The invention concerns sensitive methods to measure mRNA levels in biopsied tumor tissues, including archived paraffin-embedded biopsy material. The invention also concerns breast cancer gene sets important in the diagnosis and treatment of breast cancer, and methods for assigning the most optimal treatment options to breast cancer patient based upon knowledge derived from gene expression studies.
Overall FPET/RT-PCR Flow Chart

Receipt of Paraffin Block

Cut (Microtome) 10μm Sections

Extract RNA/Remove Protein and DNA

Analyze RNA Concentration by Ribogreen Protocol

Possible RNA Repair Steps

Possible RNA Amplification Steps

Reverse Transcription Using Gene Specific Promoters

Polymerase Chain Reaction Analysis with Optimized Probes and Primers

Data Retrieval/Analysis/Storage

FIG. 1
Alternative Scheme for Preparing Fragmented mRNA for Expression Profiling Analysis

5' → 3' purified mRNA fragments

3' → 5' ss DNA universal or gene specific template with T7 promoter

Mix, denature, anneal

Add DNA Polymerase and dNTP's

Primer Extend

IVT generated cRNA

DNase treat for pure cRNA

Figure 5
Figure 6
Progesterone Receptor Comparison

Figure 7
Gene specific amplification of RNA for Taqman Expression Profiling
Test of Concept

FIG. 8
GENE EXPRESSION PROFILING IN BIOPSIED TUMOR TISSUES

CROSS-REFERENCE

[0001] This application claims the benefit under 35 U.S.C. 119(b) of provisional applications Ser. Nos. 60/412,049, filed Sep. 18, 2002 and 60/364,890, filed Mar. 13, 2002, the entire disclosures which are hereby incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

[0002] The present invention relates to gene expression profiling in biopsied tumor tissues. In particular, the present invention concerns sensitive methods to measure mRNA levels in biopsied tumor tissues, including archived paraffin-embedded biopsy material. In addition, the invention provides a set of genes the expression of which is important in the diagnosis and treatment of breast cancer.

[0003] Oncologists have a number of treatment options available to them, including different combinations of chemotherapeutic drugs that are characterized as "standard of care," and a number of drugs that do not carry a label claim for a particular cancer, but for which there is evidence of efficacy in that cancer. Best likelihood of good treatment outcome requires that patients be assigned to optimal available cancer treatment, and that this assignment be made as quickly as possible following diagnosis.

[0004] Currently, diagnostic tests used in clinical practice are single analyte, and therefore do not capture the potential value of knowing relationships between dozens of different markers. Moreover, diagnostic tests are frequently not quantitative, relying on immunohistochemistry. This method often yields different results in different laboratories, in part because the reagents are not standardized, and in part because the interpretations are subjective and cannot be easily quantified. RNA-based tests have not often been used because of the problem of RNA degradation over time and the fact that it is difficult to obtain fresh tissue samples from patients for analysis. Fixed paraffin-embedded tissue is more readily available and methods have been established to detect RNA in fixed tissue. However, these methods typically do not allow for the study of large numbers of genes (DNA or RNA) from small amounts of material. Thus, traditionally fixed tissue has been rarely used other than for immunohistochemistry detection of proteins.

[0005] Recently, several groups have published studies concerning the classification of various cancer types by microarray gene expression analysis (see, e.g. Golub et al., Science 286:531-537 (1999); Bhattacharjee et al., Proc. Natl. Acad. Sci. USA 98:13790-13795 (2001); Chen-Hsiang et al., Bioinformatics 17 (Suppl. 1):S316-S322 (2001); Ramaswamy et al., Proc. Natl. Acad. Sci. USA 98:15149-15154 (2001)). Certain classifications of human breast cancers based on gene expression patterns have also been reported (Martin et al., Cancer Res. 60:2232-2238 (2000); West et al., Proc. Natl. Acad. Sci. USA 98:11462-11467 (2001); Sorlie et al., Proc. Natl. Acad. Sci. USA 98:10869-10874 (2001); Yan et al., Cancer Res. 61:8375-8380 (2001)). However, these studies mostly focus on improving and refining the already established classification of various types of cancer, including breast cancer, and generally do not provide new insights into the relationships of the differentially expressed genes, and do not link the findings to treatment strategies in order to improve the clinical outcome of cancer therapy.

[0006] Although modern molecular biology and biochemistry have revealed more than 100 genes whose activities influence the behavior of tumor cells, state of their differentiation, and their sensitivity or resistance to certain therapeutic drugs, with a few exceptions, the status of these genes has not been exploited for the purpose of routinely making clinical decisions about drug treatments. One notable exception is the use of estrogen receptor (ER) protein expression in breast carcinomas to select patients to treatment with anti-estrogen drugs, such as tamoxifen. Another exceptional example is the use of ErbB2 (Her2) protein expression in breast carcinomas to select patients with the Her2 antagonist drug Herceptin® (Genentech, Inc., South San Francisco, Calif.).

[0007] Despite recent advances, the challenge of cancer treatment remains to target specific treatment regimens to pathogenically distinct tumor types, and ultimately personalize tumor treatment in order to maximize outcome. Hence, a need exists for tests that simultaneously provide predictive information about patient responses to the variety of treatment options. This is particularly true for breast cancer, the biology of which is poorly understood. It is clear that the classification of breast cancer into a few subgroups, such as ErbB2+ subgroup, and subgroups characterized by low to absent gene expression of the estrogen receptor (ER) and a few additional transcriptional factors (Perou et al., Nature 406:747-752 (2000)) does not reflect the cellular and molecular heterogeneity of breast cancer, and does not allow the design of treatment strategies maximizing patient response.

SUMMARY OF THE INVENTION

[0008] The present invention provides (1) sensitive methods to measure mRNA levels in biopsied tumor tissue, (2) a set of approximately 190 genes, the expression of which is important in the diagnosis of breast cancer, and (3) the significance of abnormally low or high expression for the genes identified and included in the gene set, through activation or disruption of biochemical regulatory pathways that influence patient response to particular drugs used or potentially useful in the treatment of breast cancer. These results permit assessment of genomic evidence of the efficacy of more than a dozen relevant drugs.

[0009] The present invention accommodates the use of archived paraffin-embedded biopsy material for assay of all markers in the set, and therefore is compatible with the most widely available type of biopsy material. The invention presents an efficient method for extraction of RNA from wax-embedded, fixed tissues, which reduces cost of mass production process for acquisition of this information without sacrificing quality of the analysis. In addition, the invention describes a novel highly effective method for amplifying mRNA copy number, which permits increased assay sensitivity and the ability to monitor expression of large numbers of different genes given the limited amounts of biopsy material. The invention also captures the predictive significance of relationships between expressions of...
certain markers in the breast cancer marker set. Finally, for each member of the gene set, the invention specifies the oligonucleotide sequences to be used in the test.

[0010] In one aspect, the invention concerns a method for predicting clinical outcome for a patient diagnosed with cancer, comprising

[0011] determining the expression level of one or more genes, or their expression products, selected from the group consisting of p53BP2, cathepsin B, cathepsin L, Ki67/MiB1, and thymidine kinase in a cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference cancer tissue set,

[0012] wherein a poor outcome is predicted if:

[0013] (a) the expression level of p53BP2 is in the lower 10th percentile; or

[0014] (b) the expression level of either cathepsin B or cathepsin L is in the upper 10th percentile; or

[0015] (c) the expression level of any either Ki67/MiB1 or thymidine kinase is in the upper 10th percentile.

[0016] Poor clinical outcome can be measured, for example, in terms of shortened survival or increased risk of cancer recurrence, e.g., following surgical removal of the cancer.

[0017] In another embodiment, the inventor concerns a method of predicting the likelihood of the recurrence of cancer, following treatment, in a cancer patient, comprising determining the expression level of p27, or its expression product, in a cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference cancer tissue set, wherein an expression level in the upper 10th percentile indicates decreased risk of recurrence following treatment.

[0018] In another aspect, the invention concerns a method for classifying cancer comprising, determining the expression level of two or more genes selected from the group consisting of Bcl2, hepatocyte nuclear factor 3, ER, ErbB2, and Grb7, or their expression products, in a cancer tissue, normalized against a control gene or genes, and compared to the amount found in a reference cancer tissue set, wherein (i) tumors expressing at least one of Bcl2, hepatocyte nuclear factor 3, and ER, or their expression products, above the mean expression level in the reference tissue set are classified as having a good prognosis for disease free and overall patient survival following treatment; and (ii) tumors expressing elevated levels of ErbB2 and Grb7, or their expression products, at levels ten-fold or more above the mean expression level in the reference tissue set are classified as having poor prognosis of disease free and overall patient survival following treatment.

[0019] All types of cancer are included, such as, for example, breast cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, and brain cancer. The foregoing methods are particularly suitable for prognosis/classification of breast cancer.

[0020] In all previous aspects, in a specific embodiment, the expression level is determined using RNA obtained from a formalin-fixed, paraffin-embedded tissue sample. While all techniques of gene expression profiling, as well as proteomics techniques, are suitable for use in performing the foregoing aspects of the invention, the gene expression levels are often determined by reverse transcription polymerase chain reaction (RT-PCR).

[0021] If the source of the tissue is a formalin-fixed, paraffin embedded tissue sample, the RNA is often fragmented.

[0022] The expression data can be further subjected to multivariate analysis, for example using the Cox Proportional Hazards model.

[0023] In a further aspect, the invention concerns a method for the preparation of nucleic acid from a fixed, wax-embedded tissue specimen, comprising:

[0024] (a) incubating a section of the fixed, wax-embedded tissue specimen at a temperature of about 56° C. to 70° C. in a lysis buffer, in the presence of a protease, without prior dewaxing, to form a lysis solution;

[0025] (b) cooling the lysis solution to a temperature where the wax solidifies; and

[0026] (c) isolating the nucleic acid from the lysis solution.

[0027] The lysis buffer may comprise urea, such as 4M urea. In a particular embodiment, incubation in step (a) of the foregoing method is performed at about 65° C.

[0028] In another particular embodiment, the protease used in the foregoing method is proteinase K.

[0029] In another embodiment, the cooling in step (b) is performed at room temperature.

[0030] In a further embodiment, the nucleic acid is isolated after protein removal with 2.5 M NH₄OAc.

[0031] The nucleic acid can, for example, be total nucleic acid present in the fixed, wax-embedded tissue specimen.

[0032] In yet another embodiment, the total nucleic acid is isolated by precipitation from the lysis solution, following protein removal, with 2.5 M NH₄OAc. The precipitation may, for example, be performed with isopropanol.

[0033] The method described above may further comprise the step of removing DNA from the total nucleic acid, for example by DNase treatment.

[0034] The tissue specimen may, for example, be obtained from a tumor, and the RNA may be obtained from a microdissected portion of the tissue specimen enriched for tumor cells.

[0035] All types of tumor are included, such as, without limitation, breast cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, and brain cancer, in particular breast cancer.

[0036] The method described above may further comprise the step of subjecting the RNA to gene expression profiling.
Thus, the gene expression profile may be completed for a set of genes comprising at least two of the genes listed in Table 1.

Although all methods of gene expression profiling are contemplated, in a particular embodiment, gene expression profiling is performed by RT-PCR which may be preceded by an amplification step.

In another aspect, the invention concerns a method for preparing fragmented RNA for gene expression analysis, comprising the steps of:

(a) mixing the RNA with at least one gene-specific, single-stranded DNA scaffold under conditions such that fragments of the RNA complementary to the DNA scaffold hybridize with the DNA scaffold;

(b) extending the hybridized RNA fragments with a DNA polymerase to form a DNA-DNA duplex; and

(c) removing the DNA scaffold from the duplex.

In a specific embodiment, in step (b) of this method, the RNA may be mixed with a mixture of single-stranded DNA templates specific for each gene of interest.

The method can further comprise the step of heat-denaturing and reannealing the duplexed DNA to the DNA scaffold, with or without additional overlapping scaffolds, and further extending the duplexed sense strand with DNA polymerase prior to removal of the scaffold in step (c).

The DNA templates may be, but do not need to be, fully complementary to the gene of interest.

In a particular embodiment, at least one of the DNA templates is complementary to a specific segment of the gene of interest.

In another embodiment, the DNA templates include sequences complementary to polymorphic variants of the same gene.

The DNA template may include one or more dUTP or rNTP sites. In this case, in step (c) the DNA template may be removed by fragmenting the DNA template present in the DNA-DNA duplex formed in step (b) at the dUTP or rNTP sites.

In an important embodiment, the RNA is extracted from fixed, wax-embedded tissue specimens, and purified sufficiently to act as a substrate in an enzyme assay. The RNA purification may, but does not need to, include an oligo-dT based step.

In a further aspect, the invention concerns a method for amplifying RNA fragments in a sample comprising fragmented RNA representing at least one gene of interest, comprising the steps of:

(a) contacting the sample with a pool of single-stranded DNA scaffolds comprising an RNA polymerase promoter at the 5' end under conditions such that the RNA fragments complementary to the DNA scaffolds hybridize with the DNA scaffolds;

(b) extending the hybridized RNA fragments with a DNA polymerase along the DNA scaffolds to form DNA-DNA duplexes;

(c) amplifying the gene or genes of interest by in vitro transcription; and

(d) removing the DNA scaffolds from the duplexes.

An exemplary promoter is the T7 RNA polymerase promoter, while an exemplary DNA polymerase is DNA polymerase 1.

In step (d) the DNA scaffolds may be removed, for example, by treatment with DNase I.

In a further embodiment, the pool of single-stranded DNA scaffolds comprises partial or complete gene sequences of interest, such as a library of cDNA clones.

In a specific embodiment, the sample represents a whole genome or a fraction thereof. In a preferred embodiment, the genome is the human genome.

In another aspect, the invention concerns a method of preparing a personalized genomics profile for a patient, comprising the steps of:

(a) subjecting RNA extracted from a tissue obtained from the patient to gene expression analysis;

(b) determining the expression level in such tissue of at least two genes selected from the gene set listed in Table 1, wherein the expression level is normalized against a control gene or genes, and is compared to the amount found in a cancer tissue reference set;

(c) and creating a report summarizing the data obtained by the gene expression analysis.

The tissue obtained from the patient may, but does not have to, comprise cancer cells. Just as before, the cancer can, for example, be breast cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, or brain cancer, breast cancer being particularly preferred.

In a particular embodiment, the RNA is obtained from a microdissected portion of breast cancer tissue enriched for cancer cells. The control gene set may, for example, comprise S-actin, and ribosomal protein LPO.

The report prepared for the use of the patient or the patient's physician may include the identification of at least one drug potentially beneficial in the treatment of the patient.

Step (b) of the foregoing method may comprise the step of determining the expression level of a gene specifically influencing cellular sensitivity to a drug, where the gene can, for example, be selected from the group consisting of aldehyde dehydrogenase 1A1, aldehyde dehydrogenase 1A3, amphiurelin, ARG, BRK, BCRP, CD9, CD31, CD82/KAI-1, COX2, c-abl, c-kit, c-kit L, CYP1B1, CYP2C9, DHFR, dihydropyrimidine dehydrogenase, EGF, erigulin, ER-alpha, ErbB-1, ErbB-2, ErbB-3, ErbB-4, ER-beta, farnesyl pyrophosphate synthetase, gamma-GCS (glutamyl cysteine synthetase), GATA3, geranyl pyrophosphate synthetase, Grb7, GST-alpha, GST-pi, HB-EGF, hsp 27, human chorionic gonadotropin/CGA, IGF-1, IGF-2, IGFIR, KDR, LIV1, Lung Resistance Protein/MVP, Lot1, MDR-1, microsomal epoxide hydrolase, MMP9, MRPI, MRP2, MRP3, MRP4, PAI1, PDGF-A, PDGF-B, PDGF-C, PDGF-
D, PDGFR-alpha, PDGFR-beta, PLA2a (pleiomorphic adenoma 1), PREP prolyl endopeptidase, progesterone receptor, p52/trefoil factor 1, PTEN, PTB11b, RAR-alpha, RAR-beta2, Reduced Folate Carrier, SXR, TGF-alpha, thymidine phosphorylase, thymidine synthase, topoisomerase II-alpha, topoisomerase II-beta, VEGF, XIST, and YB-1.

[0066] In another embodiment, step (b) of the foregoing process comprises determining the expression level of a gene marker that defines a subclass of breast cancer, where the gene marker can, for example, be GRO1 oncogene alpha, Grb7, cytokeratins 5 and 17, retinol binding protein 4, hepatocyte nuclear factor 3, integrin subunit alpha 7, or lipoprotein lipase.

[0077] In a still further aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to 5-fluorouracil (5-FU) or an analog thereof, comprising the steps of:

[0078] (a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis;

[0079] (b) determining the expression level in the tissue of thymidylate synthase mRNA, wherein the expression level is normalized against a control gene or genes, and is compared to the amount found in a reference breast cancer tissue set; and

[0080] (c) predicting patient response based on the normalized thymidylate synthase mRNA level.

[0081] Step (d) of the foregoing method can further comprise determining the expression level of dihydoropyrimidine phosphorylase.

[0082] In another embodiment, step (b) of the method can further comprise determining the expression level of thymidine phosphorylase.

[0083] In yet another embodiment, a positive response to 5-FU or an analog thereof is predicted if: (i) normalized thymidylate synthase mRNA level determined in step (b) is at or below the 10th percentile; or (ii) the sum of normalized expression levels of thymidylate synthase and dihydoropyrimidine phosphorylase determined in step (b) is at or below the 25th percentile; or (iii) the sum of normalized expression levels of thymidylate synthase, dihydoropyrimidine phosphorylase, plus thymidine phosphorylase determined in step (b) is at or below the 20th percentile.

[0084] In a further embodiment, step (b) of the foregoing method the expression level of c-myc and wild-type p53 is determined. In this case, a positive response to 5-FU or an analog thereof is predicted, if the normalized expression level of c-myc relative to the normalized expression level of wild-type p53 is in the upper 10th percentile.

[0085] In a still further embodiment, in step (b) of the foregoing method, expression level of NFkB and cIAP2 is determined. In this particular embodiment, resistance to 5-FU or an analog thereof is typically predicted if the normalized expression level of NFkB and cIAP2 is at or above the 10th percentile.

[0086] In another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to methotrexate or an analog thereof, comprising the steps of:

[0087] (a) subjecting RNA extracted from a breast cancer tissue obtained from the patient in gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

[0088] (b) predicting decreased patient sensitivity to methotrexate or analog if (i) DHFR levels are more than tenfold higher than the average expression level of DHFR in the control gene set, or (ii) the normalized expression levels of members of the reduced folate carrier (RFC) family are below the 10th percentile.
In yet another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to an anthracycline or an analog thereof, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting patient resistance or decreased sensitivity to the anthracycline or analog if (i) the normalized expression level of topoisomerase IIα is below the 10th percentile, or (ii) the normalized expression level of topoisomerase IIβ is below the 10th percentile, or (iii) the combined normalized topoisomerase IIα or IIβ expression levels are below the 10th percentile.

