatt atgcaaatta ttcaaaatta 840
ttcaacctag ctaacccac ccttttgcgt gtacataagc tgccattcc cctccagcc 900
tgtggtaccc agtctcagg tgcaacccc tgcgtgtcc tctgtggcag ccttctctca 960
ttcagagctg ttccacag aggtagtga aagaactgga ttccaagt cacttgcaa 1020
gagaaaaaga aaactcagta gaagataatg gcaagtcag actggggata tgatgacaaa 1080
aatggtcctg aacaatggag caagctgtat ccattgccca atggaaataa ccaatcccct 1140
gttgatatta aaaccagtga aaccaaact gacacctctc tgaaacctat tagtgtctcc 1200
tacaaccag ccacagccaa agaaattatc aatgtggggc attcttcca tgtaaattt 1260
gaggacaacg ataaccgatc agtgctgaaa ggtggtcctt tcttgacag ctacaggctc 1320
ttcagttc atttcactg gggcagtaca aatgagcatg gtcagaaca tacagtggat 1380
ggagtcaaat atttgccga gcttcacgta gtcactgga attctgcaa gtactccagc 1440
cttgctgaag ctgcctcaaa ggctgatgtt ttggcagta ttggtgttt gatgaaggtt 1500

ggtgaggcca acccaaagct gcagaaagta ctgatgccc tccaagcaat taaaaccaag 1560
 ggcaaacgag ccccatcac aaatttgac ccctctactc tcttcttc atccctggat 1620
 ttctggacct accctggctc tctgactcat cctctcttt atgagagtgt aacttggatc 1680
 atctgtaagg agagcatcag tctcagctca gagcagctgg cacaattccg cagccttcta 1740
 tcaaagtgtg aaggtgataa cgctgtcccc atgcagcaca acaaccgccc aaccaacct 1800
 ctgaagggca gaacagttag agcttcatt tgaatattc gagaagaaac ttgtcttcc 1860
 tcaagaacac agccctgctt ctgacataat ccagttaaaa taataattt taagaaataa 1920
 atttattca atattagca gacagcatgc ctcaaatca atctgtaaaa ctaagaaact 1980
 taaatttag ttctactgc ttaattcaaa taataattag taagctagca aatagtaac 2040
 tgtaagcata agcttactt aaattcaagt ttgattgag gaattctta aaattacaac 2100
 taagtgattt gtatgtctat ttttctagt ttattgaac caataaaata atttatctc 2160
 ttctttctg ttgtgcattc agttctaaa accattaagt ttctactcca ttacattca 2220
 aaaatctaa atactttact tgcaagagta tttgtctca aatacaaaa cctaagagca 2280
 gctggagatg aaatattggg aaattcattt gcttactcct gaagacaaaa atagctga 2340
 gatgaccact ggatttaata tegtatgct ggccaacat tgctaccatt tgtgtgtct 2400
 gtgatcaaaa tgattactt ttatatagga agatgacgct tctggatatt gcttcactt 2460
 ctctcccca cgttagcaag gacaatgctt ctctgccatt attacaacta gttagttgc 2520
 atggagaatc ttactttaa aattggaaga aaagtcaca gtgaatggtt-tataaaaatg 2580
 ctaaagaagt cattctgct tagaatcata tagaaacatc atgcaatctt ttagtcagat 2640
 gtgcgctca ccttatgcta ttttatctt taattgacac acaataattg tacatgtta 2700
 tggagtatag tgtggtgtt tctgttgtt tgttgttt ttgagacaag gtctactct 2760
 gccagtcagg gtggagtgcg atggt 2785

<210> 32
 <211> 9588
 <212> DNA
 <213> human

<400> 32
 ccgaccaaca ccaacacca gctccgacgc agctcctctg cgcccttgcc gccctccgag 60
 ccacagcttt cctcccgtc ctgccccgg cccgtcgccg tctccgctc cgcagcggcc 120

tcgggagggc ccaggtagcg agcagcgacc tcgagagcct tccgcactcc cgcccggttc 180
ccgggcccgc cgctatcct tggccccctc cgctttctcc gcgcccggccc gcctcgctta 240
tgctcggcg ctgagccgct ctcccgattg cccgccgaca tgagctgcaa cggaggctcc 300
caccgcgga tcaacactct gggccgcatg atccgcgccc agtctggccc ggacctgcgc 360
tacgaggtga ccagcggcgg cgggggcacc agcaggatgt actattctcg gcgcggcgtg 420
atcaccgacc agaactcgga cggctactgt caaacgggca cgatgtccag gcaccagaac 480
cagaacacca tccaggagct gctgcagaac tgctccgact gcttgatgcg agcagagctc 540
atcgtgcagc ctgaattgaa gtatggagat ggaatacaac tgactcggag tcgagaattg 600
gatgagtgtt ttgccaggc caatgaccaa atggaaatcc tcgacagctt gatcagagag 660
atcgggcaga tgggccagcc ctgtgatgct taccagaaaa ggcttctca gctccaagag 720
caaatgcgag cctttataa agccatcagt gtccctcgag tccgcagggc cagctccaag 780
ggtggtggag gctacacttg tcagagtggc tcgggtggg atgagttcac caaacatgtc 840
accagtgaat gtttggggtg gatgaggcag caaaggcgcg agatggacat ggtggcctgg 900
ggtgtggacc tggcctcagt ggagcagcac attaacagcc accggggcat ccacaactcc 960
atcggcgact atcgctggca gctggacaaa atcaaagccg acctgcgca gaaatctcg 1020
atctaccagt tggaggagga gtatgaaaac ctgctgaaag cgtccttga gaggatggat 1080
cacctgcgac agctgcagaa catcattcag gccacgtcca gggagatcat gtggatcaat 1140
gactgcgagg aggaggagct gctgtacgac tggagcgaca agaacaccaa catcgctcag 1200
aaacaggagg ccttcccat acgcatgagt caactggaag ttaaagaaaa agagctcaat 1260
aagctgaaac aagaaagtga ccaactgtc ctcaatcagc atccagctc agacaaaatt 1320
gaggcctata tggacactct gcagacgag tggagtggg ttctcagat caccaagtgc 1380
attgatgttc atctgaaaga aaatgtgcc tacttcagt ttttgaaga ggcgcagtct 1440
actgaagcat acctgaagg gctccaggac tccatcagga agaagtacc ctgcgacaag 1500
aacatgcccc tgcagcacct gctggaacag atcaaggagc tggagaaaga acgagagaaa 1560
atccttgaat acaagcgtca ggtgcagaac ttgtaaaca agtctaagaa gattgtacag 1620
ctgaagcctc gtaaccaga ctacagaagc aataaaccca ttattctcag agctctctgt 1680
gactacaaac aagatcagaa aatcgtgcat aagggggatg agtgtatcct gaaggacaac 1740
aacgagcgca gcaagtggta cgtgacgggc ccgggaggcg ttgacatgct tgtccctct 1800