In a different aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to a docetaxel, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting reduced sensitivity to docetaxel if the normalized expression level of CYP1B1 is in the upper 10th percentile.

The invention further concerns a method for predicting the response of a patient diagnosed with breast cancer to cyclophosphamide or an analog thereof, comprising:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting reduced sensitivity to the cyclophosphamide or analog if the sum of the expression levels of aldehyde dehydrogenase 1A1 and 1A3 is more than tenfold higher than the average of their combined expression levels in the reference tissue set.

In a further aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to anti-estrogen therapy, comprising:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set that contains both specimens negative for and positive for estrogen receptor-α (ERα) and progesterone receptor-α (PRα); and

(b) predicting patient response based upon the normalized expression levels of ERα or PRα and at least one of microsomal epoxide hydrolase, pS2/trefoil factor 1, GATA3 and human chorionic gonadotropin.

In a specific embodiment, lack of response or decreased responsiveness is predicted if (i) the normalized expression level of microsomal epoxide hydrolase is in the upper 10th percentile; or (ii) the normalized expression level of pS2/trefoil factor 1, or GATA3 or human chorionic gonadotropin is at or below the corresponding average expression level in said breast cancer tissue set, regardless of the expression level of ERα or PRα in the breast cancer tissue obtained from the patient.

In another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to a taxane, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting reduced sensitivity to taxane if (i) no or minimal XIST expression is detected; or (ii) the normalized expression level of GST-π or propyl endopeptidase (PREP) is in the upper 10th percentile; or (iii) the normalized expression level of PLAG1 is in the upper 10th percentile.

The invention also concerns a method for predicting the response of a patient diagnosed with breast cancer to cisplatin or an analog thereof, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting resistance or reduced sensitivity if the normalized expression level of ERCC1 is in the upper 10th percentile.

The invention further concerns a method for predicting the response of a patient diagnosed with breast cancer to an ErbB2 or EGF/ER antagonist, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting patient response based on the normalized expression levels of at least one of Grb7, IGF1R, IGF1 and IGF2.

In particular embodiment, a positive response is predicted if the normalized expression level of Grb7 is in the upper 10th percentile, and the expression of IGF1R, IGF1 and IGF2 is not elevated above the 90th percentile.

In a further particular embodiment, a decreased responsiveness is predicted if the expression level of at least one of IGF1R, IGF1 and IGF2 is elevated.

In another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to a bis-phosphate drug, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and
(b) predicting a positive response if the breast cancer tissue obtained from the patient expresses mutant Ha-Ras and additionally expresses farnesyl pyrophosphate synthetase or geranyl pyrophosphate synthetase at a normalized expression level at or above the 90th percentile.

[0116] In yet another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to treatment with a cyclooxygenase 2 inhibitor, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting a positive response if the normalized expression level of COX2 in the breast cancer tissue obtained from the patient is at or above the 90th percentile.

[0119] The invention further concerns a method for predicting the response of a patient diagnosed with breast cancer to an EGFR-receptor (EGFR) antagonist, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting a positive response to an EGFR antagonist, if (i) the normalized expression level of EGFR is at or above the 10th percentile, and (ii) the normalized expression level of at least one of epiuregulin, TGF-α, amphiregulin, ErbB3, BRK, CD9, MMP9, CD82, and Let1 is above the 90th percentile.

[0122] In another aspect, the invention concerns a method for monitoring the response of a patient diagnosed with breast cancer to treatment with an EGFR antagonist, comprising monitoring the expression level of a gene selected from the group consisting of epiuregulin, TGF-α, amphiregulin, ErbB3, BRK, CD9, MMP9, CD82, and Let1 in the patient during treatment, wherein reduction in the expression level is indicative of positive response to such treatment.

[0123] In yet another aspect, the invention concerns a method for predicting the response of a patient diagnosed with breast cancer to a drug targeting a tyrosine kinase selected from the group consisting of abl, c-kit, PDGFR-α, PDGFR-β and ARG, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set;

(b) determining the normalized expression level of a tyrosine kinase selected from the group consisting of abl, c-kit, PDGFR-α, PDGFR-β and ARG, and the cognate ligand of the tyrosine kinase, and if the normalized expression level of the tyrosine kinase is in the upper 10th percentile,

(c) determining whether the sequence of the tyrosine kinase contains any mutation,

[0127] wherein a positive response is predicted if (i) the normalized expression level of the tyrosine kinase is in the upper 10th percentile, (ii) the sequence of the tyrosine kinase contains an activating mutation, or (iii) the normalized expression level of the tyrosine kinase is normal and the expression level of the ligand is in the upper 10th percentile.

[0128] Another aspect of the invention is a method for predicting the response of a patient diagnosed with breast cancer to treatment with an anti-angiogenic drug, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) predicting a positive response if (i) the normalized expression level of VEGF is in the upper 10th percentile and (ii) the normalized expression level of KDR or CD31 is in the upper 20th percentile.

[0131] A further aspect of the invention is a method for predicting the likelihood that a patient diagnosed with breast cancer develops resistance to a drug intersecting with the MRP-1 gene coding for the multidrug resistance protein P-glycoprotein, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis to determine the expression level of PTP1b, wherein the expression level is normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) concluding that the patient is likely to develop resistance to said drug if the normalized expression level of the MRP-1 gene is above the 90th percentile.

[0134] The invention further relates to a method for predicting the likelihood that a patient diagnosed with breast cancer develops resistance to a chemotherapeutic drug or toxin used in cancer treatment, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

(b) determining the normalized expression levels of at least one of the following genes: MDR1, SGT1, GSTδ1, SXR, BCRP YB-1, and LRP/MVP, wherein the finding of a normalized expression level in the upper 4th percentile is an indication that the patient is likely to develop resistance to the drug.

[0137] Also included herein is a method for measuring the translational efficiency of VEGF mRNA in a breast cancer tissue sample, comprising determining the expression levels of the VEGF and EIF4E mRNA in the sample, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a higher normalized EIF4E expression level for the same VEGF expression level is indicative of relatively higher translational efficiency for VEGF.

[0138] In another aspect, the invention provides a method for predicting the response of a patient diagnosed with breast
cancer to a VEGF antagonist, comprising determining the expression level of VEGF and EIF4E mRNA normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a VEGF expression level above the 90th percentile and an EIF4E expression level above the 50th percentile is a predictor of good patient response.

[0139] The invention further provides a method for predicting the likelihood of the recurrence of breast cancer in a patient diagnosed with breast cancer, comprising determining the ratio of p53:p21 mRNA expression or p53:mdm2 mRNA expression in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein an above normal ratio is indicative of a higher risk of recurrence. Typically, a higher risk of recurrence is indicated if the ratio is in the upper 10th percentile.

[0140] In yet another aspect, the invention concerns a method for predicting the likelihood of the recurrence of breast cancer in a breast cancer patient following surgery, comprising determining the expression level of cyclin D1 in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein an expression level in the upper 10th percentile indicates increased risk of recurrence following surgery. In a particular embodiment of this method, the patient is subjected to adjuvant chemotherapy, if the expression level is in the upper 10th percentile.

[0141] Another aspect of the invention is a method for predicting the likelihood of the recurrence of breast cancer in a breast cancer patient following surgery, comprising determining the expression level of APC or E-cadherin in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein an expression level in the upper 5th percentile indicates high risk of recurrence following surgery, and heightened risk of shortened survival.

[0142] A further aspect of the invention is a method for predicting the response of a patient diagnosed with breast cancer to treatment with a proapoptotic drug comprising determining the expression levels of BCL2 and c-MYC in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein (i) a BCL2 expression level in the upper 10th percentile in the absence of elevated expression of c-MYC indicates good response, and (ii) a good response is not indicated if the expression level c-MYC is elevated, regardless of the expression level of BCL2.

[0143] A still further aspect of the invention is a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising the steps of:

(a) subjecting RNA extracted from a breast cancer tissue obtained from the patient to gene expression analysis, wherein gene expression levels are normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set; and

[0144] (b) determining the normalized expression levels of NFkB and at least one gene selected from the group consisting of cIAP1, cIAP2, XIAP, and Survivin,

[0145] wherein a poor prognosis is indicated if the expression levels for NFkB and at least one of the genes selected from the group consisting of cIAP1, cIAP2, XIAP, and Survivin is in the upper 5th percentile.

[0147] The invention further concerns a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression levels of p53BP1 and p53BP2 in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor outcome is predicted if the expression level of either p53BP1 or p53BP2 is in the lower 10th percentile.

[0148] The invention additionally concerns a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression levels of uPA and PAI1 in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein (i) a poor outcome is predicted if the expression levels of uPA and PAI1 are in the upper 20th percentile, and (ii) a decreased risk of recurrence is predicted if the expression levels of uPA and PAI1 are not elevated above the mean observed in the breast cancer reference set. In a particular embodiment, poor outcome is measured in terms of shortened survival or increased risk of cancer recurrence following surgery. In another particular embodiment, uPA and PAI1 are expressed at normal levels, and the patient is subjected to adjuvant chemotherapy following surgery.

[0149] Another aspect of the invention is a method for predicting treatment outcome in a patient diagnosed with breast cancer, comprising determining the expression levels of cathepsin B and cathepsin L in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor outcome is predicted if the expression level of either cathepsin B or cathepsin L is in the upper 10th percentile. Just as before, poor treatment outcome may be measured, for example, in terms of shortened survival or increased risk of cancer recurrence.

[0150] A further aspect of the invention is a method for devising the treatment of a patient diagnosed with breast cancer, comprising the steps of:

(a) determining the expression levels of scatter factor and c-met in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, and

(b) suggesting prompt aggressive chemotherapeutic treatment if the expression levels of scatter factor and c-met or the combination of both, are above the 90th percentile.

[0153] A still further aspect of the invention is a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression levels of VEGF, CD31, and KDR in a breast cancer tissue
obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor treatment outcome is predicted if the expression level of any of VEGF, CD31, and KDR is in the upper 10th percentile.

[0154] Yet another aspect of the invention is a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression levels of Ki67/Mib1, PCNA, Pin1, and thymidine kinase in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor treatment outcome is predicted if the expression level of any of Ki67/Mib1, PCNA, Pin1, and thymidine kinase is in the upper 10th percentile.

[0155] The invention further concerns a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression level of soluble and full length CD95 in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor treatment outcome is predicted if the expression levels of soluble and full length CD95 correlate with poor patient survival.

[0156] The invention also concerns a method for predicting treatment outcome for a patient diagnosed with breast cancer, comprising determining the expression levels of IGF1, IGF1R and IGFBP3 in a breast cancer tissue obtained from the patient, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein a poor treatment outcome is predicted if the expression levels of IGF1, IGF1R and IGFBP3 is in the upper 10th percentile.

[0157] The invention additionally concerns a method for classifying breast cancer comprising, determining the expression level of two or more genes selected from the group consisting of Bel12, hepatocyte nuclear factor 3, LIV1, ER, lipoprotein lipase, retinol binding protein 4, integrin α7, cytokeratin 5, cytokeratin 17, GRO oncogene, ErbB2 and Grb7, in a breast cancer tissue, normalized against a control gene or genes, and compared to the amount found in a reference breast cancer tissue set, wherein (i) tumors expressing at least one of Bel12, hepatocyte nuclear factor 3, LIV1, and ER above the mean expression level in the reference tissue set are classified as having a good prognosis for disease free and overall patient survival following surgical removal; (ii) tumors characterized by elevated expression of at least one of lipoprotein lipase, retinol binding protein 4, integrin α7 compared to the reference tissue set are classified as having intermediate prognosis of disease free and overall patient survival following surgical removal; and (iii) tumors expressing either elevated levels of cytokeratins 5 and 17, and GRO oncogene at levels four-fold or greater above the mean expression level in the reference tissue set, or ErbB2 and Grb7 at levels ten-fold or more above the mean expression level in the reference tissue set are classified as having poor prognosis of disease free and overall patient survival following surgical removal.

[0158] Another aspect of the invention is a panel of two or more gene specific primers selected from the group consisting of the forward and reverse primers listed in Table 2.

[0159] Yet another aspect of the invention is a method for reverse transcription of a fragmented RNA population in RT-PCR amplification, comprising using a multiplicity of gene specific primers as the reverse primers in the amplification reaction. In a particular embodiment, the method uses between two and about 40,000 gene specific primers in the same amplification reaction. In another embodiment, the gene specific primers are about 18 to 24 bases, such as about 20 bases in length. In another embodiment, the primer set includes the 1m of the primers is about 58–60°C. The primers can, for example, be selected from the group consisting of the forward and reverse primers listed in Table 2.

[0160] The invention also concerns a method of reverse transcriptase driven first strand cDNA synthesis, comprising using a gene specific primer of about 18 to 24 bases in length and having a Tm optimum between about 58°C and about 60°C. In a particular embodiment, the first strand cDNA synthesis is followed by PCR DNA amplification, and the primer serves as the reverse primer that drives the PCR amplification. In another embodiment, the method uses a plurality of gene specific primers in the same first strand cDNA synthesis reaction mixture. The number of the gene specific primers can, for example, be between 2 and about 40,000.

[0161] In a different aspect, the invention concerns a method of predicting the likelihood of long-term survival of a breast cancer patient without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising determining the expression level of one or more prognostic RNA transcripts or their product in a breast cancer tissue sample obtained from said patient, normalized against the expression level of all RNA transcripts or their products in said breast cancer tissue sample, or of a reference set of RNA transcripts or their products, wherein the prognostic transcript is the transcript of one or more genes selected from the group consisting of: FOXM1, PRAME, Bcl2, STK15, CEGP1, Ki-67, GSTM1, CA9, PR, BBC3, NME1, SURV, GATA3, TFRC, YB-1, Dipyd, GSTM3, RPS6 KBI, Src, Chk1, Id1, EstR1, p27, CCNB1, XIAP, Chk2, CDC25B, IGF1R, AK055699, P13KCC2A, TGFβ3, BAG11, CYP3A4, EPcam, VEGF, pS2, Hit1, WISP1, HFN3A, NEK65B, BRCA2, EGFIR, TRK1, VDR, Contig51037, pENT1, EPBX1, IF1A, DIABLO, CDH1, IF10, EGFIR, TRK1, VDR, EPBX1, IF1A, Contig51037, CDH1, IF10, IF10a, IGFBP3, CTSD, Her2, and pENT1 indicates a decreased likelihood of long-term survival without breast cancer recurrence, and the overexpression of one or more of Bcl2, CEGP1, GSTM1, PRAME, BBC3, GATA3, Dipyd, GSTM3, Id1, EstR1, p27, XIAP, IGF1R, AK055699, P13KCC2A, TGFβ3, BAG11, pS2, WISP1, HFN3A, NEK65B, and DIABLO indicates an increased likelihood of long-term survival without breast cancer recurrence.

[0162] In a particular embodiment of this method, the expression level of at least 2, preferably at least 5, more preferably at least 10, most preferably at least 15 prognostic transcripts or their expression products is determined.

[0163] When the breast cancer is invasive breast carcinoma, including both estrogen receptor (ER) overexpressing (ER positive) and ER negative tumors, the analysis includes determination of the expression levels of the transcripts of at
least two of the following genes, or their expression products:

- FOXM1, PRAME, Bcl2, STK15, CEGP1, Ki-67, GSTM1, PR, BBC3, NME1, SURV, GATA3, TFRC, YB-1, DPYD, Src, CA9, Contig51037, RPS6K1 and Her2.

When the breast cancer is ER positive invasive breast carcinoma, the analysis includes determination of the expression levels of the transcripts of at least two of the following genes, or their expression products: PRAME, Bcl2, FOXM1, DIABLO, EPHX1, HIF1A, VEGFC, Ki-67, IGF1R, VDR, NME1, GSTM3, Contig51037, CDC25B, CTSB, p27, CDH1, and IGFBP3.

Just as before, it is preferred to determine the expression levels of at least 5, more preferably at least 10, most preferably at least 15 genes, or their respective expression products.

In a particular embodiment, the expression level of one or more prognostic RNA transcripts is determined, where RNA may, for example, be obtained from a fixed, wax-embedded breast cancer tissue specimen of the patient. The isolation of RNA can, for example, be carried out following any of the procedures described above or throughout the application, or by any other method known in the art.

In yet another aspect, the invention concerns an array comprising nucleic acids hybridizing to the following genes: FOXM1, PRAME, Bcl2, STK15, CEGP1, Ki-67, GSTM1, PR, BBC3, NME1, SURV, GATA3, TFRC, YB-1, DPYD, CA9, Contig51037, RPS6K1 and Her2, immobilized on a solid surface.

In a particular embodiment, the array comprises nucleic acids hybridizing to the following genes: FOXM1, PRAME, Bcl2, STK15, CEGP1, Ki-67, GSTM1, CA9, PR, BBC3, NME1, SURV, GATA3, TFRC, YB-1, DPYD, GSTM3, RPS6 KB1, Src, Chk1, ID1, EstR1, p27, CCNB1, XIAP, Chk2, CDC25B, IGF1R, AK055699, P13K2CA, TGFBR3, BAG11, CYP3A4, EpCAM, VEGFC, p52, hENT1, WISP1, HNF3A, NFKBp65, BRC2A2, EGF1R, TK1, VDR, Contig51037, pENT1, EPHX1, IFI1A, CDH1, HIF1A, IGFBP3, CTSB, Her2 and DIABLO.