gtggggctga tcacccctcc tcgaaccca ctggccgtgg acctctctg caagattgag 1860
cagtactacg aagccatctt ggctctgtgg aaccagctct acatcaacat gaagagcctg 1920
gtgtcttggc actactgcat gattgacata gagaagatca gggccatgac aatcgccaag 1980
ctgaaaacaa tgcggcagga agattacatg aagacgatag ccgacctga gttacattac 2040
caagagtcca tcagaaatag ccaaggctca gagatgtttg gagatgatga caagcggaaa 2100
atacagcttc agttaccga tgcgcagaag cattaccaga ccttggtcat tcagctccct 2160
ggctatcccc agcaccagac agtgaccaca actgaaatca ctcatcatgg aacctgcca 2220
atgtcaacc ataataaagt aattgaaacc aacagagaaa atgacaagca agaaacatgg 2280
atgtgatgg agctgcagaa gattcgagg cagatagagc actgcgaggg caggatgact 2340
ctcaaaaacc tccctctagc agaccagggg tcttctacc acatcacagt gaaaattaac 2400
gagcttaaga gtgtgcagaa tgattcaca gcaattgctg aggttctca ccagcttaaa 2460
gatatgctg ccaactcag aggttctgaa aagtactgt atttacagaa tgaagtatt 2520
ggactatttc agaaactgga aaatatcaat ggtgttacag atggctactt aaatagctta 2580
tgcacagtaa gggcactgct ccaggctatt ctcaaacag aagacatgtt aaaggtttat 2640
gaagccaggc tctactgagga ggaaactgtc tgcttgacc tggataaagt ggaagcttac 2700
cgctgtggac tgaagaaaat aaaaaatgac ttgaactga agaagtcgtt gttggccact 2760
atgaagacag aactacagaa agcccagcag atccactctc agactcaca gcagtatcca 2820
ctttatgac tggactggg caagtccgt gaaaaagtca cacagctgac agaccgctgg 2880
caaaggatag ataaacagat cgactttaga ttatgggacc tggagaaaca aatcaagcaa 2940
ttaggaatt atcgtgataa ctatcaggct ttctgcaagt ggctctatga tcgtaaacgc 3000
cgccaggatt ccttagaatc catgaaattt ggagattcca acacagtcac gcggttttg 3060
aatgagcaga agaactgca cagtgaataa tctggcaaac gagacaaatc agaggaagta 3120
caaaaaattg ctgaacttg cgcaattca attaaggatt atgagctcca gctggcctca 3180
tacacctcag gactggaac tctgtgaac atacctatca agaggacat gattcagtc 3240
cctctgggg tgattctga agaggctga gatgtcatg ctcggtacat tgaactactt 3300
acaagatctg gagactatta caggttcta agtgagatgc tgaagagttt ggaagatctg 3360
aagctgaaaa ataccaagat cgaagtttg gaagaggagc tcagactggc ccgagatgcc 3420
aactcgaaa actgtaataa gaacaaatc ctggatcaga acctgcagaa ataccaggca 3480

gagtgttccc agttcaaagc gaagcttgcg agcctggagg agctgaagag acaggctgag 3540
ctggatggga agtcggctaa gcaaaatcta gacaagtgt acggccaaat aaaagaactc 3600
aatgagaaga tcacccgact gacttatgag attgaagatg aaaagagaag aagaaaatct 3660
gtggaagaca gatttgacca acagaagaat gactatgacc aactgcagaa agcaaggcaa 3720
tgtgaaaagg agaaccttgg ttggcagaaa ttagagtctg agaaagccat caaggagaag 3780
gagtacgaga ttgaaaggtt gagggttcta ctgcaggaag aaggcaccgc gaagagagaa 3840
tatgaaaatg agctggcaaa ggtaagaaac cactataatg aggagatgag taatttaagg 3900
aacaagtatg aaacagagat taacattacg aagaccacca tcaaggagat atccatgcaa 3960
aaagaggatg attccaaaaa tottagaaac cagcttgata gactticaag ggaaaatcga 4020
gatctgaagg atgaaattgt caggctcaat gacagcatct tgcaggccac tgagcagcga 4080
aggcgagctg aagaaaacgc ccttcagcaa aaggcctgtg gctctgagat aatgcagaag 4140
aagcagcatc tggagataga actgaagcag gtcatgcagc agcgctctga ggacaatgcc 4200
cggcacaagc agtccttga ggaggctgcc aagaccattc aggacaaaaa taaggagatc 4260
gagagactca aagctgagtt tcaggaggag gccaaagcgc gctgggaata tgaaaatgaa 4320
ctgagtaagg taagaaacaa ttatgatgag gatatcatta gcttaaaaaa tcagtttgag 4380
accgagatca acatcaccaa gaccaccatc caccagctca ccatgcagaa ggaagaggat 4440
accagtggct accgggcctc gatagacaat ctaccccgag aaaacaggag ctatctgaa 4500
gaaataaaga ggctgaagaa cactctaacc cagaccacag agaattctag gaggttgaa 4560
gaagacatcc aacagcaaaa ggccactggc tctgaggtgt ctgagaggaa acagcagctg 4620
gaggttgagc tgagacaagt cactcagatg cgaacagagg agagcgtaag atataagcaa 4680
tctcttgatg atgctgcaa aacatccag gataaaaaca aggagataga aaggttaaaa 4740
caactgatcg acaaagaaac aaatgaccgg aaatgcctgg aagatgaaaa cgcgagatta 4800
caaaggttcc agtatgacct gcagaaagca aacagtagtg cgacggagac aataaacaaa 4860
ctgaaggttc aggagcaaga actgacacgc ctgaggatcg actatgaaag ggtttcccag 4920
gagaggactg tgaaggacca ggatatcacg cggttccaga actctctgaa agagctgcag 4980
ctgcagaagc agaaggttga agaggagctg aatcggtga agaggaccgc gtcagaagac 5040
tcctgcaaga ggaagaagct ggaggaagag ctggaaggca tgaggaggtc gctgaaggag 5100
caagccatca aaatcaccaa cctgaccag cagctggagc aggcattcat tgttaagaag 5160

aggagtgagg atgacctccg gcagcagagg gacgtgctgg atggccacct gagggaaaag 5220
cagaggaccc aggaagagct gaggaggctc tctctgagg tcgaggccct gaggcggcag 5280
ttactccagg aacaggaaag tgtcaaaca gctcacttga ggaatgagca ttccagaag 5340
gcgatagaag ataaaagcag aagcttaaat gaaagcaaaa tagaaattga gaggctgcag 5400
tctctcacag agaacctgac caaggagcac ttgatgttag aagaagaact gcggaacctg 5460
aggctggagt acgatgacct gaggagagga cgaagcgaag cggacagtga taaaaatgca 5520
accatcttgg aactaaggag ccagctgcag atcagcaaca accggaccct ggaactgcag 5580
gggctgatta atgattaca gagagagagg gaaaatttga gacaggaaat tgagaaattc 5640
caaaagcagg ctttagaggc atctaattag attcaggaat caaagaatca gtgtactcag 5700
gtgttacagg aaagagagag cttctgggtg aaaatcaaag tcctggagca agacaaggca 5760
aggctgcaga ggctggagga tgagctgaat cgtgcaaaat caactctaga ggcagaaacc 5820
agggtgaaac agcgcttga gtgtgagaaa cagcaaattc agaatgacct gaatcagtgg 5880
aagactcaat attcccga ggaggaggct attaggaaga tagaatcgga aagagaaaag 5940
agtgagagag agaagaacag tcttaggagt gagatcgaaa gactccaagc agagatcaag 6000
agaattgaag agagggtcag gcgtaagctg gaggattcta ccaggagac acagtcacag 6060
ttagaaacag aacgctccc atatcagagg gagattgata aactcagaca gcgccatat 6120
gggtcccatc gagagacca gactgagtg gtgtggaccg ttgacacctc caagctgggtg 6180
ttgatgggc tgaggaagaa ggtgacagca atgcagctct atgagtgtca gctgatcgac 6240
aaaacaacct tggacaaact attgaagggg aagaagtcag tggaagaagt tgcttctgaa 6300
atccagccat tcctcgggg tgcaggatct atcgctggag catctgcttc tctaaggaa 6360
aaatactctt tggtagaggc caagagaaag aaattaatca gccagaatc cacagtcatg 6420
cttctggagg ccagggcagc tacaggtggt ataattgac cccatcgga tgagaagctg 6480
actgtcgaca gtgccatagc tcgggacctc attgactcg atgaccgtca gcagatatat 6540
gcagcagaaa aagctatcac tggttttag gatccattt caggcaagac agtatctgtt 6600
tcagaagcca tcaagaaaaa ttgattgat agagaaaccg gaatgcgcct gctggaagcc 6660
cagattgctt caggggggtg agtagacct gtgaacagt tcttttgcc aaaagatgtc 6720
gccttgccc gggggctgat tgatagagat ttgatcgat ccctgaatga tccccgagat 6780
agtcagaaaa actttgtgga tccagtcacc aaaaagaagg tcagttacgt gcagctgaag 6840

gaacggtgca gaatcgaacc acatactggt ctgctcttgc ttacagtaca gaagagaagc 6900
atgtccctcc aaggaatcag acaacctgtg accgtcactg agctagtaga ttctgtata 6960
ttgagaccgt ccactgtcaa tgaactggaa tctggtcaga ttcttatga cgaggttggt 7020
gagagaatta aggacttcct ccagggttca agctgcatag caggcatata caatgagacc 7080
acaaaacaga agcttggcat ttatgaggcc atgaaaattg gcttagtccg acctggtact 7140
gctctggagt tgctggaagc ccaagcagct actggcttta tagtgatcc ttttagcaac 7200
ttgaggttac cagtggagga agcctacaag agaggcttgg tgggcattga gttcaaagag 7260
aagctcctgt ctgcagaacg agctgtcact gggataatg atcctgaaac aggaacatc 7320
atctcttgt tccaagccat gaataaggaa ctcatogaaa agggccacgg tattcgctta 7380
ttagaagcac agatcgcaac cggggggatc attgacccaa aggagagcca tcgtttacca 7440
gttgacatag catataagag gggctatttc aatgaggaac tcagtgagat tctctcagat 7500
ccaagtgatg ataccaaagg atttttgac cccaacactg aagaaaatct tacctatctg 7560
caactaaaag aaagatgcat taaggatgag gaaacagggc tctgtcttct gcctctgaaa 7620
gaaaagaaga aacagggtgca gacatcacia aagaataccc tcaggaagcg tagagtggtc 7680
atagttgacc cagaaaccaa taaagaaatg tctgttcagg aggcctacia gaagggccta 7740
attgattatg aaaccttcaa agaactgtgt gacgaggaat gtgaatggga agaaataacc 7800
atcacgggat cagatggctc caccaggggtg gtcttggtag atagaaagac aggcagtcag 7860
tatgatattc aagatgctat tgacaagggc ctgttgaca ggaagttctt tgatcagtac 7920
cgatccggca gcctcagcct cactcaattt gctgacatga tctcctgaa aaatggtgtc 7980
ggcaccagca gcagcatggg cagtgggtgc agcgatgatg ttttagcag ctcccgacat 8040
gaatcagtaa gtaagattc caccatatcc agcgtcagga atttaacct aaggagcagc 8100
tcttttcag acacctgga agaatcgagc ccattgcag ccatcttga cacagaaaac 8160
ctggagaaaa tctccattac agaaggtata gagcggggca tcgttgacag catcacgggt 8220
cagaggcttc tggaggctca ggctgcaca ggtggcatca tccaccaac cacgggccag 8280
aagctgtcac ttcaggacgc agtctcccag ggtgtgattg accaagacat ggccaccagc 8340
gtgaagcctg ctgagaaagc ctctataggc ttgaggggtg tgaagggaaa gaagaagatg 8400
tcagcagcag aggcagtga agaaaaatgg ctcccgatg aggctggcca gcgcttctg 8460
gagttccagt acctcacggg aggtcttgtt gacccggaag tgcattggag gataagcacc 8520