In a further aspect, the invention concerns a method of predicting the likelihood of long-term survival of a patient diagnosed with invasive breast cancer, without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising the steps of:

1. Determining the expression levels of the RNA transcripts or the expression products of genes of a gene set selected from the group consisting of:
   - Bcl2, cyclinG1, NFKBp65, NME1, EPHX1, TOP2B, D5R, TERC, SRC, DIABLO;
   - Ki67, XAP, hENT1, TS, CD9, p27, cyclinG1, p52, NFKBp65, CYP3A4;
   - GSTM1, XAP, Ki67, TS, cyclinG1, p27, CYP3A4, p52, NFKBp65, ErbB3;
   - PR, NME1, XAP, upa, cyclinG1, Contig51037, TERC, EPHX1, ALDH1A3, CTSL;
   - CA9, NME1, TERC, cyclinG1, EPHX1, DPYD, Src, TOP2B, NFKBp65, VEGFC;
   - TFRC, XAP, Ki67, TS, cyclinG1, p27, CYP3A4, p52, ErbB3, NFKBp65;

2. Subjecting the data obtained in step (a) to statistical analysis; and

3. Determining whether the likelihood of said long-term survival has increased or decreased.

In a still further aspect, the invention concerns a method of predicting the likelihood of long-term survival of a patient diagnosed with estrogen receptor (ER)-positive invasive breast cancer, without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising the steps of:

1. Determining the expression levels of the RNA transcripts or the expression products of genes of a gene set selected from the group consisting of:
   - PRAME, p27, IGFBP2, HIF1A, TIMP2, IL2, CYP3A4, ID1, EstrR1, DIABLO;
   - Contig51037, EPHX1, Ki67, TIMP2, cyclinG1, Dpyd, CYP3A4, TP, AIB1, CYP2C8;
   - Becl2, hENT1, FOXM1, Contig51037, cyclinG1, Contig46653, PTEN, CYP3A4, TIMP2, AREG;
   - HIF1A, PRAME, p27, IGFBP2, TIMP2, ILT2, CYP3A4, ID1, EstrR1, DIABLO;
   - IGFB1R, PRAME, EPHX1, Contig51037, cyclinG1, Becl2, NME1, Pten, TBP, TIMP2;
   - FOXM1, Contig51037, VEGFC, TP, HIF1A, Dpyd, RAD51C, DCR3, cyclinG1, BAG1;
   - EPHX1, Contig51037, Ki67, TIMP2, cyclinG1, Dpyd, CYP3A4, TP, AIB1, CYP2C8;
[0198] (h) Ki67, VEGFC, VDR, GSTM3, p27, upa, ITGA7, rhoC, TERC, Pin1;
[0199] (i) CDC25B, Contig51037, hENT1, Bcl2, HLAG, TERC, NME1, upa, ID1, CYP7;
[0200] (j) VEGFC, Ki67, VDR, GSTM3, p27, upa, ITGA7, rhoC, TERC, Pin1;
[0201] (k) CTSB, PRAME, p27, IGBPBP2, EPHX1, CTS, BAD, DR5, DCR3, XIAP;
[0202] (l) DIABLO, Ki67, hENT1, TIMP2, ID1, p27, KRT19, IGBPBP2, TS, PDGF;
[0203] (m) p27, PRAME, IGBPBP2, HIF1A, TIMP2, ILT2, CYP3A4, ID1, ESR1, DIABLO;
[0204] (a) CDH1, PRAME, VEGFC, HIF1A, DPYD, TIMP2, CYP3A4, ESR1, RBP4, p27;
[0205] (o) IGBPBP, PRAME, p27, Bcl2, XIAP, ESR1, Ki67, TS, Src, VEGF;
[0206] (p) GSTM3, PRAME, p27, IGBPBP3, XIAP, FGF2, hENT1, PTEN, ESR1, APC;
[0207] (q) hENT1, Bcl2, FOXM1, Contig51037, CyclinG1, Contig46653, PTEN, CYP3A4, TIMP2, AREG;
[0208] (r) STK15, VEGFC, PRAME, p27, GCLC, hENT1, ID1, TIMP2, ESR1, MCP1;
[0209] (s) NME1, PRAME, p27, IGBPBP3, XIAP, PTEN, hENT1, Bcl2, CYP3A4, HLAG;
[0210] (t) VDR, Bcl2, p27, hENT1, p53, PI3KCA, EIF4E, TFR, MCM3, ID1;
[0211] (u) EIF4E, Contig51037, EPHX1, cyclinG1, Bcl2, DR5, TBP, PTEN, NME1, HER2;
[0212] (v) CCNB1, PRAME, VEGFC, HIF1A, hENT1, GCLC, TIMP2, ID1, p27, upa;
[0213] (w) ID1, PRAME, DIABLO, hENT1, p27, PDGFRa, NME1, BNI, BRCa1, TP;
[0214] (x) FBX05, PRAME, IGBPBP3, p27, GSTM3, hENT1, XIAP, FGF2, TS, PTEN;
[0215] (y) GUS, H1A1A, VEGFC, GSTM3, DPYD, hENT1, EBX05, CA9, CYP, KRT18; and
[0216] (z) Bcl6, Bcl2, hENT1, Contig51037, HLAG, CD9, ID1, BRCa1, BIN1, HBEFG;
[0217] (2) subjecting the data obtained in step (1) to statistical analysis; and
[0218] (3) determining whether the likelihood of said long-term survival has increased or decreased.
[0219] In a different aspect, the invention concerns an array comprising nucleic acids hybridizing to a gene set selected from the group consisting of:

[0220] (a) Bcl2, cyclinG1, NFKBp65, NME1, EPHX1, TOP2A, DR5, TERC, Src, DIABLO;
[0221] (b) Ki67, XIAP, hENT1, TS, CD9, p27, cyclinG1, pS2, NFKBp65, CYP3A4;
[0222] (c) GSTM1, XIAP, Ki67, TS, cyclinG1, p27, CYP3A4, pS2, NFKBp65, ErbB3;
[0223] (d) PR, NME1, —XIAP, upa, cyclinG1, Contig51037, TERC, EPHX1, ALDH1A3, CTS;
[0224] (e) CA9, NME1, TERC, cyclinG1, EPHX1, DPYD, Src, TOP2A, NFKBp65, VEGFC;
[0225] (f) TFR, XIAP, Ki67, TS, cyclinG1, p27, CYP3A4, pS2, ErbB3, NFKBp65;
[0226] (g) Bcl2, PRAME, cyclinG1, FOXM1, NFKBp65, TS, XIAP, Ki67, CYP3A4, p27;
[0227] (h) FOXM1, cyclinG1, XIAP, Contig51037, PRAME, TS, Ki67, PDGFRa, p27, NFKBp65;
[0228] (i) PRAME, FOXM1, cyclinG1, XIAP, Contig51037, TS, Ki67, PDGFRa, p27, NFKBp65;
[0229] (j) Ki67, XIAP, PRAME, hENT1, contig51037, TS, CD9, p27, ErbB3, cyclinG1;
[0230] (k) STK15, XIAP, PRAME, PLAUR, p27, CTS, CD18, PREP, p53, RPS6 KB1;
[0231] (l) GSTM1, XIAP, PRAME, p27, Contig51037, ErbB3, GSTp, EREG, ID1, PLAUR;
[0232] (m) PR, PRAME, NME1, XIAP, PLAUR, cyclinG1, Contig51037, TERC, EPHX1, DR5;
[0233] (a) CA9, FOXM1, cyclinG1, XIAP, TS, Ki67, NFKBp65, CYP3A4, GSTM3, p27;
[0234] (o) TFR, XIAP, PRAME, p27, Contig51037, ErbB3, DPYD, TERC, NME1, VEGFC; and
[0235] (p) CEGP1, PRAME, hENT1, XIAP, Contig51037, ErbB3, DPYD, NFKBp65, ID1, TS, immobilized on a solid surface.

[0236] In an additional aspect, the invention concerns an array comprising nucleic acids hybridizing to a gene set selected from the group consisting of:

[0237] (a) PRAME, p27, IGBPBP2, HIF1A, TIMP2, ILT2, CYP3A4, ID1, ESR1, DIABLO;
[0238] (b) Contig51037, EPHX1, Ki67, TIMP2, cyclinG1, DPYD, CYP3A4, TP, AIB1, CYP2C8;
[0239] (c) Bcl2, hENT1, FOXM1, Contig51037, cyclinG1, Contig46653, PTEN, CYP3A4, TIMP2, AREG;
[0240] (d) HIF1A, PRAME, p27, IGBPBP2, TIMP2, ILT2, CYP3A4, ID1, ESR1, DIABLO;
[0241] (e) IGFR, PRAME, EPHX1, Contig51037, cyclinG1, Bcl2, NME1, PTEN, TBP, TIMP2;
[0242] (f) FOXM1, Contig51037, VEGFC, TBP, HIF1A, DPYD, RAD51C, DCR3, cyclinG1, BAG1;
[0243] (g) EPHX1, Contig51037, Ki67, TIMP2, cyclinG1, DPYD, CYP3A4, TP, AIB1, CYP2C8;
[0244] (h) Ki67, VEGFC, VDR, GSTM3, p27, upa, ITGA7, rhoC, TERC, Pin1;
[0245] (i) CDC25B, Contig51037, hENT1, Bcl2, HLAG, TERC, NME1, upa, ID1, CYP;
[0246] (j) VEGFC, Ki67, VDR, GSTM3, p27, upa, ITGA7, rhoC, TERC, Pin1;
FIG. 2 is a flow chart showing the steps of an RNA extraction method according to the invention alongside a flow chart of a representative commercial method.

FIG. 3 is a scheme illustrating the steps of an improved method for preparing fragmented mRNA for expression profiling analysis.

FIG. 4 illustrates methods for amplification of RNA prior to RT-PCR.

FIG. 5 illustrates an alternative scheme for repair and amplification of fragmented mRNA.

FIG. 6 shows the measurement of estrogen receptor mRNA levels in 40 FPE breast cancer specimens via RT-PCR. Three 10 micron sections were used for each measurement. Each data point represents the average of triplicate measurements.

FIG. 7 shows the results of the measurement of progesterone receptor mRNA levels in 40 FPE breast cancer specimens via RT-PCR performed as described in the legend of FIG. 6 above.

FIG. 8 shows results from an IVT/RT-PCR experiment.

FIG. 9 is a representation of the expression of 92 genes across 70 FPE breast cancer specimens. The y-axis shows expression as cycle threshold times. These genes are a subset of the genes listed in Table 1.

Table 1 shows a breast cancer gene list.

Table 2 sets forth amplicon and primer sequences used for amplification of fragmented mRNA.

Table 3 shows the Accession Nos. and SEQ ID NOS of the breast cancer genes examined.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

A. Definitions

Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), and March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, N.Y. 1992), provide one skilled in the art with a general guide to many of the terms used in the present application.

One skilled in the art will recognize many methods and materials similar or equivalent to those described herein, which could be used in the practice of the present invention. Indeed, the present invention is in no way limited to the methods and materials described. For purposes of the present invention, the following terms are defined below.

The term “microarray” refers to an ordered arrangement of hybridizable array elements, preferably polynucleotide probes, on a substrate.

The term “polynucleotide,” when used in singular or plural, generally refers to any polynucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, poly-
nucleotides as defined herein include, without limitation, single- and double-stranded DNA, DNA including single- and double-stranded regions, single- and double-stranded RNA, and RNA including single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or include single- and double-stranded regions. In addition, the term “polynucleotide” as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. The term “polynucleotide” specifically includes DNAs and RNAs that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are “polynucleotides” as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritiated bases, are included within the term “polynucleotides” as defined herein. In general, the term “polynucleotide” embraces all chemically, enzymatically and/or metabolically modified forms of unmodified polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells.

[0280] The term “oligonucleotide” refers to a relatively short polynucleotide, including, without limitation, single-stranded deoxyribonucleotides, single- or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs. Oligonucleotides, such as single-stranded DNA probe oligonucleotides, are often synthesized by chemical methods, for example using automated oligonucleotide synthesizers that are commercially available. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

[0281] The terms “differentially expressed gene,” “differential gene expression” and their synonyms, which are used interchangeably, refer to a gene whose expression is activated to a higher or lower level in a subject suffering from a disease, specifically cancer, such as breast cancer, relative to its expression in a normal or control subject. The terms also include genes whose expression is activated to a higher or lower level at different stages of the same disease. It is also understood that a differentially expressed gene may be either activated or inhibited at the nucleic acid level or protein level, or may be subject to alternative splicing to result in a different polypeptide product. Such differences may be evidenced by a change in mRNA levels, surface expression, secretion or other partitioning of a polypeptide, for example. Differential gene expression may include a comparison of expression between two or more genes, or a comparison of the ratios of the expression between two or more genes, or even a comparison of two differently processed products of the same gene, which differ between normal subjects and subjects suffering from a disease, specifically cancer, or between various stages of the same disease. Differential expression includes both quantitative, as well as qualitative, differences in the temporal or cellular expression pattern in a gene or its expression products among, for example, normal and diseased cells, or among cells which have undergone different disease events or disease stages. For the purpose of this invention, “differential gene expression” is considered to be present when there is at least an about two-fold, preferably at least about four-fold, more preferably at least about six-fold, most preferably at least about ten-fold difference between the expression of a given gene in normal and diseased subjects, or in various stages of disease development in a diseased subject.

[0282] The phrase “gene amplification” refers to a process by which multiple copies of a gene or gene fragment are formed in a particular cell or cell line. The duplicated region (a stretch of amplified DNA) is often referred to as “ampli-con.” Usually, the amount of the messenger RNA (mRNA) produced, i.e., the level of gene expression, also increases in the proportion of the number of copies made of the particular gene expressed.

[0283] The term “prognosis” is used herein to refer to the prediction of the likelihood of cancer-attributable death or progression, including recurrence, metastatic spread, and drug resistance, of a neoplastic disease, such as breast cancer. The term “prediction” is used herein to refer to the likelihood that a patient will respond either favorably or unfavorably to a drug or set of drugs, and also the extent of those responses. The predictive methods of the present invention can be used clinically to make treatment decisions by choosing the most appropriate treatment modalities for any particular patient. The predictive methods of the present invention are valuable tools in predicting if a patient is likely to respond favorably to a treatment regimen, such as surgical intervention, chemotherapy with a given drug or drug combination, and/or radiation therapy.

[0284] The term “increased resistance” to a particular drug or treatment option, when used in accordance with the present invention, means decreased response to a standard dose of the drug or to a standard treatment protocol.

[0285] The term “decreased sensitivity” to a particular drug or treatment option, when used in accordance with the present invention, means decreased response to a standard dose of the drug or to a standard treatment protocol, where decreased response can be compensated for (at least partially) by increasing the dose of drug, or the intensity of treatment.

[0286] “Patient response” can be assessed using any endpoint indicating a benefit to the patient, including, without limitation, (1) inhibition, to some extent, of tumor growth, including slowing down and complete growth arrest; (2) reduction in the number of tumor cells; (3) reduction in tumor size; (4) inhibition (i.e., reduction, slowing down or complete stopping) of tumor cell infiltration into adjacent peripheral organs and/or tissues; (5) inhibition (i.e. reduction, slowing down or complete stopping) of metastasis; (6) enhancement of anti-tumor immune response, which may, but does not have to, result in the regression or rejection of the tumor; (7) relief, to some extent, of one or more symptoms associated with the tumor; (8) increase in the length of survival following treatment; and/or (9) decreased mortality at a given point of time following treatment.

[0287] The term “treatment” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of
treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. In tumor (e.g., cancer) treatment, a therapeutic agent may directly decrease the pathology of tumor cells, or render the tumor cells more susceptible to treatment by other therapeutic agents, e.g., radiation and/or chemotherapy.

[0288] The term “tumor,” as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

[0289] The terms “cancer” and “cancerous” refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include but are not limited to, breast cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, and brain cancer.

[0290] The “pathology” of cancer includes all phenomena that compromise the well-being of the patient. This includes, without limitation, abnormal or uncontrollable cell growth, metastasis, interference with the normal functioning of neighboring cells, release of cytokines or other secretory products at abnormal levels, suppression or aggravation of inflammatory or immunological response, neoplasia, premalignancy, malignancy, invasion of surrounding or distant tissues or organs, such as lymph nodes, etc.

[0291] “Stringency” of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

[0292] “Stringent conditions” or “high stringency conditions”, as defined herein, typically: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50° C.; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42° C.; or (3) employ 50% formamide, 5xSSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5x Denhardt’s solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42° C., with washes at 42° C. in 0.2xSSC (sodium chloride/sodium citrate) and 50% formamide at 55° C., followed by a high-stringency wash consisting of 0.1xSSC containing EDTA at 55° C.

[0293] “Moderately stringent conditions” may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and % SDS) less stringent that those described above. An example of moderately stringent conditions is overnight incubation at 37° C. in a solution comprising: 20% formamide, 5xSSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt’s solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1xSSC at about 37-50° C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like. In the context of the present invention, reference to “at least one,” “at least two,” “at least five,” etc. of the genes listed in any particular gene set means any one or any and all combinations of the genes listed.

[0294] The terms “splicing” and “RNA splicing” are used interchangeably and refer to RNA processing that removes introns and joins exons to produce mature mRNA with continuous coding sequence that moves into the cytoplasm of an eukaryotic cell.

[0295] In theory, the term “exon” refers to any segment of an interrupted gene that is represented in the mature RNA product (B. Lewin. Genes IV, Cell Press, Cambridge Mass. 1990). In theory the term “intron” refers to any segment of DNA that is transcribed but removed from within the transcript by splicing together the exons on either side of it. Operationally, exon sequences occur in the mRNA sequence of a gene as defined by Ref. Seq ID numbers. Operationally, intron sequences are the intervening sequences within the genomic DNA of a gene, bracketed by exon sequences and having GT and AG splice consensus sequences at their 5′ and 3′ boundaries.

B. Detailed Description


[0297] 1. Gene Expression Profiling

[0298] In general, methods of gene expression profiling can be divided into two large groups: methods based on hybridization analysis of polynucleotides, and methods based on sequencing of polynucleotides. The most com-
monly used methods known in the art for the quantification of mRNA expression in a sample include northern blotting and in situ hybridization (Parker & Barnes, *Methods in Molecular Biology* 106:247-283 (1999)); RNAse protection assays (Hod, *Biotechniques* 13:852-854 (1992)); and reverse transcription polymerase chain reaction (RT-PCR) (Weis et al., *Trends in Genetics* 8:263-264 (1992)). Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. Representative methods for sequencing-based gene expression analysis include Serial Analysis of Gene Expression (SAGE), and gene expression analysis by massively parallel signature sequencing (MPSS).

2. Reverse Transcriptase PCR (RT-PCR)

Of the techniques listed above, the most sensitive and most flexible quantitative method is RT-PCR, which can be used to compare mRNA levels in different sample populations, in normal and tumor tissues, with or without drug treatment, to characterize patterns of gene expression, to discriminate between closely related mRNAs, and to analyze RNA structure.

The first step is the isolation of mRNA from a target sample. The starting material is typically total RNA isolated from human tumors or tumor cell lines, and corresponding normal tissues or cell lines, respectively. Thus RNA can be isolated from a variety of primary tumors, including breast, lung, colon, prostate, brain, liver, kidney, pancreas, spleen, thymus, testis, ovary, uterus, etc., tumor or tumor cell lines, with pooled DNA from healthy donors. If the source of mRNA is a primary tumor, mRNA can be extracted, for example, from frozen or archived paraffin-embedded and fixed (e.g. formalin-fixed) tissue samples.

General methods for mRNA extraction are well known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel et al., *Current Protocols of Molecular Biology*, John Wiley and Sons (1997). Methods for RNA extraction from paraffin-embedded tissues are disclosed, for example, in Rupp and Locker, *Lab Invest.* 56:A67 (1987), and De Andries et al., *BioTechniques* 18:42044 (1995). In particular, RNA isolation can be performed using purification kits, buffer set and protease from commercial manufacturers, such as Qiagen; according to the manufacturer’s instructions. For example, total RNA from cells in culture can be isolated using Qiagen RNaseasy minicolumns. Other commercially available RNA isolation kits include MasterPure™ Complete DNA and RNA Purification Kit (EPICENTRE®, Madison, Wis.), and Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples can be isolated using RNA Stat-60 (Tel-Test). RNA prepared from tumor can be isolated, for example, by cesium chloride density gradient centrifugation.

As RNA cannot serve as a template for PCR, the first step in gene expression profiling by RT-PCR is the reverse transcription of the RNA template into cDNA, followed by its exponential amplification in a PCR reaction. The two most commonly used reverse transcriptases are avian myeloblastosis virus reverse transcriptase (AMV-RT) and Moloney murine leukemia virus reverse transcriptase (MMLV-RT). The reverse transcription step is typically primed using specific primers, random hexamers, or oligo-dT primers, depending on the circumstances and the goal of expression profiling. For example, extracted RNA can be reverse-transcribed using a GeneAmp RNA PCR kit (Perkin Elmer, CA, USA), following the manufacturer’s instructions. The derived cDNA can then be used as a template in the subsequent PCR reaction.

Although the PCR step can use a variety of thermostable DNA-dependent DNA polymerases, it typically employs the Taq DNA polymerase, which has a 5'-3' nucleic acid activity but lacks a 3'-5' proofreading endonuclease activity. Thus, TaqMan®PCR typically utilizes the 5'-nucleic acid activity of Taq or Th polymerase to hydrolyze a hybridization probe bound to its target amplicon, but any enzyme with equivalent 5'-nucleic acid activity can be used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

TaqMan® RT-PCR can be performed using commercially available equipment, such as, for example, ABI PRISM 7700™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, Calif., USA), or LightCycler (Roche Molecular Biochemicals, Mannheim, Germany). In a preferred embodiment, the 5'-nucleic acid procedure is run on a real-time quantitative PCR device such as the ABI PRISM 7700™ Sequence Detection System™. The system consists of a thermocycler, laser, charge-coupled device (CCD), camera and computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

5'-Nucleic assay data are initially expressed as Ct, or the threshold cycle. As discussed above, fluorescence values are recorded during every cycle and represent the amount of product amplified to that point in the amplification reaction. The point when the fluorescent signal is first recorded as statistically significant is the threshold cycle (Ct).

To minimize errors and the effect of sample-to-sample variation, RT-PCR is usually performed using an internal standard. The ideal internal standard is expressed at a constant level among different tissues, and is unaffected by the experimental treatment. RNAs most frequently used to normalize patterns of gene expression are mRNAs for the housekeeping genes glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) and β-actin.

A more recent variation of the RT-PCR technique is the real time quantitative PCR, which measures PCR
product accumulation through a dual-labeled fluorogenic probe (i.e., TaqMan® probe). Real time PCR is compatible both with quantitative competitive PCR, where internal competitor for each target sequence is used for normalization, and with quantitative comparative PCR using a normalization gene contained within the sample, or a housekeeping gene for RT-PCR. For further details see, e.g., Held et al., Genome Research 6:986-994 (1996).

(0309) 3. Microarrays

(0310) Differential gene expression can also be identified, or confirmed using the microarray technique. Thus, the expression profile of breast cancer-associated genes can be measured in either fresh or paraffin-embedded tumor tissue, using microarray technology. In this method, polynucleotide sequences of interest are plated, or arrayed, on a microchip substrate. The arrayed sequences are then hybridized with specific DNA probes from cells or tissues of interest. Just as in the RT-PCR method, the source of mRNA typically is total RNA isolated from human tumors or tumor cell lines, and corresponding normal tissues or cell lines. Thus RNA can be isolated from a variety of primary tumors or tumor cell lines. If the source of mRNA is a primary tumor, mRNA can be extracted, for example, from frozen or archived paraffin-embedded and fixed (e.g. formalin-fixed) tissue samples, which are routinely prepared and preserved in everyday clinical practice.

(0311) In a specific embodiment of the microarray technique, PCR amplified inserts of cDNA clones are applied to a substrate in a dense array. Preferably at least 10,000 nucleotide sequences are applied to the substrate. The microarrayed genes, immobilized on the microchip at 10,000 elements each, are suitable for hybridization under stringent conditions. Fluorescently labeled cDNA probes may be generated through incorporation of fluorescent nucleotides by reverse transcription of RNA extracted from tissues of interest. Labeled cDNA probes applied to the chip hybridize with specificity to each spot of DNA on the array. After stringent washing to remove non-specifically bound probes, the chip is scanned by confocal laser microscopy or by another detection method, such as a CCD camera. Quantification of hybridization of each arrayed element allows for assessment of corresponding mRNA abundance. With dual color fluorescence, separately labeled cDNA probes generated from two sources of RNA are hybridized pairwise to the array. The relative abundance of the transcripts from the two sources corresponding to each specified gene is thus determined simultaneously. The miniaturized scale of the hybridization affords a convenient and rapid evaluation of the expression pattern for large numbers of genes. Such methods have been shown to have the sensitivity required to detect rare transcripts, which are expressed at a few copies per cell, and to reproducibly detect at least approximately two-fold differences in the expression levels (Scheda et al., Proc. Natl. Acad. Sci. USA 93(2):106-149 (1996)). Microarray analysis can be performed by commercially available equipment, following manufacturer’s protocols, such as by using the Affymetrix GeneChip technology, or Incyte’s microarray technology.

(0312) The development of microarray methods for large-scale analysis of gene expression makes it possible to search systematically for molecular markers of cancer classification and outcome prediction in a variety of tumor types.

(0313) 4. Serial Analysis of Gene Expression (SAGE)

(0314) Serial analysis of gene expression (SAGE) is a method that allows the simultaneous and quantitative analysis of a large number of gene transcripts, without the need of providing an individual hybridization probe for each transcript. First, a short sequence tag (about 10-14 bp) is generated that contains sufficient information to uniquely identify a transcript, provided that the tag is obtained from a unique position within each transcript. Then, many transcripts are linked together to form long serial molecules, that can be sequenced, revealing the identity of the multiple tags simultaneously. The expression pattern of any population of transcripts can be quantitatively evaluated by determining the abundance of individual tags, and identifying the gene corresponding to each tag. For more details see, e.g., Velculescu et al., Science 270:484-487 (1995); and Velculescu et al., Cell 88:243-51 (1997).

(0315) 5. Gene Expression Analysis by Massively Parallel Signature Sequencing (MPSs)

(0316) This method, described by Brenner et al., Nature Biotechnology 18:630-634 (2000), is a sequencing approach that combines non-gel-based signature sequencing with in vitro cloning of millions of templates on separate 5 µm diameter microbeads. First, a microbead library of DNA templates is constructed by in vitro cloning. This is followed by the assembly of a planar array of the template-containing microbeads in a flow cell at a high density (typically greater than 3x10^8 microbeads/cm²). The free ends of the cloned templates on each microbead are analyzed simultaneously, using a fluorescence-based signature sequencing method that does not require DNA fragment separation. This method has been shown to simultaneously and accurately provide, in a single operation, hundreds of thousands of gene signature sequences from a yeast cDNA library.

(0317) 6. General Description of the mRNA Isolation, Purification and Amplification Methods of the Invention

(0318) The steps of a representative protocol of the invention, including mRNA isolation, purification, primer extension and amplification are illustrated in FIG. 1. As shown in FIG. 1, this representative process starts with cutting about 10 µm thick sections of paraffin-embedded tumor tissue samples. The RNA is then extracted, and protein and DNA are removed, following the method of the invention described below. After analysis of the RNA concentration, RNA repair and/or amplification steps may be included, if necessary, and RNA is reverse transcribed using gene specific promoters followed by RT-PCR. Finally, the data are analyzed to identify the best treatment option(s) available to the patient on the basis of the characteristic gene expression pattern identified in the tumor sample examined. The individual steps of this protocol will be discussed in greater detail below.

(0319) 7. Improved Method for Isolation of Nucleic Acid from Archived Tissue Specimens

(0320) As discussed above, in the first step of the method of the invention, total RNA is extracted from the source material of interest, including fixed, paraffin-embedded tissue specimens, and purified sufficiently to act as a substrate in an enzyme assay. Despite the availability of commercial products, and the extensive knowledge available concerning the isolation of nucleic acid, such as RNA, from tissues,
isolation of nucleic acid (RNA) from fixed, paraffin-embedded tissue specimens (FPET) is not without difficulty. In one aspect, the present invention concerns an improved method for the isolation of nucleic acid from archived, e.g. FPET tissue specimens. Measured levels of mRNA species are useful for defining the physiological or pathological status of cells and tissues. RT-PCR (which is discussed above) is one of the most sensitive, reproducible and quantitative methods for this "gene expression profiling". Paraffin-embedded, formalin-fixed tissue is the most widely available material for such studies. Several laboratories have demonstrated that it is possible to successfully use fixed-paraffin-embedded tissue (FPET) as a source of RNA for RT-PCR (Stanta et al., Biotecniques 11:304-308 (1991); Stanta et al., Methods Mol. Biol. 86:23-26 (1998); Jackson et al., Lancet 1:1391 (1989); Jackson et al., J. Clin. Pathol. 43:499-504 (1990); Finke et al., Biotecniques 14:448-453 (1993); Goldsworthy et al., Mol. Carcinog. 25:86-91 (1999); Stanta and Bonin, Biotecniques 24:271-276 (1998); Godfrey et al., J. Mol. Diagnostics. 2:84 (2000); Specht et al., J. Mol. Med. 78:327 (2000); Specht et al., Am. J. Pathol. 158:419-429 (2001)). This allows gene expression profiling to be carried out on the most commonly available source of human biopsy specimens, and therefore potentially to create new valuable diagnostic and therapeutic information.

The most widely used protocols utilize hazardous organic solvents, such as xylene, or octane (Finke et al., supra) to dewax the tissue in the paraffin blocks before nucleic acid (RNA and/or DNA) extraction. Obligatory organic solvent removal (e.g., with ethanol) and rehydration steps follow, which necessitate multiple manipulations, and addition of substantial total time to the protocol, which can take up to several days. Commercial kits and protocols for RNA extraction from FPET [MasterPure™, Complete DNA and RNA Purification Kit (EPICENTRE®, Madison, Wis.); Paraffin Block RNA Isolation Kit (Ambion, Inc.) and RNeasy™ Mini kit (Qiagen, Chatsworth, Calif.)] use xylene for deparaffinization, in procedures which typically require multiple centrifugations and ethanol buffer changes, and incubations following incubation with xylene.

The present invention provides an improved nucleic acid extraction protocol that produces nucleic acid, in particular RNA, sufficiently intact for gene expression measurements. The key step in the nucleic acid extraction protocol herein is the performance of de-waxing without the use of any organic solvent, thereby eliminating the need for multiple manipulations associated with the removal of the organic solvent, and substantially reducing the total time to the protocol. According to the invention, wax, e.g., paraffin is removed from wax-embedded tissue samples by incubation at 65-75°C in a lysis buffer that solubilizes the tissue and hydrolyzes the protein, following by cooling to solidify the wax.

FIG. 2 shows a flow chart of an RNA extraction protocol of the present invention in comparison with a representative commercial method, using xylene to remove wax. The times required for individual steps in the processes and for the overall processes are shown in the chart. As shown, the commercial process requires approximately 50% more time than the process of the invention.

The lysis buffer can be any buffer known for cell lysis. It is, however, preferred that oligo-dT-based methods of selectively purifying polyadenylated mRNA not be used to isolate RNA for the present invention, since the bulk of the mRNA molecules are expected to be fragmented and therefore will not have an intact polyadenylated tail, and will not be recovered or available for subsequent analytical assays. Otherwise, any number of standard nucleic acid purification schemes can be used. These include chaotrope and organic solvent extractions, extraction using glass beads or filters, salting out and precipitation based methods, or any of the purification methods known in the art to recover total RNA or total nucleic acids from a biological source.

Lysis buffers are commercially available, such as, for example, from Qiagen, Epicentre, or Ambion. A preferred group of lysis buffers typically contains urea, and Proteinase K or other protease. Proteinase K is very useful in the isolation of high quality, undamaged DNA or RNA, since most mammalian DNases and RNases are rapidly inactivated by this enzyme, especially in the presence of 0.5-1% sodium dodecyl sulfate (SDS). This is particularly important in the case of RNA, which is more susceptible to degradation than DNA. While DNases require metal ions for activity, and can therefore be easily inactivated by chelating agents, such as EDTA, there is no similar co-factor requirement for RNases.

Cooling and resultant solidification of the wax permits easy separation of the wax from the total nucleic acid, which can be conveniently precipitated, e.g. by isopropanol. Further processing depends on the intended purpose. If the proposed method of RNA analysis is subject to bias by contaminating DNA in an extract, the RNA extract can be further treated, e.g. by DNase, post purification to specifically remove DNA while preserving RNA. For example, if the goal is to isolate high quality RNA for subsequent RT-PCR amplification, nucleic acid precipitation is followed by the removal of DNA, usually by DNase treatment. However, DNA can be removed at various stages of nucleic acid isolation, by DNase or other techniques well known in the art.

While the advantages of the nucleic acid extraction protocol of the invention are most apparent for the isolation of RNA from archived, paraffin embedded tissue samples, the wax removal step of the present invention, which does not involve the use of an organic solvent, can also be included in any conventional protocol for the extraction of total nucleic acid (RNA and DNA) or DNA only. All of these aspects are specifically within the scope of the invention.

By using heat followed by cooling to remove paraffin, the process of the present invention saves valuable processing time, and eliminates a series of manipulations, thereby potentially increasing the yield of nucleic acid. Indeed, experimental evidence presented in the examples below, demonstrates that the method of the present invention does not compromise RNA yield.

8. 5'-Multiplexed Gene Specific Priming of Reverse Transcription

RT-PCR requires reverse transcription of the test RNA population as a first step. The most commonly used primer for reverse transcription is oligo-dT, which works well when RNA is intact. However, this primer will not be effective when RNA is highly fragmented as is the case in FPET tissues.
The present invention includes the use of gene specific primers, which are roughly 20 bases in length with a Tm optimum between about 58° C. and 60° C. These primers will also serve as the reverse primers that drive PCR DNA amplification.

Another aspect of the invention is the inclusion of multiple gene-specific primers in the same reaction mixture. The number of such different primers can vary greatly and can be as low as two and as high as 40,000 or more. Table 2 displays examples of reverse primers that can be successfully used in carrying out the methods of the invention. FIG. 9 shows expression data obtained using this multiplexed gene-specific priming strategy. Specifically, FIG. 9 is a representation of the expression of 92 genes (a subset of genes listed in Table 1) across 70 FPE breast cancer specimens. The y-axis shows expression as cycle threshold times.

An alternative approach is based on the use of random hexamers as primers for cDNA synthesis. However, we have experimentally demonstrated that the method of using a multiplicity of gene-specific primers is superior over the known approach using random hexamers.

9. Preparation of Fragmented mRNA for Expression Profiling Assays

It is of interest to analyze the abundance of specific mRNA species in biological samples, since this expression profile provides an index of the physiological state of the sample. mRNA is notoriously difficult to extract and maintain in its native state, consequently, mRNA recovered from biological sources is often fragmented or somewhat degraded. This is especially true of human tissue specimens which have been chemically fixed and stored for extended periods of time.

In one aspect, the present invention provides a means of preparing the mRNA extracted from various sources, including archived tissue specimens, for expression profiling in a way that its relative abundance is preserved and the mRNA's interest can be successfully measured. This method is useful as a means of preparing mRNA for analysis by any of the known expression profiling methods, including RT-PCR coupled with 5' exonuclease of reporter probes (TaqMan® type assays), as discussed above, flap endonuclease assays (Clearave® and Invader® type assays), oligonucleotide hybridization arrays, cDNA hybridization arrays, oligonucleotide ligation assays, 3' single nucleotide extension assays and other assays designed to assess the abundance of specific mRNA sequences in a biological sample.

According to the method of the invention, total RNA is extracted from the source material and sufficiently purified to act as a substrate in an enzyme assay. The extraction procedure, including a new and improved way of removing the wax (e.g. paraffin) used for embedding the tissue samples, has been discussed above. It has also been noted that it is preferred that oligo-dT based methods of selectively purifying polyadenylated mRNA not be used to isolate RNA for this invention since the bulk of the mRNA is expected to be fragmented, will not be polyadenylated and, therefore, will not be recovered and available for subsequent analytical assays if an oligo-dT based method is used.

A diagram of an improved method for repairing fragmented RNA is shown in FIG. 3. The fragmented RNA purified from the tissue sample is mixed with universal or gene-specific, single-stranded, DNA templates for each mRNA species of interest. These templates may be full length DNA copies of the mRNA derived from cloned gene sources, they may be fragments of the gene representing only the segment of the gene to be assayed, they may be a series of long oligonucleotides representing either the full length gene or the specific segment(s) of interest. The template can represent either a single consensus sequence or be a mixture of polymorphic variants of the gene. This DNA template, or scaffold, will preferably include one or more dUTP or rNTP sites in its length. This will provide a means of removing the template prior to carrying out subsequent analytical steps to avoid its acting as a substrate or target in later analysis assays. This removal is accomplished by treating the sample with uracil-DNA glycosylase (UDG) and heating it to cause strand breaks where UDG has generated abasic sites. In the case of rNTP's, the sample can be heated in the presence of a basic buffer (pH~10) to induce strand breaks where rNTP's are located in the template.

The single stranded DNA template is mixed with the purified RNA, the mixture is denatured and annealed so that the RNA fragments complementary to the DNA template effectively become primers that can be extended along the single stranded DNA templates. DNA polymerase 1 requires a primer for extension but will efficiently use either a DNA or an RNA primer. Therefore in the presence of DNA polymerase 1 and dNTP's, the fragmented RNA can be extended along the complementary DNA templates. In order to increase the efficiency of the extension, this reaction can be thermally cycled, allowing overlapping templates and extension products to hybridize and extend until the overall population of fragmented RNA becomes represented as double stranded DNA extended from RNA fragment primers.

Following the generation of this “repaired” RNA, the sample should be treated with UDG or heat-treated in a mildly based solution to fragment the DNA template (scaffold) and prevent it from participating in subsequent analytical reactions.

The product resulting from this enzyme extension can then be used as a template in a standard enzyme profiling assay that includes amplification and detectable signal generation such as fluorescent, chemiluminescent, colorimetric or other common read outs from enzyme based assays. For example, for TaqMan® type assays, this double stranded DNA product is added as the template in a standard assay; and, for array hybridization, this product acts as the cDNA template for the cRNA labelling reaction typically used to generate single-stranded, labeled RNA for array hybridization.