gaagaagcca tccggaaggg gttcatagat ggccgcgccc cacagaggct gcaagacacc 8580
 agcagctatg ccaaaatcct gacctgcccc aaaaccaaataaaaaatc ctataaggat 8640
 gccataaatc gctccatggg agaagatatc actgggctgc gcctctgga agccgcctcc 8700
 gtgtcgtcca agggcttacc cagcccttac aacatgtctt cggctccggg gtcccgtcc 8760
 ggctcccgct cgggatctcg ctccggatct cgtccgggt cccgcagtgg gtcccgaga 8820
 ggaagcttgg acgccacagg gaattctcc tactcttatt cctactcatt tagcagtagt 8880
 tctatgggc actagtagtc agttgggagt ggtgctata cctgacttc atttatga 8940
 attccactt tattaataa tagaaaagaa aatccccgtg ctgcagtag agtgatagga 9000
 cattctatgc ttacagaaaa tatagccatg attgaaatca aatagtaaag gctgtctgg 9060
 cttttatct tctagctca tctaaataa gcagtacact tggatgcagt gcgtctgaag 9120
 tgctaatacag ttgtaacaat agcacaaatc gaacttagga ttgttctt ctctctgtg 9180
 ttctgattt tgatcaattc ttaatttg gaagcctata atacagttt ctattctgg 9240
 agataaaaat taaatggatc actgatatt tagtcattct gctctcacc taaatattc 9300
 catattctgt attaggagaa aattaccctc ccagcaccag ccccccctc aaaccccaa 9360
 cccaaaacca agcatttgg aatgagctc cttagttc agagtgtgga ttgtataacc 9420
 catatactct tcgatgtact tgttgggtt ggtattaatt tgactgtgca tgacagcggc 9480
 aatctttct ttgtcaaag tttctgtt atttggctg tcatattcga tgtacttaa 9540
 ggtgtctta tgaagttgc tattctggca ataaacttt agacttt 9588

<210> 33
 <211> 366
 <212> DNA
 <213> human

<220>
 <221> misc_feature
 <222> (351)..(351)
 <223> any kind of base

<400> 33
 gaagtgccat ttatattat acaaaaatat tacataattc agttagtatt ggtgacataa 60
 tttagttagt atgggtgata taatggatc aatttttagc atctaataaa gatctttta 120
 tgagtcccat ataaaatatg tgaacaaagc aatctgtca taagattgt gatgatttag 180

gagaaagtac tttagataa ttttttctg tctcttctg aactctctca acagtagttc 240
tcttagatt agagccagca ggtcggccat aacagtttc ttcaaattg ggcaacagtt 300
atacaaatgc ttgaattca agacaacata ttaaagggc tatgaactgg naatctaacc 360
tgggtt 366

<210> 34
<211> 1466
<212> DNA
<213> human

<400> 34
agccccaagc ttaccacctg caccgggaga gctgtgtgc accatgtggg tcccggttgt 60
cttctcacc ctgtccgtga cgtggattgg tgctgcaccc ctcatctgt ctggattgt 120
gggaggctgg gagtgcgaga agcattccca accctggcag gtgcttggg cctctcgtgg 180
cagggcagtc tggggcggg ttctgggtga ccccagtggt gtcctcacag ctgccactg 240
catcaggaac aaaagcgtga tcttgctggg tcggcacagc ctgtttcctc ctgaagacac 300
aggccaggta ttacaggta gccacagctt cccacacccg ctctacgata tgagcctcct 360
gaagaatcga ttctcaggc caggtgatga ctccagccac gacctcatgc tgctccgct 420
gtcagagcct gccagctca cggatgctgt gaaggctatg gacctgcca cccaggagcc 480
agcactgggg accacctgct acgcctcagg ctggggcagc attgaaccag aggagtctt 540
gaccccaaag aaacttcagt gtgtggacct ccatgttatt tccaatgacg tgtgtgcgca 600
agttcacctc cagaaggtga ccaagttcat gctgtgtgct ggacgctgga cagggggcaa 660
aagcacctgc tcgggtgatt ctggggggcc acttgctgt aatgggtgct ttcaaggat 720
cacgtcatgg ggcagtgaac catgtgccct gcccgaaagg ccttccctgt acaccaaggt 780
ggtgcattac cggaagtga tcaaggacac catcgtggcc aaccctgag caccctatc 840
aaccctctat ttagtaaac ttggaacctt ggaaatgacc agccaagac tcaagcctcc 900
ccagttctac tgaccttct ccttaggtgt gaggtccagg gttgctagga aaagaaatca 960
gcagacacag gttagacca gagtgttct taaatggtgt aatttctcc tctctgtgc 1020
ctggggaata ctggccatgc ctggagacat atcactcaat ttctctgagg acacagatag 1080
gatggggtgt ctgtgttatt tgtgggttac agagatgaaa gaggggtggg atccacactg 1140
agagagtga gagtgacatg tgctggacac tgcctatgaa gcactgagca gaagctggag 1200

gcacaacgca ccagacacac acagcaagga tggagctgaa aacataaccc actctgtcct 1260
ggaggcactg ggaagcctag agaaggctgt gagccaagga gggagggctt tccttggca 1320
tgggatgggg atgaagtaag gagagggact ggacccctg gaagctgatt cactatgggg 1380
ggagggtgat tgaagtcctc cagacaaccc tcagatttga tgatttcta gtagaactca 1440
cagaaataaa gagctgttat actgtg 1466

<210> 35
<211> 187
<212> DNA
<213> human

<400> 35
gatctggcgc attccggctg acactctcgt ttatttggac tgtaagctg acctctatga 60
ataattactt cagccccga gtgtcccg gccaaagctc ttggccaaac ttacaccta 120
gcttctgata agtctgggc caagctaagc agcatctatc aatcatccct tcagctcctg 180
attgatc 187

<210> 36
<211> 2913
<212> DNA
<213> human

<400> 36
actgggtacc gaggactggg tgtgtttaag gcagacagcc aggtgaggat ccagctact 60
ggggcctgct gtcactcct gggagtacc gggggcagg agcctagggg actctgcac 120
ttacatcca gccatgctaa ttacacittt tggcaaagga aacagctagg agcagttct 180
ttactccta cagccccgt ttctcagtg ttacacctg aattattact gggctagagg 240
gaaggcagcc tctgaagtgt ggcaggagga ggggaagtct gcctgcatct tgggtgtct 300
gtcagatgcc agcactaata acctggctt tgtaggcct gtcagtgtc tcaggaatga 360
aaggggaccc ctgagaggtg ctcatgacca gcaggctgtg aatgctctt acccaccacc 420
ctcactcct cgttaaagat ggtgctacct gccacacagc agacatctgg tcgtgcaca 480
cccgaagac cccaaggcag tctgccctt gtccagccac acgccagcac ccacctcct 540
ggccctgcc tcggcctccc cagaccagct gcaccagcc cccaacacgc acccctctc 600
cagatgtgtg cagggcctca tttgcagag caaagacaga tgttcagcc acacgttta 660
ttaacttcta aaacctgtgc tcaggacact cttaacagt catgaaaagt ttgatcact 720