This method of preparing template has the advantage of recovering information from mRNA fragments too short to effectively act as templates in standard cDNA generation schemes. In addition, this method acts to preserve the specific locations in mRNA sequences targeted by specific analysis assays. For example, TaqMan® type assays rely on a single contiguous sequence in a cDNA copy of mRNA to act as a PCR amplification template targeted by a labeled reporter probe. If mRNA strand breaks occur in this sequence, the assay will not detect that template and will...
underestimate the quantity of that mRNA in the original sample. This target preparation method minimizes that effect from RNA fragmentation.

[0344] The extension product formed in the RNA primer extension assay can be controlled by controlling the input quantity of the single stranded DNA template and by doing limited cycling of the extension reaction. This is important in preserving the relative abundance of the mRNA sequences targeted for analysis.

[0345] This method has the added advantage of not requiring parallel preparation for each target sequence since it is easily multiplexed. It is also possible to use large pools of random sequence long oligonucleotides or full libraries of cloned sequences to extend the entire population of mRNA sequences in the sample extract for whole expressed genome analysis rather than targeted gene specific analysis.

[0346] 10. Amplification of mRNA Species Prior to RT-PCR

[0347] Due to the limited amount and poor quality of mRNA that can be isolated from FFPE, a new procedure that could accurately amplify mRNAs of interest would be very useful, particularly for real time quantitation of gene expression (TaqMan®) and especially for quantitatively large number (>50) of genes>50 to 10,000.

[0348] Current protocols (e.g. Eberwine, Biotechniques 20:584-91 (1996)) are optimized for mRNA amplification from small amount of total or poly A+ RNA mainly for microarray analysis. The present invention provides a protocol optimized for amplification of small amounts of fragmented total RNA (average size about 60-150 bps), utilizing gene-specific sequences as primers, as illustrated in FIG. 4.

[0349] The amplification procedure of the invention uses a very large number, typically as many as 100-190,000 gene specific primers (GSP’s) in one reverse transcription run. Each GSP contains an RNA polymerase promoter, e.g. a T7 DNA-dependent RNA polymerase promoter, at the 5’ end for subsequent RNA amplification. GSP’s are preferred as primers because of the small size of the RNA. Current protocols utilize dT primers, which would not adequately represent all reverse transcripts of mRNAs due to the small size of the FFPE RNA. GSP’s can be designed by optimizing usual parameters, such as length, Tm, etc. For example, GSP’s can be designed using the Primer Express® (Applied Biosystems), or Primer 3 (MIT) software programs. Typically, at least 3 sets per gene are designed, and the ones giving the lowest Ct on FFPE RNA (best performers) are selected.

[0350] Second strand cDNA synthesis is performed by standard procedures (see FIG. 4, Method 1), or by GSP primers and Taq pol under PCR conditions (e.g., 95° C., 10 min (Taq activation) then 60° C., 45 sec). The advantages of the latter method are that the second gene specific primer, SG2, adds additional specificity (and potentially more efficient second strand synthesis) and the option of performing several cycles of PCR, if more starting DNA is necessary for RNA amplification by T7 RNA polymerase. RNA amplification is then performed under standard conditions to generate multiple copies of cRNA, which is then used in a standard TaqMan® reaction.

[0351] Although this process is illustrated by using T7-based RNA amplification, a person skilled in the art will understand that other RNA polymerase promoters that do not require a primer, such as T3 or Sp6 can also be used, and are within the scope of the invention.

[0352] 11. A Method of Elongation of Fragmented RNA and Subsequent Amplification

[0353] This method, which combines and modifies the inventions described in sections 9 and 10 above, is illustrated in FIG. 5. The procedure begins with elongation of fragmented mRNA. This occurs as described above except that the scaffold DNAs are tagged with the T7 RNA polymerase promoter sequence at their 5’ ends, leading to double-stranded DNA extended from RNA fragments. The template sequences need to be removed after in vitro transcription. These templates can include dUTP or rNTP nucleotides, enabling enzymatic removal of the templates as described in section 9, or the templates can be removed by DNaseI treatment.

[0354] The template DNA can be a population representing different mRNAs of any number. A high sequence complexity source of DNA templates (scaffolds) can be generated by pooling RNA from a variety of cells or tissues. In one embodiment, these RNAs are converted into double stranded DNA and cloned into phagemids. Single stranded DNA can then be rescued by phagemid growth and single stranded DNA isolation from purified phagemids.

[0355] This invention is useful because it increases gene expression profile signals two different ways: both by increasing test mRNA polynucleotide sequence length and by in vitro transcription amplification. An additional advantage is that it eliminates the need to carry out reverse transcription optimization with gene specific primers tagged with the T7 RNA polymerase promoter sequence, and thus, is comparatively fast and economical.

[0356] This invention can be used with a variety of different methods to profile gene expression, e.g., RT-PCR or a variety of DNA array methods. Just as in the previous protocol, this approach is illustrated by using a T7 promoter but the invention is not so limited. A person skilled in the art will appreciate, however, that other RNA polymerase promoters, such as T3 or Sp6 can also be used.


[0358] An important aspect of the present invention is to use the measured expression of certain genes by breast cancer tissue to match patients to best drugs or drug combinations, and to provide prognostic information. For this purpose it is necessary to correct for (normalize away) both differences in the amount of RNA assayed and variability in the quality of the RNA used. Therefore, the assay measures and incorporates the expression of certain normalizing genes, including well known housekeeping genes, such as GAPDH and Cyp1. Alternatively, normalization can be based on the mean or median signal (Ct) of all of the assayed genes or a large subset thereof (global normalization approach). On a gene-by-gene basis, measured normalized amount of a patient tumor mRNA is compared to the amount found in a breast cancer tissue reference set. The number (N) of breast cancer tissues in this reference set should be sufficiently high to ensure that different reference sets (as a whole) behave essentially the same way. If this condition is
met, the identity of the individual breast cancer tissues present in a particular set will have no significant impact on the relative amounts of the genes assayed. Usually, the breast cancer tissue reference set consists of at least about 30, preferably at least about 40 different FPE breast cancer tissue specimens. Unless noted otherwise, normalized expression levels for each mRNA/tested tumor/patient will be expressed as a percentage of the expression level measured in the reference set. More specifically, the reference set of a sufficiently high number (e.g., 40) tumors yields a distribution of normalized levels of each mRNA species. The level measured in a particular tumor sample to be analyzed falls at some percentile within this range, which can be determined by methods well known in the art. Below, unless noted otherwise, reference to expression levels of a gene assume normalized expression relative to the reference set although this is not always explicitly stated.

0359 The breast cancer gene set is shown in Table 1. The gene Accession Numbers, and the SEQ ID Nos. for the forward primer, reverse primer and amplicon sequences that can be used for gene amplification, are listed in Table 2. The basis for inclusion of markers, as well as the clinical significance of mRNA level variations with respect to the reference set, is indicated below. Genes are grouped into subsets based on the type of clinical significance indicated by their expression levels: A. Prediction of patient response to drugs used in breast cancer treatment, or to drugs that are approved for other indications and could be used off-label in the treatment of breast cancer. B. Prognostic for survival or recurrence of cancer.

C. Prediction of Patient Response to Therapeutic Drugs

0360 Molecules that Specifically Influence Cellular Sensitivity to Drugs

0361 Table 1 lists 74 genes (shown in italics) that specifically influence cellular sensitivity to potent drugs, which are also listed. Most of the drugs shown are approved and already used to treat breast cancer (e.g., anthracyclines; cyclophosphamide; methotrexate; 5-FU and analogues). Several of the drugs are used to treat breast cancer off-label or are in clinical development phase (e.g., bisphosphonates and anti-VEGF mAb). Several of the drugs have not been widely used to treat breast cancer but are used in other cancers in which the indicated target is expressed (e.g., Celebrex is used to treat familial colon cancer; cisplatin is used to treat ovarian and other cancers.)

0362 Patient response to 5FU is indicated if normalized thymidylate synthase mRNA amount is at or below the 15th percentile, or the sum of expression of thymidylate synthase plus dihydropyrimidine phosphorylase is at or below the 25th percentile, or the sum of expression of these mRNAs plus thymidylate phosphorylase is at or below the 20th percentile. Patients with dihydropyrimidine dehydrogenase below 5th percentile are at risk of adverse response to 5FU, or analogs such as Xeloda.

0363 When levels of thymidylate synthase, and dihydropyrimidine dehydrogenase, are within the acceptable range as defined in the preceding paragraph, amplification of c-myc mRNA in the upper 15%, against a background of wild-type p53 [as defined below] predicts a beneficial response to 5FU (see D. Arango et al., Cancer Res. 61:4910-4915 (2001)). In the presence of normal levels of thymidylate synthase and dihydropyrimidine dehydrogenase, levels of NFκB and cIAP2 in the upper 10% indicate resistance of breast tumors to the chemotherapeutic drug 5FU.

0364 Patient resistance to anthracyclines is indicated if the normalized mRNA level of topoisomerase IIα is below the 10th percentile, or the topoisomerase IIα normalized mRNA level is below the 10th percentile or if the combined normalized topoisomerase IIα and β signals are below the 10th percentile.

0365 Patient sensitivity to methotrexate is compromised if DHFR levels are more than tenfold higher than the average reference set level for this mRNA species, or if reduced folate carrier levels are below 10th percentile.

0366 Patients whose tumors express CYP1B1 in the upper 10%, have reduced likelihood of responding to docetaxel.

0367 The sum of signals for aldehyde dehydrogenase 1A1 and 1A3, when more than tenfold higher than the reference set average, indicates reduced likelihood of response to cyclophosphamide.

0368 Currently, estrogen and progesterone receptor expression as measured by immunohistochemistry is used to select patients for anti-estrogen therapy. We have demonstrated RT-PCR assays for estrogen and progesterone receptor mRNA levels that predict levels of these proteins as determined by a standard clinical diagnostic tests, with high degree of concordance (FIGS. 6 and 7).

0369 Patients whose tumors express ERα or PR mRNA in the upper 70%, are likely to respond to tamoxifen or other anti-estrogens (thus, operationally, lower levels of ERα than this are to defined ERα-negative). However, when the signal for microsomal epoxide hydrolase is in the upper 10% or when mRNAs for p53/trefoil factor, GATA3 or human chorionic gonadotropin are at or below average levels found in ERα-negative tumors, anti-estrogen therapy will not be beneficial.

0370 Absence of XIST signal compromises the likelihood of response to taxanes, as does elevation of the GST-π or prolyl endopeptidase [PREP] signal in the upper 10%. Elevation of PLAG1 in the upper 10% decreases sensitivity to taxanes.

0371 Expression of ERCC1 mRNA in the upper 10% indicate significant risk of resistance to cisplatin or analogs.

0372 An RT-PCR assay of Her2 mRNA expression predicts Her2 overexpression as measured by a standard diagnostic test, with high degree of concordance (data not shown). Patients whose tumors express Her2 (normalized to cyp.1) in the upper 10% have increased likelihood of beneficial response to treatment with Herceptin or other ErbB2 antagonists. Measurement of expression of Grb7 mRNA serves as a test for HER2 gene amplification, because the Grb7 gene is closely linked to Her2. When Her2 is expression is high as defined above in this paragraph, similarly elevated Grb7 indicates Her2 gene amplification. Overexpression of IGFR1 and or IGFl or IGF2 decreases likelihood of beneficial response to Herceptin and also to EGFR antagonists.

0373 Patients whose tumors express mutant Ha-Ras, and also express farnesyl pyrophosphate synthetase or geranyl
pyrophosphonate synthetase mRNAs at levels above the tenth percentile comprise a group that is especially likely to exhibit a beneficial response to bis-phosphonate drugs.

[0374] Cox2 is a key control enzyme in the synthesis of prostaglandins. It is frequently expressed at elevated levels in subsets of various types of carcinomas including carcinoma of the breast. Expression of this gene is controlled at the transcriptional level, so RT-PCR serves a valid indicator of the cellular enzyme activity. Nonclinical research has shown that cox2 promotes tumor angiogenesis, suggesting that this enzyme is a promising drug target in solid tumors. Several Cox2 antagonists are marketed products for use in anti-inflammatory conditions. Treatment of familial adenomatous polyposis patients with the cox2 inhibitor Celebrex significantly decreased the number and size of neoplastic polyps. No cox2 inhibitor has yet been approved for treatment of breast cancer, but generally this class of drugs is safe and could be prescribed off-label in breast cancers in which cox2 is over-expressed. Tumors expressing COX2 at levels in the upper ten percent have increased chance of beneficial response to Celebrex or other cyclooxygenase 2 inhibitors.

[0375] The tyrosine kinases ErbB1 (EGFR), ErbB3 (Her3) and ErbB4 (Her4); also the ligands TGFi alpha, amphiregulin, heparin-binding EGF-like growth factor, and epiregulin; also BRK, a non-receptor kinase. Several drugs in clinical development block the EGFR receptor: ErbB2-4, the indicated ligands, and BRK also increase the activity of the EGFR pathway. Breast cancer patients whose tumors express high levels of EGFR or EGFR and abnormally high levels of the other indicated activators of the EGFR pathway are potential candidates for treatment with an EGFR antagonist.

[0376] Patients whose tumors express less than 10% of the average level of EGFR mRNA observed in the reference panel are relatively less likely to respond to EGFR antagonists such as Iressa, or ImClone 225. In cases in which the EGFR is above this low range, the additional presence of epiregulin, TGFialpha, amphiregulin, or ErbB3, or BRK, CD9, MMP9, or L01 at levels above the 90th percentile predisposes to response to EGFR antagonists. Epiregulin gene expression, in particular, is a good surrogate marker for EGFR activation, and can be used to not only to predict response to EGFR antagonists, but also to monitor response to EGFR antagonists [taking fine needle biopsies to provide tumor tissue during treatment]. Levels of CD82 above the 90th percentile suggest poorer efficacy from EGFR antagonists.

[0377] The tyrosine kinases abl, c-kit, PDGFRalpha, PDGFRbeta, and ARG; also, the signal transmitting ligands c-kit ligand, PDGFA, B, C and D. The listed tyrosine kinases are all targets of the drug Gleevec™ (imatinib mesylate, Novartis), and the listed ligands stimulate one or more of the listed tyrosine kinases. In the two indications for which Gleevec™ is approved, tyrosine kinase targets (bcr-abl and c-kit) are overexpressed and also contain activating mutations. A finding that one of the Gleevec™ target tyrosine kinase targets is expressed in breast cancer tissue will prompt a second stage of analysis wherein the gene will be sequenced to determine whether it is mutated. That a mutation found is an activating mutation can be proved by methods known in the art, such as, for example, by measuring kinase enzyme activity or by measuring phosphorylation status of the particular kinase, relative to the corresponding wild-type kinase. Breast cancer patients whose tumors express high levels of mRNAs encoding Gleevec™ target tyrosine kinases, specifically, in the upper ten percentile, or mRNAs for Gleevec™ target tyrosine kinases in the average range and mRNAs for their cognate growth stimulating ligands in the upper ten percentile, are particularly good candidates for treatment with Gleevec™.

[0378] VEGF is a potent and pathologically important angiogenic factor. (See below under Prognostic Indicators.) When VEGF mRNA levels are in the upper ten percentile, aggressive treatment is warranted. Several levels particularly suggest the value of treatment with anti-angiogenic drugs, including VEGF antagonists, such as anti-VEGF antibodies. Additionally, KDR or CD31 mRNA level in the upper 20 percentile further increases likelihood of benefit from VEGF antagonists.

[0379] Farnesyl pyrophosphatase synthetase and geranyl pyrophosphatase synthetase. These enzymes are targets of commercialized bisphosphonate drugs, which were developed originally for treatment of osteoporosis but recently have begun to prescribe them off-label in breast cancer. Elevated levels of mRNAs encoding these enzymes in breast cancer tissue, above the 90th percentile, suggest use of bisphosphonates as a treatment option.

[0380] 2. Multidrug Resistance Factors

[0381] These factors include 10 Genes: gamma glutamyl cysteine synthetase [GCS]; GST-α; GST-π; MDR-1; MRPI-4; breast cancer resistance protein [BCRP]; lung resistance protein [MVP]; SXR; YB-1.

[0382] GCS and both GST-α and GST-π regulate glutathione levels, which decrease cellular sensitivity to chemotherapy drugs and other toxins by reductive derivatization. Glutathione is a necessary cofactor for multi-drug resistant pumps, MDR-1 and the MRPs. MDR1 and MRPs function to actively transport out of cells several important chemotherapeutic drugs used in breast cancer.

[0383] GSTs, MDR-1, and MRP-1, have all been studied extensively to determine possible have prognostic or predictive significance in human cancer. However, a great deal of disagreement exists in the literature with respect to these questions. Recently, new members of the MRP family have been identified: MRP-2, MRP-3, MRP-4, BCRP, and lung resistance protein [major vault protein]. These have substrate specificities that overlap with those of MDR-1 and MRP-1. The incorporation of all of these relevant ABC family members as well as glutathione synthetic enzymes into the present invention captures the contribution of this family to drug resistance, in a way that single or double analytic assays cannot.

[0384] MRP-1, the gene coding for the multidrug resistance protein.

[0385] P-glycoprotein, is not regulated primarily at the transcriptional level. However, p-glycoprotein stimulates the transcription of PTP1b. An embodiment of the present invention is the use of the level of the mRNA for the phosphatase PTP1b as a surrogate measure of MRP-1/p-glycoprotein activity.

[0386] The gene SXR is also an activator of multidrug resistance, as it stimulates transcription of certain multidrug resistance factors.
The impact of multidrug resistance factors with respect to chemotherapeutic agents used in breast cancer is as follows. Beneficial response to doxorubicin is compromised when the mRNA levels of either MDR1, GST\(\alpha\), GST\(\beta\), SXR, BCRP YB-1, or LRP/MVP are in the upper four percentiles. Beneficial response to methotrexate is inhibited if mRNA levels of any of MRP1, MRP2, MRP3, or MRP4 or gamma-glutamyl cysteine synthetase are in the upper four percentiles.

3. Eukaryotic Translation Initiation Factor 4E [eIF4E]

eIF4E mRNA levels provide evidence of protein expression and so expands the capability of RT-PCR to indicate variation in gene expression. Thus, one claim of the present invention is the use of eIF4E as an added indicator of gene expression of certain genes [e.g., cyclinD1, mdm2, VEGF, and others]. For example, in two tissue specimens containing the same amount of normalized VEGF mRNA, it is likely that the tissue containing the higher normalized level of eIF4E exhibits the greater level of VEGF gene expression.

The background is as follows. A key point in the regulation of mRNA translation is selection of mRNAs by the eIF4E complex to bind to the 43S ribosomal subunit. The protein eIF4E [the m7G CAP-binding protein] is often limiting because more mRNAs than eIF4E copies exist in cells. Highly structured 5'UTRs or highly GC-rich ones are inefficiently translated, and these often code for genes that carry out functions relevant to cancer [e.g., cyclinD1, mdm2, and VEGF]. eIF4E is itself regulated at the transcriptional/mRNA level. Thus, expression of eIF4E provides added indication of increased activity of a number of proteins.

It is also noteworthy that overexpression of eIF4E transforms cultured cells, and hence is an oncogene. Overexpression of eIF4E occurs in several different types of carcinomas but is particularly significant in breast cancer. eIF4E is typically expressed at very low levels in normal breast tissue.

D. Prognostic Indicators

1. DNA Repair Enzymes

Loss of BRCA1 or BRCA2 activity via mutation represents the critical oncogenic step in the most common type[s] of familial breast cancer. The levels of mRNAs of these important enzymes are abnormal in subsets of sporadic breast cancer as well. Loss of signals from either [to within the lower ten percent] heightens risk of short survival.

2. Cell Cycle Regulators

Cell cycle regulators include 14 genes: c-MYC; c-Src; Cyclin D1; Ha-Ras; mdm2; p14ARF; p21WAFL1/CIP; p16INK4a/p14; p23; p27; p53; P13K; PKC-epsilon; PKC-delta.

The gene for p53 [TP53] is mutated in a large fraction of breast cancers. Frequently p53 levels are elevated when loss of function mutation occurs. When the mutation is dominant-negative, it creates survival value for the cancer cell because growth is promoted and apoptosis is inhibited. Thousands of different p53 mutations have been found in human cancer, and the functional consequences of many of them are not clear. A large body of academic literature addresses the prognostic and predictive significance of mutated p53 and the results are highly conflicting. The present invention provides a functional genomic measure of p53 activity, as follows. The activated wild type p53 molecule triggers translocation of the cell cycle inhibitor p21. Thus, the ratio of p53 to p21 should be low when p53 is wild-type and activated. When p53 is detectable and the ratio of p53 to p21 is elevated in tumors relative to normal breast, it signifies nonfunctional or dominant negative p53. The cancer literature provides evidence for this as born out by poor prognosis.

Mdm2 is an important p53 regulator. Activated wildtype p53 stimulates transcription of mdm2. The mdm2 protein binds p53 and promotes its proteolytic destruction. Thus, abnormally low levels of mdm2 in the presence of normal or higher levels of p53 indicate that p53 is mutated and inactivated.

One aspect of the present invention is the use of ratios of mRNAs levels p53:p21 and p53:mdm2 to provide a picture of p53 status. Evidence for dominant negative mutation of p53 (as indicated by high p53:p21 and/or high p53:mdm2 mRNA ratios Specifically in the upper ten percentile) presages higher risk of recurrence in breast cancer and therefore weighs toward a decision to use chemotherapy in node negative post surgery breast cancer.

Another important cell cycle regulator is p27, which in the activated form blocks cell cycle progression at the level of cdk4. The protein is regulated primarily via phosphorylation/dephosphorylation, rather than at the transcriptional level. However, levels of p27 mRNAs do vary. Therefore a level of p27 mRNA in the upper ten percentile indicates reduced risk of recurrence of breast cancer post surgery.

Cyclin D1 is a principle positive regulator of entry into S phase of the cell cycle. The gene for cyclin D1 is amplified in about 20% of breast cancer patients, and therefore promotes tumor promotes tumor growth in those cases. One aspect of the present invention is use of cyclin D1 mRNA levels for diagnostic purposes in breast cancer. A level of cyclin D1 mRNA in the upper ten percentile suggests high risk of recurrence in breast cancer following surgery and suggests particular benefit of adjuvant chemotherapy.

3. Other Tumor Suppressors and Related Proteins

These include APC and E-cadherin. It has long been known that the tumor suppressor APC is lost in about 50% of colon cancers, with concomitant transcriptional upregulation of E-cadherin, an important cell adhesion molecule and growth suppressor. Recently, it has been found that the APC gene silenced in 15-40% of breast cancers. Likewise, the E-cadherin gene is silenced [via CpG island methylation] in about 30% of breast cancers. An abnormally low level of APC and/or E-cadherin mRNA in the lower 5 percentile suggests high risk of recurrence in breast cancer following surgery and heightened risk of shortened survival.

4. Regulators of Apoptosis

These include BCL2 family members BCL2, Bcl-xl, Bak, Box and related factors, NFk-B and related factors, and also p53BP1/ASPP1 and p53BP2/ASPP2.
Bax and Bak are pro-apoptotic and Bcl2 and Bcl-x1 are anti-apoptotic. Therefore, the ratios of these factors influence the resistance or sensitivity of a cell to toxic (pro-apoptotic) drugs. In breast cancer, unlike other cancers, elevated level of Bcl2 (in the upper ten percentile) correlates with good outcome. This reflects the fact that Bcl2 has growth inhibitory activity as well as anti-apoptotic activity, and in breast cancer the significance of the former activity outweighs the significance of the latter. The impact of Bcl2 is in turn dependent on the status of the growth stimulating transcription factor c-MYC. The gene for c-MYC is amplified in about 20% of breast cancers. When c-MYC message levels are abnormally elevated relative to Bcl2 (such that this ratio is in the upper ten percentile), then elevated level of Bcl2 mRNA is no longer a positive indicator.

NFκ-B is another important anti-apoptotic factor. Originally, recognized as a pro-inflammatory transcription factor, it is now clear that it prevents programmed cell death in response to several extracellular toxic factors [such as tumor necrosis factor]. The activity of this transcription factor is regulated principally via phosphorylation/dephosphorylation events. However, levels of NFκ-B nevertheless do vary from cell to cell, and elevated levels should correlate with increased resistance to apoptosis. Importantly for present purposes, NFκ-B exerts its anti-apoptotic activity largely through its stimulation of transcription of mRNAs encoding certain members of the IAP [inhibitor of apoptosis] family of proteins, specifically cIAP1, cIAP2, XIAP, and Survivin. Thus, abnormally elevated levels of mRNAs for these IAPs and for NFκ-B any in the upper 5 percentile signifies activation of the NFκ-B anti-apoptotic pathway. This suggests high risk of recurrence in breast cancer following chemotherapy and therefore poor prognosis. One embodiment of the present invention is the inclusion in the gene set of the above apoptotic regulators, and the above-outlined use of combinations and ratios of the levels of their mRNAs for prognosis in breast cancer.

The proteins p53BP1 and 2 bind to p53 and promote transcriptional activation of pro-apoptotic genes. The levels of p53BP1 and 2 are suppressed in a significant fraction of breast cancers, correlating with poor prognosis. When either is expressed in the lower tenth percentile poor prognosis is indicated.

5. Factors that Control Cell Invasion and Angiogenesis

These include uPA, PAI1, cathepsinsB, G and L, scatter factor [HGF], c-met, KDR, VEGF, and CD31. The plasminogen activator uPA and its serpin regulator PAI1 promote breakdown of extracellular matrices and tumor cell invasion. Abnormally elevated levels of both mRNAs in malignant breast tumors (in the upper twenty percentile) signify an increased risk of shortened survival, increased recurrence in breast cancer patients post surgery, and increased importance of receiving adjuvant chemotherapy. On the other hand, node negative patients whose tumors do not express elevated levels of these mRNA species are less likely to have recurrence of this cancer and could more seriously consider whether the benefits of standard chemotherapy justifies the associated toxicity.

Cathepsins B or L, when expressed in the upper ten percentile, predict poor disease-free and overall survival. In particular, cathepsin L predicts short survival in node positive patients.

Scatter factor and its cognate receptor c-met promote cell motility and invasion, cell growth, and angiogenesis. In breast cancer elevated levels of mRNAs encoding these factors should prompt aggressive treatment with chemotherapeutic drugs, when expression of either, or the combination, is above the 90th percentile.

VEGF is a central positive regulator of angiogenesis, and elevated levels in solid tumors predict short survival [note many references showing that elevated level of VEGF predicts short survival]. Inhibitors of VEGF therefore slow the growth of solid tumors in animals and humans. VEGF activity is controlled at the level of transcription. VEGF mRNA levels in the upper ten percentile indicate significantly worse than average prognosis. Other markers of vascularization, CD31 [PECAM], and KDR indicate high vessel density in tumors and that the tumor will be particularly malignant and aggressive, and hence that an aggressive therapeutic strategy is warranted.

6. Markers for Immune and Inflammatory Cells and Processes

These markers include the genes for Immunglobulin light chain λ, CD18, CD3, CD68, Fas [CD95], and Fas Ligand.

Several lines of evidence suggest that the mechanisms of action of certain drugs used in breast cancer entail activation of the host immune/inflammatory response (For example, Herceptin®). One aspect of the present invention is the inclusion in the gene set of markers for inflammatory and immune cells, and markers that predict tumor resistance to immune surveillance. Immunoglobulin light chain lambda is a marker for immunoglobulin producing cells. CD18 is a marker for all white cells. CD3 is a marker for T-cells. CD68 is a marker for macrophages.

CD95 and Fas ligand are a receptor: ligand pair that mediate one of two major pathways by which cytotoxic T cells and NK cells kill targeted cells. Decreased expression of CD95 and increased expression of Fas Ligand indicates poor prognosis in breast cancer. Both CD95 and Fas Ligand are transmembrane proteins, and need to be membrane anchored to trigger cell death. Certain tumor cells produce a truncated soluble variant of CD95, created as a result of alternative splicing of the CD95 mRNA. This blocks NK cell and cytotoxic T cell Fas Ligand-mediated killing of the tumors cells. Presence of soluble CD95 correlates with poor survival in breast cancer. The gene set includes both soluble and full-length variants of CD95.

7. Cell Proliferation Markers

The gene set includes the cell proliferation markers Ki67/Mib1, PCNA, Pin1, and thymidine kinase. High levels of expression of proliferation markers associate with high histologic grade, and short survival. High levels of thymidine kinase in the upper ten percentile suggest in decreased risk of short survival. Pin1 is a prolyl isomerase that stimulates cell growth, in part through the transcriptional activation of the cyclin D1 gene, and levels in the upper ten percentile contribute to a negative prognostic profile.

8. Other Growth Factors and Receptors

This gene set includes IGF1, IGF2, IGFBP3, IGFR1, FGF2, FGFR1, CSF-1R/frms, CSF-1, II.6 and II.8. All of these proteins are expressed in breast cancer. Most
stimulate tumor growth. However, expression of the growth factor FGF2 correlates with good outcome. Some have anti-apoptotic activity, prominently IGF1. Activation of the IGF1 axis via elevated IGF1, IGF1R, or IGFBP3 (as indicated by the sum of these signals in the upper ten percentile) inhibits tumor cell death and strongly contributes to a poor prognostic profile.

9. Gene Expression Markers that Define Subclasses of Breast Cancer

These include: GRO1 oncogene alpha, Grb7, cytokeratin 5 and 17, retinoblastoma protein 4, hepatocyte nuclear factor 3, integrin alpha 7, and lipoprotein lipase. These markers subset breast cancer into different cell types that are phenotypically different at the level of gene expression. Tumors expressing signals for Bcl2, hepatocyte nuclear factor 3, IGF1 and ER above the mean have the best prognosis for disease free and overall survival following surgical removal of the cancer. Another category of breast cancer tumor type, characterized by elevated expression of lipoprotein lipase, retinoblastoma protein 4, and integrin alpha 7, carry intermediate prognosis. Tumors expressing either elevated levels of cytokeratin 5, and 17, GRO oncogene at levels four-fold or greater above the mean, or ErbB2 and Grb7 at levels ten-fold or more above the mean, have worst prognosis.

Although throughout the present description, including the Examples below, various aspects of the invention are explained with reference to gene expression studies, the invention can be performed in a similar manner, and similar results can be reached by applying proteomics techniques that are well known in the art. The proteome is the totality of the proteins present in a sample (e.g. tissue, organism, or cell culture) at a certain point of time. Proteomics includes, among other things, study of the global changes of protein expression in a sample (also referred to as "expression proteomics"). Proteomics typically includes the following steps: (1) separation of individual proteins in a sample by 2-D gel electrophoresis (2-D PAGE); (2) identification of the individual proteins recovered from the gel, e.g. my mass spectrometry and/or N-terminal sequencing, and (3) analysis of the data using bioinformatics. Proteomics methods are valuable supplements to other methods of gene expression profiling, and can be used, alone or in combination with other methods of the present invention, to detect the products of the gene markers of the present invention.

Further details of the invention will be described in the following non-limiting Examples.

EXAMPLE 1

Isolation of RNA from Formalin-Fixed, Paraffin-Embedded (FFPE) Tissue Specimens

A. Protocols

1. EPICENTRE® Xylene Protocol

RNA Isolation

(1) Cut 1-6 sections (each 10 µm thick) of paraffin-embedded tissue per sample using a clean microtome blade and place into a 1.5 ml eppendorf tube.

(2) To extract paraffin, add 1 ml of xylene and invert the tubes for 10 minutes by rocking on a nutator.

(3) Pellet the sections by centrifugation for 10 minutes at 14,000 x g in an eppendorf microcentrifuge.

(4) Remove the xylene, leaving some in the bottom to avoid dislodging the pellet.

(5) Repeat steps 2-4.

(6) Add 1 ml of 100% ethanol and invert for 3 minutes by rocking on the nutator.

(7) Pellet the debris by centrifugation for 10 minutes at 14,000 x g in an eppendorf microcentrifuge.

(8) Remove the ethanol, leaving some at the bottom to avoid the pellet.

(9) Repeat steps 6-8 twice.

(10) Remove all of the remaining ethanol.

(11) For each sample, add 2 µl of 50 µg/µl Proteinase K to 300 µl of Tissue and Cell Lysis Solution.

(12) Add 300 µl of Tissue and Cell Lysis Solution containing the Proteinase K to each sample and mix thoroughly.

(13) Incubate at 65°C for 90 minutes (vortex mixing every 5 minutes). Visually monitor the remaining tissue fragment. If still visible after 30 minutes, add an additional 2 µl of 50 µg/µl Proteinase K and continue incubating at 65°C until fragment dissolves.

(14) Place the samples on ice for 3-5 minutes and proceed with protein removal and total nucleic acid precipitation.

(15) Protein Removal and Precipitation of Total Nucleic Acid

(16) Add 150 µl of MPC Protein Precipitation Reagent to each lysed sample and vortex vigorously for 10 seconds.

(17) Pellet the debris by centrifugation for 10 minutes at 14,000 x g in an eppendorf microcentrifuge.

(18) Transfer the supernatant into clean eppendorf tubes and discard the pellet.

(19) Add 500 µl of isopropanol to the recovered supernatant and thoroughly mix by rocking on the nutator for 3 minutes.

(20) Pellet the RNA/DNA by centrifugation at 4°C for 10 minutes at 14,000 x g in an eppendorf microcentrifuge.

(21) Remove all of the isopropanol with a pipet, being careful not to dislodge the pellet.

Removal of Contaminating DNA from RNA Preparations

(22) Prepare 200 µl of DNase I solution for each sample by adding 5 µl of RNase-Free DNase I (1 U/µl) to 195 µl of 1× DNase Buffer.

(23) Completely resuspend the pelleted RNA in 200 µl of DNase I solution by vortexing.

(24) Incubate the samples at 37°C for 60 minutes.

(25) Add 200 µl of 2x T and C Lysis Solution to each sample and vortex for 5 seconds.
(0455) (5) Add 200 µl of MPC Protein Precipitation Reagent, mix by vortexing for 10 seconds and place on ice for 3-5 minutes.

(0456) (6) Pellet the debris by centrifugation for 10 minutes at 14,000xg in an eppendorf microcentrifuge.

(0457) (7) Transfer the supernatant containing the RNA to clean eppendorf tubes and discard the pellet. (Be careful to avoid transferring the pellet.)

(0458) (8) Add 500 µl of isopropanol to each supernatant and rock samples on the nutator for 3 minutes.

(0459) (9) Pellet the RNA by centrifugation at 4°C for 10 minutes at 14,000xg in an eppendorf microcentrifuge.

(0460) (10) Remove the isopropanol, leaving some at the bottom to avoid dislodging the pellet.

(0461) (11) Rinse twice with 1 ml of 75% ethanol. Centrifuge briefly if the RNA pellet is dislodged.

(0462) (12) Remove ethanol carefully.

(0463) (13) Set under fume hood for about 3 minutes to remove residual ethanol.

(0464) (14) Resuspend the RNA in 30 µl of TE Buffer and store at ~30°C.

II. Hot Wax/Urea Protocol of the Invention

(0465) RNA Isolation

(0467) (1) Cut 3 sections (each 10 µm thick) of paraffin-embedded tissue using a clean microtome blade and place into a 1.5 ml eppendorf tube.

(0468) (2) Add 300 µl of lysis buffer (10 mM Tris 7.5, 0.5% sodium lauryl sarcosine, 0.1 mM EDTA pH 7.5, 4M Urea) containing 330 µg/ml Proteinase K (added freshly from a 50 µg/ml stock solution) and vortex briefly.

(0469) (3) Incubate at 65°C for 90 minutes (vortex mixing every 5 minutes). Visually monitor the tissue fragment. If still visible after 30 minutes, add an additional 2 µl of 50 µg/ml Proteinase K and continue incubating at 65°C until fragment dissolves.

(0470) (4) Centrifuge for 5 minutes at 14,000xg and transfer upper aqueous phase to new tube, being careful not to disrupt the paraffin seal.

(0471) (5) Place the samples on ice for 3-5 minutes and proceed with protein removal and total nucleic acid precipitation.

(0472) Protein Removal and Precipitation of Total Nucleic Acid

(0473) (1) Add 150 µl of 7.5M NH4OAc to each lysed sample and vortex vigorously for 10 seconds.

(0474) (2) Pellet the debris by centrifugation for 10 minutes at 14,000xg in an eppendorf microcentrifuge.

(0475) (3) Transfer the supernatant into clean eppendorf tubes and discard the pellet.

(0476) (4) Add 500 µl of isopropanol to the recovered supernatant and thoroughly mix by rocking on the nutator for 3 minutes.

(0477) (5) Pellet the RNA/DNA by centrifugation at 4°C for 10 minutes at 14,000xg in an eppendorf microcentrifuge.

(0478) (6) Remove all of the isopropanol with a pipet, being careful not to dislodge the pellet.

(0479) Removal of Contaminating DNA from RNA Preparations

(0480) (1) Add 45 µl of 1× DNase I buffer (10 mM Tris-Cl, pH 7, 5, 2.5 mM MgCl2, 0.1 mM CaCl2) and 5 µl of RNase-Free DNase I (2 U/µl, Ambion) to each sample.

(0481) (2) Incubate the samples at 37°C for 60 minutes. Inactivate the DNasel by heating at 70°C for 5 minutes.