gccacagtca ggaccttgt gtggggctct gatctgatgt tcggtctcat catctccaa 780
accagcagtc gttgtaccc caacctctg ctacggggt cataccccca aatgatttc 840
ctgattatg tattcccta caaagggtt tctataccta gcatctgct ccagcatgag 900
aagggggaat aggtgagacc cattgccag tagcagacgg ggacctggg gagaaaatgg 960
cagagcctgt tggagactcc ctgtccag ctgaccagcc aatgggatt ctctccctc 1020
cactgtctcc cacaagtag aagaatctg gtacattag ccatgagcc tggcacagat 1080
ccctatctag acataggcc cttagacat gacttgga ttgaccagcc tgttgcaat 1140
gggtcgggga ggcagaggg atgtcacac cagtaattct catcccga atgttgga 1200
tcacctggg agagtcaca aaatactgt gcaggggtcc cacctctgat gatctgagt 1260
gggtggctg ggtgtggcc caggcatcat gatgttcag gccccaggt gactcttag 1320
gcagcccagc taagccccta gaccttga attcccca aatgacctca gagggcccga 1380
ttgagggaa atgcctaact tcaggggccc taagaatccc ccaggagca tgtgaaatgc 1440
agataccagg cccaccccca gagatgagct gaggtgggtc aggggtgaag tgcagggatc 1500
agtgttttc acaagctcca tactccagg aaatggtgt gtggtgggc ccgtagaaaa 1560
cattctgaga gtctgttc ctgtccctg gtgcacgtg ggtggaatcc cagtggcct 1620
gccttgagga ggatgtcat taacgtgta ggggagacag agacagctcc acctgcccc 1680
tgtcccaccg gggacctca aaaactcat ggatgttaga gcaagcagcc atgtgcagc 1740
agaggatgag gctggcgat ttagtaagag cctctgtgt tgggtgag ttcttctct 1800
agtgccctg tcatctggcc tctggataac ccacctctc tccctatcc taaaattaca 1860
gatggcgaaa gacggccaca ttagtgaga cccctaagg cctccaacta ggtgggtcc 1920
acagtggccc ctggtgatg gaccacacac tctctccct cctctggctc aggactacgg 1980
tgtgaaatta gggagatatg aatgtcttc ttgaaaact ctctcccag tctcccact 2040
ttgctgggg gtcttggtc aaggccagct ttgactagg gctgttgcg actaccagct 2100
gtctcattt gctgtactgc aaactcaggc ttggtccaa gcttatggg gccctgtct 2160
tcccctagta ggtgtgtt tgggtcaca tctgtcata ccttcagag agcttctcc 2220
cagcctctac atcagggaga gaggtagga gggaggagca ttaaggatt agaagaagga 2280
ctaaagtaca acagccttg aggaactgcc aggaactaag ggcgagcact ggagaaggca 2340
acctgggacc cctgcgctt ctgagcagga agaccaagac ctcaggggc cctaagcact 2400

gaaaacatca ttctcatcc ccaagccctg gcatccccct gttcttctaa aataattctt 2460
 ttctaggtaa ttctgattgc aaaattctgg atgggttcat ccaagctgac ctttgcctgt 2520
 tttcccttc ccaacaaggc ctactttt ggagccacct tagttggtgc ctaggcagag 2580
 gggcagtcag cagtgggtat caggatcctg gctctatggg ttgccttct cctggtctgt 2640
 aaagccctcg caggcagga ctcttagat agctgcttc ttagggcatg gcatgtggtg 2700
 ggtgggtaat gaatggaaga gaggaatga gtgatcaagg gagggaggag ggagtggagt 2760
 ggagatttct catccttcc tgtaattta tgacatctc ctgcctatga gtccttgact 2820
 ctggagttt acaaagcagt cacatttcaa ataaaagtct gggaaagcaa cacatcatcg 2880
 ccaacttta atttgctaa ataaggatat tag 2913

<210> 37
 <211> 1466
 <212> DNA
 <213> human

<400> 37
 agccccaagc ttaccacctg caccggaga gctgtgtgc accatgtggg tcccgggtgt 60
 ctctcacc ctgtccgtga cgtggattgg tgctgaccc ctcatctgt ctggattgt 120
 gggaggctgg gactgagaga agcattccca accctggcag gtgcttggg cctctcgtgg 180
 cagggcagtc tggggcgtg ttctggtgca ccccatgtg gtctcacag ctgccactg 240
 catcaggaac aaaagcgtga tctgtctggg tggcacagc ctgttcatc ctgaagacac 300
 aggccaggta ttacaggta gccacagctt cccacaccg ctctacgata tgagctcct 360
 gaagaatga ttctcaggc cagggtatga ctccagccac gacctcatg tgctccgct 420
 gtcagagcct gccagctca cggatgctgt gaaggatg gacctgcca cccaggagcc 480
 agcactgggg accacctgct acgctcagg ctggggcagc attgaaccag aggagttct 540
 gaccccaaag aaactcagt gtgtggacct ccatgttatt tccaatgacg tgtgtgcga 600
 agttcacctt cagaaggta ccaagttcat gctgtgtgt ggacgtgga cagggggcaa 660
 aagcacctgc tcgggtgatt ctgggggcc actgtctgt aatggtgtg tcaaggtat 720
 cacgtcatgg ggcagtgaac catgtgcct gccgaaagg ccttccctgt acaccaagg 780
 ggtgcattac cggaagtga tcaaggacac catcgtggc aaccctgag caccctatc 840
 aacccttat ttagtaaac ttggaacct ggaaatgacc agccaagac tcaagcctc 900

ccagttctac tgaccttgt ccttaggtgt gaggtccagg gttgctagga aaagaaatca 960
 gcagacacag gtgtagacca gagtgttct taaatggtgt aatttgtcc tctctgtgc 1020
 ctggggaata ctggccatgc ctggagacat atcactcaat ttctctgagg acacagatag 1080
 gatggggtgt ctgtgttatt tgtgggttac agagatgaaa gaggggtggg atccacactg 1140
 agagagtgga gagtgacatg tgctggacac tgtccatgaa gcactgagca gaagctggag 1200
 gcacaacgca ccagacactc acagcaagga tggagctgaa aacataaccc actctgtcct 1260
 ggaggcactg ggaagcctag agaaggctgt gagccaagga gggaggggtct tccttggca 1320
 tgggatgggg atgaagtaag gagagggact ggacccctg gaagctgatt cactatgggg 1380
 ggagggtgat tgaagtctc cagacaaccc tcagattga tgatttcta gtagaactca 1440
 cagaaataaa gagctgttat actgtg 1466

<210> 38
 <211> 462
 <212> DNA
 <213> human

<220>
 <221> misc_feature
 <222> (197)..(197)
 <223> any kind of base

<220>
 <221> misc_feature
 <222> (116)..(116)
 <223> any kind of base

<220>
 <221> misc_feature
 <222> (334)..(334)
 <223> any kind of base

<220>
 <221> misc_feature
 <222> (402)..(402)
 <223> any kind of base

<220>
 <221> misc_feature
 <222> (429)..(429)
 <223> any kind of base

<220>
<221> misc_feature
<222> (438)..(438)
<223> any kind of base

<220>
<221> misc_feature
<222> (443)..(443)
<223> any kind of base

<220>
<221> misc_feature
<222> (459)..(459)
<223> any kind of base

<400> 38
taaggtttta taattatttt tatttttctt ttctttttt ttatggctt ggatgacact 60

ttatttcag atccaatact agaagttgtt tccatgtca cattttcctt cctggnntaa 120

aaaaaagagt tgtattttt tttttgctt tttttaaatt atactttaag ttttagggta 180

catgtgcaca acgcagnggt tagctacata tgtatacatg tgccatgttg gcgtgctgca 240

tccagtaact cgtcatttaa cattaggtat atctccaaat gctatccttc cccccattgt 300

attttcata gcttaaaaaat cattgacata ggantaattc caactaaagt acggtattaa 360

atccctgggg gaataaattt tgtcttaaca agggtaaggt tngtgaaaag gatggttttg 420

tcacagggna aaaggganat ccnccattt taaaaccnc ct 462

<210> 39
<211> 1490
<212> DNA
<213> human

<400> 39
ctcgtgcccc ccacggaggg gactgctctc ccccgtgca tcctttctgt gaggtacctt 60

accacctca gcacctgaga gggtgaaata gaattctaac ctcgacattc gggaagtgtt 120

tttgagaagt ctcggtcggt aagggaagtc ttccaagtcc gtgcagcact aacgtattgg 180

cacctgcctc ctcttcggcc acccccaga tgaggcagct gtgactgtgt caagggaagc 240

cacgactctg accatagtct tctctcagct tccactgccg tctccacagg aaaccagaa 300

gttctgtgaa caagtcacatg ctgccatcaa ggcatttatt gcagtgtact attgcttcc 360

aaaggatcag gggatcaccc tgagaaagct ggtacggggc gccaccctgg acatcgtgga 420

tggcatggct cagctcatgg aagtacttcc cgtcactcca actcagagcc ctgagaacaa 480
 tgaccttatt tctacaaca gtgtctgggt tgcgtgccag cagatgcctc agataccaag 540
 agataacaaa gctgcagctc tttgatgct gaccaagaat gtggattttg tgaaggatgc 600
 acatgaagaa atggagcagg ctgtggaaga atgtgaccct tactctggcc tcttgaatga 660
 tactgaggag aacaactctg acaaccacaa tcatgaggat gatgtgttg gggttccag 720
 caatcaggac ttgtattgt cagaggacga tcaagagctc ataatccat gccttgcgct 780
 ggtgagagca tcaaagcct gcctgaagaa aattcggatg ttagtggcag agaatgggaa 840
 gaaggatcag gtggcacagc tggatgacat tgtggatatt tctgatgaaa tcagccctag 900
 tgtggatgat ttggctctga gcatatatcc acctatgtgt cacctgaccg tgcgaaatcaa 960
 ttctgcgaaa ctgtatctg ttttaaagaa ggcactgaa attacaaaag caagtcattg 1020
 gaccctcag ccagaagata gttggatccc ttacttatt aatgccattg atcattgcat 1080
 gaatagaatc aaggagctca ctcagagtga actgaatta tgactttca ggctcattg 1140
 tactctcttc cctctcctc gtcattgta ggctctgata cctgcttta aaatggagct 1200
 agaatgcttg ctggatgaa agggagtgc tatctatatt tagcaagaga cactattacc 1260
 aaagattgtt ggtaggcca gattgacacc tattataaa ccatacgct atattttct 1320
 gtgctatata tgaaaaataa ttgatgatt tctcattcct gactcattc tcagagattc 1380
 ctaggaaagc tgccttattc tcttttgca gtaaagatg ttgtttcat tgtaaagatg 1440
 ttgatgtct caataaaatg ctaactgcc agtgattaaa aaaaaaaaaa 1490

<210> 40
 <211> 1677
 <212> DNA
 <213> human

<400> 40
 ctgacccta ttatagtggt ctctaaaggt ggtgttatta tgtttctag agcacttca 60
 ttatacaaac gtcaaggaat ccgagtaat gtgcttgcc ctgagttgt tgaaacagac 120
 atgggcacaa tgatcgggcc caaattcctt agtatgatgg ggggctttgt acctatggaa 180
 atggtggtga aagggtcttt tgagctcctc actgatgaga ataaagccgg cgattgccta 240
 tggattacta atcggcgagg tcttgagtac tggcccacc cactcagaaga agcaaaglac 300
 ttgctgcgtt ctacacgttc caggagaaga actgaatata aagctccacc aattaaacta 360

ccgagagatt ttgagaaaat agttgttcag accttgactc acaacttcg gaatgctacc 420
 agtglagtaa gagcaccact gagattacct atcaaaccac actatgttct tgtgaagata 480
 atctatgctg gtgtaaatgc tagtgatgta aattttagct caggctcgcta ttttggtggc 540
 aataacagtg acactgcac cgccttcg tttgatgcag gatttgaggc tgtgggagta 600
 attgcagcag ttggggattc tgttactgac ttgaaagttg gcatgccttg tgcgttcacg 660
 acttttggag gctatgctga atttacaatg attccttcga aatacgccct tccaatgcct 720
 agaccagaac cggaaggtgt tgccatgctt acatcaggat taacagcttc aattgctcta 780
 gaaaaggcag gacagatgga atctggaaaa gtggctcttg ttactgctgc ggcaggagga 840
 actggtcagt ttgctgttca gcttgcaaaa ttactgtgta ataccgtgtt tgccacttgt 900
 ggaggtgggg caaaggccaa gcttctgaaa gaattgggag tcgacagagt catagactat 960
 cacagtgaag atataaaaac ggttctaagg aaagagttcc cgaaaggtat tgatatcatc 1020
 tacgaatctg ttggtgggga catgttaaag ttgtgcttgg atgctttggc agtccatgga 1080
 cgactcattg tcattggcat gatttctcag tatcaaggag aaaatggttg gacgccatca 1140
 aaatatcctg gactatgtga gaagctcttg tcaaagagtc aaactgtggc tggcttttc 1200
 ctggtgcaat atatgcacat gtaccaagaa caccttaaca agttattga ccttactct 1260
 tccggaaaac taaaggttgc tgtggatcca aagagattta taggccttca ttctgttgc 1320
 gatgctgttg agtatctcca ttacggcaaa agcgttggga aggtggtgt ctgcgtggac 1380
 ccgacctcg gtcatcaagt agccaaatta tgaatgaaca cgggtgcaaa tacagaaaga 1440
 agtgaagttt tcaattctta gtctagagat tgtctcgaa tgttactgaa aatagctgct 1500
 agaccagtgc tggaaatatt atttcaatg cttttcaat ttggattac ttgaaagaat 1560
 aatccattta tgtataccat gtttatgtt acactataca acaactatga gcagaagaaa 1620
 gcgagatata taaaaataa attataatcc ttcatctta aaaaaaaaaa aaaaaaa 1677