(0482) B. Results

(0483) Experimental evidence demonstrates that the hot RNA extraction protocol of the invention does not compromise RNA yield. Using 19 FPE breast cancer specimens, extracting RNA from three adjacent sections in the same specimen, RNA yields were measured via capillary electrophoresis with florescence detection (Agilent Bioanalyzer). Average RNA yields in nanograms and standard deviations with the invented and commercial methods, respectively, were: 139+/−21 versus 141+/−34.

(0484) Also, it was found that the urea-containing lysis buffer of the present invention can be substituted for the EPICENTRE® T&C lysis buffer, and the 7.5 M NH4OAc reagent used for protein precipitation in accordance with the present invention can be substituted for the EPICENTRE® MPC protein precipitation solution with neither significant compromise of RNA yield nor TaqMan® efficiency.

EXAMPLE 2

(0485) Amplification of mRNA Species Prior to RT-PCR

(0486) The method described in section 10 above was used with RNA isolated from fixed, paraffin-embedded breast cancer tissue. TaqMan® analyses were performed with first strand cDNA generated with the T7-GSP primer (unamplified(T7-GSP)), T7 amplified RNA (amplified(T7-GSP)). RNA was amplified according to step 2 of FIG. 4. As a control, TaqMan® was also performed with cDNA generated with an unmodified GSP (amplified(GSP)). An equivalent amount of initial template (1 ng/well) was used in each TaqMan® reaction.

(0487) The results are shown in FIG. 8. In vitro transcription increased RT-PCR signal intensity by more than 10 fold, and for certain genes by more than 100 fold relative to controls in which the RT-PCR primers were the same primers used in method 2 for the generation of double-stranded DNA for in vitro transcription (GSP-T7, and GSP). Also shown in FIG. 8 are RT-PCR data generated when standard optimized RT-PCR primers (i.e., lacking T7 tails) were used. As shown, compared to this control, the new method yielded substantial increases in RT-PCR signal (from 4 to 64 fold in this experiment).

(0488) The new method requires that each T7-GSP sequence be optimized so that the increase in the RT-PCR signal is the same for each gene, relative to the standard optimized RT-PCR (with non-T7 tailed primers).
EXAMPLE 3

A Study of Gene Expression in Premalignant and Malignant Breast Tumors

A gene expression study was designed and conducted with the primary goal to molecularly characterize gene expression in paraffin-embedded, fixed tissue samples of invasive breast ductal carcinoma, and to explore the correlation between such molecular profiles and disease-free survival. A further objective of the study was to compare the molecular profiles in tissue samples of invasive breast cancer with the molecular profiles obtained in ductal carcinoma in situ. The study was further designed to obtain data on the molecular profiles in lobular carcinoma in situ and in paraffin-embedded, fixed tissue samples of invasive lobular carcinoma.

Molecular assays were performed on paraffin-embedded, formalin-fixed primary breast tumor tissues obtained from 202 individual patients diagnosed with breast cancer. All patients underwent surgery with diagnosis of invasive ductal carcinoma of the breast, pure ductal carcinoma in situ (DCIS), lobular carcinoma of the breast, or pure lobular carcinoma in situ (LCIS). Patients were included in the study only if histopathologic assessment, performed as described in the Materials and Methods section, indicated adequate amounts of tumor tissue and homogeneous pathology.

The individuals participating in the study were divided into the following groups:

- Group 1: Pure ductal carcinoma in situ (DCIS); n=18
- Group 2: Invasive ductal carcinoma n=130
- Group 3: Pure lobular carcinoma in situ (LCIS); n=7
- Group 4: Invasive lobular carcinoma n=16

Materials and Methods

Each representative tumor block was characterized by standard histopathology for diagnosis, semi-quantitative assessment of amount of tumor, and tumor grade. A total of 6 sections (10 microns in thickness each) were prepared and placed in two Costar Brand Microcentrifuge Tubes (Polypropylene, 1.7 mL tubes, clear; 3 sections in each tube). If the tumor constituted less than 30% of the total specimen area, the sample may have been crudely dissected by the pathologist, using gross microdissection, putting the tumor tissue directly into the Costar tube.

If more than one tumor block was obtained as part of the surgical procedure, all tumor blocks were subjected to the same characterization, as described above, and the block most representative of the pathology was used for analysis.

Gene Expression Analysis

mRNA was extracted and purified from fixed, paraffin-embedded tissue samples, and prepared for gene expression analysis as described in chapters 7-11 above. Molecular assays of quantitative gene expression were performed by RT-PCR, using the ABI PRISM 7900™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, Calif., USA). ABI PRISM 7900™ consists of a thermocycler, laser, charge-coupled device (CCD), camera and computer. The system amplifies samples in a 384-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 384 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

Analysis and Results

Tumor tissue was analyzed for 185 cancer-related genes and 7 reference genes. The threshold cycle (CT) values for each patient were normalized based on the median of all genes for that particular patient. Clinical outcome data were available for all patients from a review of registry data and selected patient charts. Outcomes were classified as:

- 0 died due to breast cancer or to unknown cause or alive with breast cancer recurrence;
- 1 alive without breast cancer recurrence or died due to a cause other than breast cancer

Analysis was performed by:
1. Analysis of the relationship between normalized gene expression and the binary outcomes of 0 or 1.
2. Analysis of the relationship between normalized gene expression and the time to outcome (0 or 1 as defined above) where patients who were alive without breast cancer recurrence or who died due to a cause other than breast cancer were censored. This approach was used to evaluate the prognostic impact of individual genes and also sets of multiple genes.

Analysis of 147 Patients with Invasive Breast Carcinoma by Binary Approach

In the first (binary) approach, analysis was performed on all 146 patients with invasive breast carcinoma. A test was performed on the group of patients classified as 0 or 1 and the p-values for the differences between the groups for each gene were calculated.

The following Table 4 lists the 45 genes for which the p-value for the differences between the groups was <0.05.

<table>
<thead>
<tr>
<th>Gene/SEQ_ID No</th>
<th>Mean CT Alive</th>
<th>Mean CT Deceased</th>
<th>t-value</th>
<th>Degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXM1</td>
<td>33.66</td>
<td>32.52</td>
<td>3.92</td>
<td>144</td>
<td>0.0001</td>
</tr>
<tr>
<td>FRAME</td>
<td>35.45</td>
<td>33.84</td>
<td>3.71</td>
<td>144</td>
<td>0.0003</td>
</tr>
<tr>
<td>BC12</td>
<td>29.52</td>
<td>29.32</td>
<td>-3.53</td>
<td>144</td>
<td>0.0006</td>
</tr>
<tr>
<td>STK15</td>
<td>30.82</td>
<td>30.10</td>
<td>3.49</td>
<td>144</td>
<td>0.0006</td>
</tr>
<tr>
<td>CREB1</td>
<td>29.12</td>
<td>30.86</td>
<td>-3.39</td>
<td>144</td>
<td>0.0009</td>
</tr>
<tr>
<td>KI-67</td>
<td>30.57</td>
<td>29.62</td>
<td>3.34</td>
<td>144</td>
<td>0.0011</td>
</tr>
<tr>
<td>GSTM1</td>
<td>30.62</td>
<td>31.63</td>
<td>-3.27</td>
<td>144</td>
<td>0.0014</td>
</tr>
<tr>
<td>CA9</td>
<td>34.96</td>
<td>33.54</td>
<td>3.18</td>
<td>144</td>
<td>0.0018</td>
</tr>
<tr>
<td>FR</td>
<td>29.56</td>
<td>31.22</td>
<td>-3.16</td>
<td>144</td>
<td>0.0019</td>
</tr>
<tr>
<td>EBC3</td>
<td>31.54</td>
<td>32.10</td>
<td>-3.10</td>
<td>144</td>
<td>0.0023</td>
</tr>
</tbody>
</table>
TABLE 4-continued

<table>
<thead>
<tr>
<th>Gene/SEQ ID NO:</th>
<th>Mean CT Alive</th>
<th>Mean CT Deceased</th>
<th>t-value</th>
<th>Degrees of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NME1</td>
<td>27.31</td>
<td>26.68</td>
<td>3.04</td>
<td>144</td>
<td>0.0028</td>
</tr>
<tr>
<td>SURV</td>
<td>31.64</td>
<td>30.68</td>
<td>2.92</td>
<td>144</td>
<td>0.0041</td>
</tr>
<tr>
<td>GATA3</td>
<td>26.06</td>
<td>26.99</td>
<td>-2.91</td>
<td>144</td>
<td>0.0042</td>
</tr>
<tr>
<td>TFRC</td>
<td>28.96</td>
<td>28.48</td>
<td>2.87</td>
<td>144</td>
<td>0.0047</td>
</tr>
<tr>
<td>YB-1</td>
<td>26.72</td>
<td>26.41</td>
<td>2.79</td>
<td>144</td>
<td>0.0060</td>
</tr>
<tr>
<td>DPYD</td>
<td>28.51</td>
<td>28.84</td>
<td>-2.67</td>
<td>144</td>
<td>0.0084</td>
</tr>
<tr>
<td>GSTM3</td>
<td>28.21</td>
<td>29.03</td>
<td>-2.63</td>
<td>144</td>
<td>0.0095</td>
</tr>
<tr>
<td>RPS6KB1</td>
<td>31.18</td>
<td>30.61</td>
<td>2.61</td>
<td>144</td>
<td>0.0099</td>
</tr>
<tr>
<td>Src</td>
<td>27.97</td>
<td>27.69</td>
<td>2.59</td>
<td>144</td>
<td>0.0105</td>
</tr>
<tr>
<td>Chk1</td>
<td>32.63</td>
<td>31.99</td>
<td>2.57</td>
<td>144</td>
<td>0.0113</td>
</tr>
<tr>
<td>ID1</td>
<td>20.73</td>
<td>29.13</td>
<td>-2.48</td>
<td>144</td>
<td>0.0141</td>
</tr>
<tr>
<td>EstR1</td>
<td>24.22</td>
<td>25.40</td>
<td>-2.44</td>
<td>144</td>
<td>0.0160</td>
</tr>
<tr>
<td>p27</td>
<td>27.15</td>
<td>27.51</td>
<td>-2.41</td>
<td>144</td>
<td>0.0174</td>
</tr>
<tr>
<td>CCNB1</td>
<td>31.63</td>
<td>30.87</td>
<td>2.40</td>
<td>144</td>
<td>0.0176</td>
</tr>
<tr>
<td>XIAP</td>
<td>30.27</td>
<td>30.51</td>
<td>-2.40</td>
<td>144</td>
<td>0.0178</td>
</tr>
<tr>
<td>Chk2</td>
<td>31.48</td>
<td>31.11</td>
<td>2.39</td>
<td>144</td>
<td>0.0179</td>
</tr>
<tr>
<td>CDC25B</td>
<td>29.75</td>
<td>29.39</td>
<td>2.37</td>
<td>144</td>
<td>0.0193</td>
</tr>
<tr>
<td>IGFlR</td>
<td>28.85</td>
<td>29.44</td>
<td>-2.34</td>
<td>144</td>
<td>0.0209</td>
</tr>
<tr>
<td>AK055699</td>
<td>33.23</td>
<td>34.11</td>
<td>-2.28</td>
<td>144</td>
<td>0.0242</td>
</tr>
<tr>
<td>PI3KC2A</td>
<td>31.07</td>
<td>31.42</td>
<td>-2.25</td>
<td>144</td>
<td>0.0257</td>
</tr>
<tr>
<td>TGFBlB</td>
<td>28.42</td>
<td>28.85</td>
<td>-2.25</td>
<td>144</td>
<td>0.0258</td>
</tr>
<tr>
<td>RAG1I</td>
<td>28.40</td>
<td>28.75</td>
<td>-2.24</td>
<td>144</td>
<td>0.0269</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>35.70</td>
<td>35.32</td>
<td>2.17</td>
<td>144</td>
<td>0.0317</td>
</tr>
<tr>
<td>EpCAM</td>
<td>28.73</td>
<td>28.34</td>
<td>2.16</td>
<td>144</td>
<td>0.0321</td>
</tr>
<tr>
<td>VEGFC</td>
<td>32.28</td>
<td>31.82</td>
<td>2.16</td>
<td>144</td>
<td>0.0326</td>
</tr>
<tr>
<td>p53</td>
<td>20.96</td>
<td>30.60</td>
<td>-2.14</td>
<td>144</td>
<td>0.0341</td>
</tr>
<tr>
<td>hENT1</td>
<td>27.19</td>
<td>26.91</td>
<td>2.12</td>
<td>144</td>
<td>0.0357</td>
</tr>
<tr>
<td>WISP1</td>
<td>31.20</td>
<td>31.64</td>
<td>-2.10</td>
<td>144</td>
<td>0.0377</td>
</tr>
<tr>
<td>HNF3A</td>
<td>27.89</td>
<td>28.64</td>
<td>-2.09</td>
<td>144</td>
<td>0.0384</td>
</tr>
<tr>
<td>NFkBe65</td>
<td>33.22</td>
<td>33.00</td>
<td>-2.08</td>
<td>144</td>
<td>0.0396</td>
</tr>
<tr>
<td>BRCA2</td>
<td>33.06</td>
<td>32.62</td>
<td>2.08</td>
<td>144</td>
<td>0.0397</td>
</tr>
<tr>
<td>EGFR</td>
<td>30.68</td>
<td>30.13</td>
<td>2.06</td>
<td>144</td>
<td>0.0414</td>
</tr>
<tr>
<td>TK1</td>
<td>32.27</td>
<td>31.72</td>
<td>2.02</td>
<td>144</td>
<td>0.0453</td>
</tr>
<tr>
<td>VDR</td>
<td>30.08</td>
<td>29.73</td>
<td>1.99</td>
<td>144</td>
<td>0.0488</td>
</tr>
</tbody>
</table>

[0505] In the foregoing Table 4, lower (negative) t-values indicate higher expression (or lower CTs), associated with better outcomes, and, inversely, higher (positive) t-values indicate higher expression (lower CTs) associated with worse outcomes. Thus, for example, elevated expression of the FOXM1 gene (t-value=3.92, CT mean alive>CT mean deceased) indicates a reduced likelihood of disease free survival. Similarly, elevated expression of the CEGP1 gene (t-value=-3.39; CT mean alive<CT mean deceased) indicates an increased likelihood of disease free survival.

[0506] Based on the data set forth in Table 4, the overexpression of any of the following genes in breast cancer indicates a reduced likelihood of survival without cancer recurrence following surgery: FOXM1; PRAE; SKT15; Ki-67; CA9; NME1; SURV; TFRC; YB-1; RPS6 KB1; Src; Chk1; CCNB1; Chk2; CDC25B; CYP3A4; EpCAM; VEGFC; hENT1; BRCA2; EGFR; TK1; VDR.

[0507] Based on the data set forth in Table 4, the overexpression of any of the following genes in breast cancer indicates a better prognosis for survival without cancer recurrence following surgery: Bcl2; CEGP1; GSTM1; PR; BBC3; GATA3; DPYD; GSTM3; ID1; EstR1; p27; XIAP; IGF1R; AK055699; PI3KC2A; TGFBlB; BAG11; p52; WISP1; HNF3A; NFkB65.


[0509] 108 patients with normalized CT for estrogen receptor (ER)>25.2 (i.e., ER positive patients) were subjected to separate analysis. A t-test was performed on the groups of patients classified as 0 or 1 and the p-values for the differences between the groups for each gene were calculated. The following Table 5 lists the 12 genes where the p-value for the differences between the groups was <0.05.

TABLE 5

<table>
<thead>
<tr>
<th>Gene/SEQ ID NO:</th>
<th>Mean CT Alive</th>
<th>Mean CT Deceased</th>
<th>t-value</th>
<th>Degrees of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRAME</td>
<td>35.54</td>
<td>33.88</td>
<td>3.03</td>
<td>106</td>
<td>0.0031</td>
</tr>
<tr>
<td>Bcl2</td>
<td>28.24</td>
<td>28.87</td>
<td>-2.70</td>
<td>106</td>
<td>0.0092</td>
</tr>
<tr>
<td>FOXM1</td>
<td>33.82</td>
<td>32.85</td>
<td>2.66</td>
<td>106</td>
<td>0.0099</td>
</tr>
<tr>
<td>DIABLO</td>
<td>30.33</td>
<td>30.71</td>
<td>-2.47</td>
<td>106</td>
<td>0.0153</td>
</tr>
<tr>
<td>EPCH1</td>
<td>29.62</td>
<td>28.03</td>
<td>2.44</td>
<td>106</td>
<td>0.0163</td>
</tr>
<tr>
<td>HIF1A</td>
<td>29.37</td>
<td>28.88</td>
<td>2.40</td>
<td>106</td>
<td>0.0180</td>
</tr>
<tr>
<td>VEGFC</td>
<td>32.39</td>
<td>31.69</td>
<td>2.39</td>
<td>106</td>
<td>0.0197</td>
</tr>
<tr>
<td>KL-67</td>
<td>30.73</td>
<td>29.82</td>
<td>2.38</td>
<td>106</td>
<td>0.0191</td>
</tr>
<tr>
<td>IGF1R</td>
<td>29.60</td>
<td>29.18</td>
<td>-2.37</td>
<td>106</td>
<td>0.0194</td>
</tr>
<tr>
<td>VDR</td>
<td>30.14</td>
<td>29.60</td>
<td>2.17</td>
<td>106</td>
<td>0.0322</td>
</tr>
<tr>
<td>NME1</td>
<td>27.34</td>
<td>26.80</td>
<td>2.03</td>
<td>106</td>
<td>0.0452</td>
</tr>
<tr>
<td>GSTM3</td>
<td>28.08</td>
<td>28.92</td>
<td>-2.00</td>
<td>106</td>
<td>0.0485</td>
</tr>
</tbody>
</table>

[0510] For each gene, a classification algorithm was utilized to identify the best threshold value (CT) for using each gene alone in predicting clinical outcome.

[0511] Based on the data set forth in Table 5, overexpression of the following genes in ER-positive cancer is indica-
tive of a reduced likelihood of survival without cancer recurrence following surgery: PRAME; FOXM1; EPHX1; HIF1A; VEGFC; Ki-67; VDR; NME1. Some of these genes (PRAME; FOXM1; VEGFC; Ki-67; VDR; and NME1) were also identified as indicators of poor prognosis in the previous analysis, not limited to ER-positive breast cancer. The overexpression of the remaining genes (EPHX1 and HIF1A) appears to be negative indicator of disease free survival in ER-positive breast cancer only. Based on the data set forth in Table 5, overexpression of the following genes in ER-positive cancer is indicative of a better prognosis for survival without cancer recurrence following surgery: Bcl-2; DIABLO; IGF1R; GSTM3. Of the latter genes, Bcl-2; IGF1R; and GSTM3 have also been identified as indicators of good prognosis in the previous analysis, not limited to ER-positive breast cancer. The overexpression of DIABLO appears to be positive indicator of disease free survival in ER-positive breast cancer only.