<210> 41
 <211> 1330
 <212> DNA
 <213> human

<400> 41
 atggcgcagt gggacagct cactgatcaa caggaggaca ctgatagctg ttcagaatct 60
 gtgaagtttg atgctcgtc caatacagct ttgcttccc caaatctaa aaatggccct 120

ccacttcaag aaaagctgaa atccttcaaa gctgcactga ttgcccttta tctcctgtg 180
 ttgctgttc tcattcctat catcgcaata atggcagctc aactcctgaa gtgggaaatg 240
 aagaattgca cagttgggtc aattaatgca aacagtgtat cctccagtct cctgggaaga 300
 ggaaatgaca gtgaagatga agtgagattt cgagaagttg ttatggaaca cattagcaag 360
 atggagaaaa gaatccaata tatttcagat actgaagaaa atctcgtaga ttacagagcat 420
 ttcaaaaatt tcagtgtagc aactgatcaa cgatttgctg atgttctct ccaactaagt 480
 accttgggtc ccacagtcca gggacatggg aatgccgtag atgaaatcac cagggtccta 540
 ataagtctga ataccacgct gcttgattg cacctctatg tagaaacact gaatgtcaaa 600
 ttccaggaga atacacttaa agggcaagag gaaatcagca aattaaagga gcgtgtgcac 660
 aatgcatcag cagaaattat gtctatgaaa gaagaacaag tgcatttga acaggaaata 720
 aaaagagaag tgaaagtct gaataacatc actaatgatc tcaggctgaa agattgggaa 780
 catctcaga cgttgagaaa tatcacttta attcaaggtc ctctggacc cccaggagaa 840
 aaaggagata gaggtccaac cggagaaagt ggtccaccag gcgttcagg tccagtaggt 900
 cctccaggtc ttaagggta tcgaggatct attggcttc cggaagtcg aggatatcca 960
 ggacaatcag ggaagactgg gaggacagga tatctggac caaaaggcca aaaggagaa 1020
 aaaggcagtg gaagcatcct gactcctct gcgactgtcc gactggttg tggccgtggc 1080
 cctcatgagg gtagagtga gatattgcac aatggacagt ggggcacagt ttgtgatgat 1140
 cactgggaac tgcgtgccgg gcaggtgtc tgcaggagct tgggataccg aggtgttaag 1200
 agtgtgcaca agaaagctta ttttgacaa ggtactggtc ccattggct gaatgaagta 1260
 ccctgtttgg ggatggagtc atccattgaa gagtgcacaaa tcagacagtg gggcgtgaga 1320
 gtctgttcac 1330

<210> 42
 <211> 431
 <212> DNA
 <213> human

 <220>
 <221> misc_feature
 <222> (97)..(97)
 <223> any kind of base

 <220>

<221> misc_feature
<222> (347)..(347)
<223> any kind of base

<220>
<221> misc_feature
<222> (349)..(349)
<223> any kind of base

<220>
<221> misc_feature
<222> (361)..(361)
<223> any kind of base

<220>
<221> misc_feature
<222> (362)..(362)
<223> any kind of base

<220>
<221> misc_feature
<222> (363)..(363)
<223> any kind of base

<220>
<221> misc_feature
<222> (401)..(401)
<223> any kind of base

<220>
<221> misc_feature
<222> (428)..(428)
<223> any kind of base

<400> 42
ctttttatat ttattttcat cgctacacaa acatttttta ggagtttgat tctacctcca 60

tttggttag atatacaaac tctaccccat gagggantgt atggtgtatt tctagattta 120

gcaacaattt tcttgaaaaa tgtacaatac tatagaaaaa tgaagatagt aaataccagg 180

tataagttaa taacagtgtt tctttgttc agtaataatg aactgtgtac tagcactgaa 240

ctttaggccc tcctatttgc gtattttctg ttgtatatt tttaaataga ggaattgtga 300

ttataatatt attattttgg aatatcctaa atcataaatt caaaacntna ttagttttt 360

nnntttttt ttaagatgg agtcccgtt tgtcccaggc nggagtgacag tggcatgatc 420

tcagctcnct g

431

<210> 43
<211> 669
<212> DNA
<213> human

<220>
<221> misc_feature
<222> (641)..(641)
<223> any kind of base

<400> 43
ttcttttga aaaccaaaaca tgctttattt cattttttc acaatttatt taaacatctc 60

acatatcaaa aataggtaga atttaatttt tctgcttgcc caagaaacaa agcttctgtg 120
gaaccatgga agaagatgaa aatgagactg gcaaagaaca aatgctgaat ctgaagaaga 180
ggacaacttt gggcaaataa tctgcatact ttaattggg aataagatgg aaaatatgaa 240
tgctaaatca aatttttaa aaaatacacc acacgatata actcaatata ggagtatttc 300
ttctcaaatt ctctagcac catcaacatt ctcaagtat ctgaaatact attaattagc 360
acctttgtat tatgaacaaa acaaaacaag gacctcagtt catctctgtc taggtcagca 420
cctaacaatg tggatcacac tcatgggaaa gtgttttgag gtagtttaaa ctttggaag 480
tttgggtttt aaacttcctt ctgtggaaga tattcaaaag ccacaagtgg tgcaaatgtt 540
tatgggtttt attttcaat tttattttg gtttcttac aaagggtgac attttcata 600
acagggtgaa gagtgttgaa aaaaaaattt caatttttg ngggaacggg ggaaggagtt 660
aatgaaact 669

<210> 44
<211> 287
<212> DNA
<213> human

<400> 44
gccggagagt ctacaatgtt acccagcatg ctgttggcat tgttgtaaac aaacaagtta 60
agggcaagat tcttgccaag agaattaatg tgcgtattga gcacattaag cactctaaga 120
gccgagatag ctctctgaaa cgtgtgaagg aaaatgatca gaaaaagaaa gaagccaaag 180
agaaaggtag ctgggttcaa ctaaagcgcc acgctgctcc acccagagaa gcacactttg 240
tgagaaccaa tgggaaggag cctgagctgc tggaacctat tccttat 287

<210> 45
<211> 383
<212> DNA
<213> human

<220>
<221> misc_feature
<222> (147)..(147)
<223> any kind of base

<220>
<221> misc_feature
<222> (309)..(309)
<223> any kind of base

<220>
<221> misc_feature
<222> (349)..(349)
<223> any kind of base

<220>
<221> misc_feature
<222> (365)..(365)
<223> any kind of base

<220>
<221> misc_feature
<222> (372)..(372)
<223> any kind of base

<220>
<221> misc_feature
<222> (380)..(380)
<223> any kind of base

<400> 45
ggaacggaaa aggagaattc aagtgtgacc ctcatgaggc aacgtgttat gatgatggga 60
agacatacca cgtaggagaa cagtggcaga aggaatatct cggtgccatt tgctcctgca 120
catgctttgg aggccagcgg ggctcgnctt gtgacaactg ccgcagacct ggggggtgaa 180
cccagtcccg aaggcactac tggccagtc tacaaccagt attcttcaga galaccattc 240
agagaacaaa cactaatgtt taattgccc aatttgagtg ctcatgcct ttaggatgt 300
tacaggctng acagagaagg tttcccgag gagttaaata atcttttnc catttccga 360

ggggncaagg cntgttttn ttt

383

<210> 46
<211> 523
<212> DNA
<213> human

<400> 46
cagaggggca gggcggacgg ctaggagttc aagaaacatc ctggctgag ggaaaggctg 60
cagctgcacc gccatgaata agcttttcag cttctggaag aggaagaatg agaccgcag 120
ccagggtac aaccttcgag aaaaggattt aaagaaactt cacagagctg cttcagtcgg 180
ggatttgaag aagctgaagg aataccttca gatcaagaaa tatgatgtaa atatgcagga 240
ctatgaatac agaacacctt tgcacctagc ctgtgctaata ggacatacag atgtgttact 300
tttctaatt gagcaacaat gcaagataaa tgtccgggat agtgaaaaca aatccccatt 360
gattaaggca gtacagtgtc aaaatgagga ttgtgtact attctgctaa actttggtgc 420
agaccagat ctgagggata ttcgtataa tactgttctt cactatgctg ttgtggta 480
aagttgtca ttagttgaaa aactgctga atacgaagct gat 523

<210> 47
<211> 390
<212> DNA
<213> human

<400> 47
tccaaggta tggcaaaaaca tctgaagttc atcgccagga ctgtgatggt acaggaaggg 60
aacgtggaaa ggcatacag gaccctaac agaactctca ctatggaagg gctcattgag 120
gacattaagc atcggcggtat ttatgagaag ccatgccgcc gcgacagagg gaaagctatg 180
aaagggtccg gcggatctac aacatggaaa tggctcgcaa gatcaacttc ttgatcgaa 240
agaatcgggc agatccgtgg cagggtgtgt gaggcctgtg ggtgggacac cagtgcgaaa 300
ccctcatcca gtttctctc catctctttt cttgtacaa tccatttcc tattaccatt 360
ctctgaata aactcaaalc acatgtctgc 390

<210> 48
<211> 669
<212> DNA
<213> human

<220>

<221> misc_feature
<222> (641)..(641)
<223> any kind of base

<400> 48
ttcttttga aaaccaaaaca tgcatttatt catTTTTtTc acaatttatt taaacatctc 60

acatatataa aataggataga atttaatttt tctgcttgcc caagaaacaa agcttctgtg 120
gaaccatgga agaagatgaa aatgagactg gcaaagaaca aatgctgaat ctgaagaaga 180
ggacaacttt gggcaaataa tctgcatact ttaattggg aataagatgg aaaatatgaa 240
tgctaaatca aatttttaa aaaatacacc acacgataca actcaataga ggagtatttc 300
ttctcaaatt ctctagcac catcaacatt ctcaagtat ctgaaatact attaattagc 360
acctttgtat tatgaacaaa acaaaacaag gacctcagtt catctctgtc taggtcagca 420
cctaacaatg tggatcacac tcatgggaaa gtgtttgag gtagtttaa cctttggaag 480
tttgggtttt aaacttcct ctgtggaaga tattcaaaag ccacaagtgg tgcaaatgtt 540
tatggttttt attttcaat tttattttg gtttcttac aaagggtgac attttcata 600
acagggtgaa gagtgttgaa aaaaaaattt caatttttg ngggaacggg ggaaggagt 660
aatgaaact 669

<210> 49
<211> 431
<212> DNA
<213> human

<220>
<221> misc_feature
<222> (97)..(97)
<223> any kind of base

<220>
<221> misc_feature
<222> (347)..(347)
<223> any kind of base

<220>
<221> misc_feature
<222> (349)..(349)
<223> any kind of base

<220>

<221> misc_feature
<222> (361)..(361)
<223> any kind of base

<220>
<221> misc_feature
<222> (362)..(362)
<223> any kind of base

<220>
<221> misc_feature
<222> (363)..(363)
<223> any kind of base

<220>
<221> misc_feature
<222> (401)..(401)
<223> any kind of base

<220>
<221> misc_feature
<222> (428)..(428)
<223> any kind of base

<400> 49
ctttttatat ttattttcat cgctacacaa acatttttta ggagtttgat tctacctcca 60
tttgggttag atatacaaac tctaccccat gagggantgt atggtgtatt tctagattta 120
gcaacaattt tcttgaaaaa tgtacaatac tatagaaaaa tgaagatagt aaataccagg 180
tataagttaa taacagtggt tcttttggtc agtaataatg aactgtgtac tagcactgaa 240
ctttagggccc tctatttgc gtatttctg ttgtatatt tttaaataga ggaattgtga 300
ttataatatt attattttgg aataccttaa atcataaatt caaaacnina ttagttttt 360
nnntttttt ttaagatgg agtcccgctt tgtcccaggc nggagtgcag tggcatgatc 420
tcagctcnct g 431