[0512] Analysis of Multiple Genes and Indicators of Outcome

[0513] Two approaches were taken in order to determine whether using multiple genes would provide better discrimination between outcomes.

[0514] First, a discrimination analysis was performed using a forward stepwise approach. Models were generated that classified outcome with greater discrimination than was obtained with any single gene alone.

[0515] According to a second approach (time-to-event approach), for each gene a Cox Proportional Hazards model (see, e.g., Cox, D. R., and Oakes, D. (1984). Analysis of Survival Data. Chapman and Hall, London, New York) was defined with time to recurrence or death as the dependent variable, and the expression level of the gene as the independent variable. The genes that have a p-value<0.05 in the Cox model were identified. For each gene, the Cox model provides the relative risk (RR) of recurrence or death for a unit change in the expression of the gene. One can choose to partition the patients into subgroups at any threshold value of the measured expression (on the CT scale), where all patients with expression values above the threshold have higher risk, and all patients with expression values below the threshold have lower risk, or vice versa, depending on whether the gene is an indicator of good (RR>1.01) or poor (RR<1.01) prognosis. Thus, any threshold value will define subgroups of patients with respectively increased or decreased risk. The results are summarized in the following Tables 6 and 7.

### Table 6

Cox Model Results for 146 Patients with Invasive Breast Cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Relative Risk (RR)</th>
<th>SE Relative Risk</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEBP1</td>
<td>1.25</td>
<td>0.07</td>
<td>0.0014</td>
</tr>
<tr>
<td>GSTM1</td>
<td>1.40</td>
<td>0.11</td>
<td>0.0014</td>
</tr>
<tr>
<td>Ki67</td>
<td>0.62</td>
<td>0.15</td>
<td>0.0016</td>
</tr>
<tr>
<td>FR</td>
<td>1.23</td>
<td>0.07</td>
<td>0.0017</td>
</tr>
<tr>
<td>Contig51037</td>
<td>0.81</td>
<td>0.07</td>
<td>0.0022</td>
</tr>
<tr>
<td>NME1</td>
<td>0.64</td>
<td>0.15</td>
<td>0.0023</td>
</tr>
<tr>
<td>YB-1</td>
<td>0.39</td>
<td>0.32</td>
<td>0.0033</td>
</tr>
<tr>
<td>TP53</td>
<td>0.53</td>
<td>0.21</td>
<td>0.0035</td>
</tr>
<tr>
<td>BC3</td>
<td>1.72</td>
<td>0.19</td>
<td>0.0036</td>
</tr>
<tr>
<td>GATA3</td>
<td>1.32</td>
<td>0.10</td>
<td>0.0039</td>
</tr>
<tr>
<td>CA9</td>
<td>0.81</td>
<td>0.07</td>
<td>0.0049</td>
</tr>
<tr>
<td>SUBV</td>
<td>0.69</td>
<td>0.13</td>
<td>0.0049</td>
</tr>
<tr>
<td>DPDY</td>
<td>2.58</td>
<td>0.34</td>
<td>0.0052</td>
</tr>
<tr>
<td>RPS6KB1</td>
<td>0.60</td>
<td>0.18</td>
<td>0.0055</td>
</tr>
<tr>
<td>GSTM3</td>
<td>1.36</td>
<td>0.12</td>
<td>0.0078</td>
</tr>
<tr>
<td>Src2</td>
<td>0.39</td>
<td>0.36</td>
<td>0.0094</td>
</tr>
<tr>
<td>TGFB3</td>
<td>1.61</td>
<td>0.19</td>
<td>0.0109</td>
</tr>
<tr>
<td>CDC25B</td>
<td>0.54</td>
<td>0.25</td>
<td>0.0122</td>
</tr>
<tr>
<td>XIAP</td>
<td>3.20</td>
<td>0.47</td>
<td>0.0126</td>
</tr>
<tr>
<td>CCNB1</td>
<td>0.60</td>
<td>0.16</td>
<td>0.0151</td>
</tr>
<tr>
<td>IGF1R</td>
<td>1.42</td>
<td>0.15</td>
<td>0.0153</td>
</tr>
<tr>
<td>Chk1</td>
<td>0.68</td>
<td>0.16</td>
<td>0.0155</td>
</tr>
<tr>
<td>ID1</td>
<td>1.80</td>
<td>0.25</td>
<td>0.0164</td>
</tr>
<tr>
<td>p27</td>
<td>1.69</td>
<td>0.22</td>
<td>0.0168</td>
</tr>
<tr>
<td>Chk2</td>
<td>0.52</td>
<td>0.27</td>
<td>0.0175</td>
</tr>
<tr>
<td>EstR1</td>
<td>1.17</td>
<td>0.07</td>
<td>0.0196</td>
</tr>
<tr>
<td>HNF3A</td>
<td>1.21</td>
<td>0.08</td>
<td>0.0206</td>
</tr>
<tr>
<td>p62</td>
<td>1.12</td>
<td>0.05</td>
<td>0.0230</td>
</tr>
<tr>
<td>EAG1</td>
<td>1.08</td>
<td>0.29</td>
<td>0.0266</td>
</tr>
<tr>
<td>AK055699</td>
<td>1.24</td>
<td>0.10</td>
<td>0.0276</td>
</tr>
<tr>
<td>pSNT1</td>
<td>0.51</td>
<td>0.31</td>
<td>0.0293</td>
</tr>
<tr>
<td>EpCAM</td>
<td>0.62</td>
<td>0.22</td>
<td>0.0310</td>
</tr>
<tr>
<td>WISP1</td>
<td>1.39</td>
<td>0.16</td>
<td>0.0338</td>
</tr>
<tr>
<td>VEGFC</td>
<td>0.62</td>
<td>0.23</td>
<td>0.0364</td>
</tr>
<tr>
<td>TK1</td>
<td>0.73</td>
<td>0.15</td>
<td>0.0382</td>
</tr>
</tbody>
</table>
TABLE 6-continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Relative Risk (RR)</th>
<th>SE Relative Risk</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFKBp65</td>
<td>1.32</td>
<td>0.14</td>
<td>0.0384</td>
</tr>
<tr>
<td>BRCA2</td>
<td>0.66</td>
<td>0.20</td>
<td>0.0404</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>0.60</td>
<td>0.25</td>
<td>0.0417</td>
</tr>
<tr>
<td>EGFR</td>
<td>0.72</td>
<td>0.16</td>
<td>0.0436</td>
</tr>
</tbody>
</table>

TABLE 7

<table>
<thead>
<tr>
<th>Gene</th>
<th>Relative Risk (RR)</th>
<th>SE Relative Risk</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRAME</td>
<td>0.75</td>
<td>0.10</td>
<td>0.0045</td>
</tr>
<tr>
<td>Contig51037</td>
<td>0.75</td>
<td>0.11</td>
<td>0.0060</td>
</tr>
<tr>
<td>Bcl2</td>
<td>2.11</td>
<td>0.28</td>
<td>0.0075</td>
</tr>
<tr>
<td>NIP1A</td>
<td>0.42</td>
<td>0.34</td>
<td>0.0117</td>
</tr>
<tr>
<td>IGF1R</td>
<td>1.92</td>
<td>0.26</td>
<td>0.0117</td>
</tr>
<tr>
<td>FOXM1</td>
<td>0.54</td>
<td>0.24</td>
<td>0.0119</td>
</tr>
<tr>
<td>EPXH1</td>
<td>0.43</td>
<td>0.33</td>
<td>0.0120</td>
</tr>
<tr>
<td>Ki67</td>
<td>0.60</td>
<td>0.21</td>
<td>0.0160</td>
</tr>
<tr>
<td>CDC25B</td>
<td>0.41</td>
<td>0.38</td>
<td>0.0200</td>
</tr>
<tr>
<td>VEGF</td>
<td>0.45</td>
<td>0.37</td>
<td>0.0208</td>
</tr>
<tr>
<td>CTSS</td>
<td>0.32</td>
<td>0.53</td>
<td>0.0328</td>
</tr>
<tr>
<td>DIABLO</td>
<td>2.91</td>
<td>0.50</td>
<td>0.0328</td>
</tr>
<tr>
<td>p27</td>
<td>1.83</td>
<td>0.28</td>
<td>0.0341</td>
</tr>
<tr>
<td>CDH1</td>
<td>0.57</td>
<td>0.27</td>
<td>0.0352</td>
</tr>
<tr>
<td>IGFBI</td>
<td>0.45</td>
<td>0.40</td>
<td>0.0499</td>
</tr>
</tbody>
</table>

The binary and time-to-event analyses, with few exceptions, identified the same genes as prognostic markers. For example, comparison of Tables 4 and 6 shows that, with the exception of a single gene, the two analyses generated the same list of top 15 markers (as defined by the smallest p values). Furthermore, when both analyses identified the same gene, they were concordant with respect to the direction (positive or negative sign) of the correlation with survival/recurrence. Overall, these results strengthen the conclusion that the identified markers have significant prognostic value.

For Cox models comprising more than two genes (multivariate models), stepwise entry of each individual gene into the model is performed, where the first gene entered is pre-selected from among those genes having significant univariate p-values, and the gene selected for entry into the model at each subsequent step is the gene that best improves the fit of the model to the data. This analysis can be performed with any total number of genes. In the analysis the results of which are shown below, stepwise entry was performed for up to 10 genes.

Multivariate analysis is performed using the following equation:

\[ RR = \exp(\text{coef}(\text{gene 1}) \times \text{Ct}(\text{gene 1}) + \text{coef}(\text{gene 2}) \times \text{Ct}(\text{gene 2}) + \ldots) \]

In this equation, coefficients for genes that are predictors of beneficial outcome are positive numbers and coefficients for genes that are predictors of unfavorable outcome are negative numbers. The "Ct" values in the equation are Ct values, i.e. reflect the difference between the average normalized Ct value for a population and the normalized Ct measured for the patient in question. The convention used in the present analysis has been that Ct values below and above the population average have positive signs and negative signs, respectively (reflecting greater or lesser mRNA abundance). The relative risk (RR) calculated by solving this equation will indicate if the patient has an enhanced or reduced chance of long-term survival without cancer recurrence.

Multivariate Gene Analysis of 147 Patients with Invasive Breast Carcinoma

(a) A multivariate stepwise analysis, using the Cox Proportional Hazards Model, was performed on the gene expression data obtained for all 147 patients with invasive breast carcinoma. Genes CEGPI, FOXM1, STK15 and PRAME were excluded from this analysis. The following ten-gene sets have been identified by this analysis as having particularly strong predictive value of patient survival without cancer recurrence following surgical removal of primary tumor.

1. Bcl2, cyclinG1, NFkB65, NME1, EPHX1, TOP2B, DR5, TERC, Src, DIABLO;
2. Ki67, XIAP, hENT1, TS, CD9, p27, cyclinG1, p52, NFkB65, CYP3A4;
3. GSTM1, XIAP, Ki67, TS, cyclinG1, p27, CYP3A4, p52, NFkB65, ErbB3;
4. PR, NME1, XIAP, upa, cyclinG1, Contig51037, TERC, EPHX1, ALDH1A3, CTSL;
5. CA9, NME1, TERC, cyclinG1, EPHX1, DPYD, Src, TOP2B, NFkB65, VEGFC;
6. TFRC, XIAP, Ki67, TS, cyclinG1, p27, CYP3A4, p52, ErbB3, NFkB65.

(b) A multivariate stepwise analysis, using the Cox Proportional Hazards Model, was performed on the gene expression data obtained for all 147 patients with invasive breast carcinoma, using an interrogation set including a reduced number of genes. The following ten-gene sets have been identified by this analysis as having particularly strong predictive value of patient survival without cancer recurrence following surgical removal of primary tumor.

1. Bcl2, PRAME, cyclinG1, FOXM1, NFkB65, TS, XIAP, Ki67, CYP3A4, p27;
[0531] 2. FOXM1, cyclinG1, XIAP, Contig51037, PRAME, TS, Ki67, PDGFRα, p27, NFKBp65;
[0532] 3. PRAME, FOXM1, cyclinG1, XIAP, Contig51037, TS, Ki67, PDGFRα, p27, NFKBp65;
[0533] 4. Ki67, XIAP, PRAME, hENT1, Contig51037, TS, CD9, p27, ErbB3, cyclinG1;
[0534] 5. STK15, XIAP, PRAME, PLAUR, p27, CTS1, CD18, PREP, p53, RPS6 KB1;
[0535] 6. GSTM1, XIAP, PRAME, p27, Contig51037, ErbB3, GSTp, EREG, ID1, PLAUR;
[0536] 7. PR, PRAME, NME1, XIAP, PLAUR, cyclinG1, Contig51037, TERC, EPHX1, DR5;
[0537] 8. CA9, FOXM1, cyclinG1, XIAP, TS, Ki67, NFKBp65, CYP3A4, GSTM3, p27;
[0538] 9. TFRC, XIAP, PRAME, p27, Contig51037, ErbB3, DPYD, TERC, NME1, VEGFC;
[0539] 10. CECP1, PRAME, hENT1, XIAP, Contig51037, ErbB3, DPYD, NFKBp65, ID1, TS.

(0540) Multivariate Analysis of Patients with ER Positive Invasive Breast Carcinoma

[0541] A multivariate stepwise analysis, using the Cox Proportional Hazards Model, was performed on the genome expression data obtained for patients with ER positive invasive breast carcinoma. The following ten-gene sets have been identified by this analysis as having particularly strong predictive value of patient survival without cancer recurrence following surgical removal of primary tumor.

[0542] 1. PRAME, p27, IGBP2, HIF1A, TIMP2, ILT2, CYP3A4, ID1, EstR1, DIABLO;
[0543] 2. Contig51037, EPHX1, Ki67, TIMP2, cyclinG1, DPYD, CYP3A4, TP, AIB1, CYP2C8;
[0544] 3. Bcl2, hENT1, FOXM1, Contig51037, cyclinG1, Contig46653, PTEN, CYP3A4, TIMP2, AREG;
[0545] 4. HIF1A, PRAME, p27, IGBP2, TIMP2, ILT2, CYP3A4, ID1, EstR1, DIABLO;
[0546] 5. IGF1R, PRAME, EPHX1, Contig51037, cyclinG1, Bcl2, NME1, PTEN, TBP, TIMP2;
[0547] 6. FOXM1, Contig51037, VEGFC, TBP, HIF1A, DPYD, RAD51C, DCR3, cyclinG1, BAG1;
[0548] 7. EPHX1, Contig51037, Ki67, TIMP2, cyclinG1, DPYD, CYP3A4, TP, AIB1, CYP2C8;
[0549] 8. Ki67, VEGFC, VDR, GSTM3, p27, uPA, ITGA7, THOC, TERC, Pim1;
[0550] 9. CDC25B, Contig51037, hENT1, Bcl2, ILAG, TERC, NME1, uPA, ID1, CYP;
[0551] 10. VEGFC, Ki67, VDR, GSTM3, p27, uPA, ITGA7, THOC, TERC, Pim1;
[0552] 11. CTSS, PRAME, p27, IGBP2, EPHX1, CTS1, BAD, DR5, DCR3, XIAP;
[0553] 12. DIABLO, Ki67, hENT1, TIMP2, ID1, p27, KRT19, IGBP2, TS, PDGFB;
[0554] 13. p27, PRAME, IGBP2, HIF1A, TIMP2, ILT2, CYP3A4, ID1, EstR1, DIABLO;
[0555] 14. CDEI1, PRAME, VEGFC, HIF1A, DPYD, TIMP2, CYP3A4, EstR1, RBP4, p27;
[0556] 15. IGBP2, PRAME, p27, Bcl2, XIAP, EstR1, Ki67, TS, Src, VEGF;
[0557] 16. GSTM3, PRAME, p27, IGBP2, XIAP, FGF2, hENT1, PTEN, EstR1, APC;
[0558] 17. hENT1, Bcl2, FOXM1, Contig51037, CyclinG1, Contig46653, PTEN, CYP3A4, TIMP2, AREG;
[0559] 18. STK15, VEGFC, PRAME, p27, GCLC, hENT1, ID1, TIMP2, EstR1, MCP1;
[0560] 19. NME1, PRAM, p27, IGBP2, XIAP, PTEN, hENT1, Bcl2, CYP3A4, HLAG;
[0561] 20. VDR, Bcl2, p27, hENT1, p53, PI3KC2A, EIF4E, TFRC, MCM3, ID1;
[0562] 21. EIF4E, Contig51037, EPHX1, cyclinG1, Bcl2, DR5, TBP, PTEN, NME1, HER2;
[0563] 22. CCNB1, PRAME, VEGFC, HIF1A, hENT1, GCLC, TIMP2, ID1, p27, uPA;
[0564] 23. ID1, PRAME, DIABLO, hENT1, p27, PDGFRα, NME1, BIN1, BRCA1, TP;
[0565] 24. FBXO5, PRAME, IGBP2, p27, GSTM3, hENT1, XIAP, FGF2, TS, PTEN;
[0566] 25. GUS, HLA1A, VEGFC, GSTM3, DPYD, hENT1, FBXO5, CA9, CYP, KRT18;
[0567] 26. Bcl2, Bcl2, hENT1, Contig51037, HLAG, CD9, ID1, BRCA1, BIN1, HBEFG.

[0568] It is noteworthy that many of the foregoing gene sets include genes that alone did not have sufficient predictive value to qualify as prognostic markers under the standards discussed above, but in combination with other genes, their presence provides valuable information about the likelihood of long-term patient survival without cancer recurrence

[0569] All references cited throughout the disclosure are hereby expressly incorporated by reference.

[0570] While the present invention has been described with reference to what are considered to be the specific embodiments, it is to be understood that the invention is not limited to such embodiments. To the contrary, the invention is intended to cover various modifications and equivalents included within the spirit and scope of the appended claims. For example, while the disclosure focuses on the identification of various breast cancer associated genes and gene sets, and on the diagnosis and treatment of breast cancer, similar genes, gene sets and methods concerning other types of cancer are specifically within the scope herein.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ADD3 (adducin 3 gamma)*</td>
</tr>
<tr>
<td>2. AKT1/Protein Kinase B</td>
</tr>
<tr>
<td>3. PRAME (protein)</td>
</tr>
</tbody>
</table>
### TABLE 1-continued

| 71. FasL | 72. FGF R 1* |
| 73. FGF2 [bFGF] | 74. 35BP1 |
| 75. 35BP2 | 76. GALC |
| (galactosylceramidase)* | Gamma-GCS (glutamyl cysteine synthetase) |
| 77. GABA3 | 78. geranylgeranyl pyrophosphate synthetase |
| 80. G-CSF | 81. GPC3 |
| 82. gravin* [AK AP258] | 83. GRO1 oncogene alpha |
| 84. Grb7 | 85. GST-alpha |
| 86. GST-pi | 87. Hsc 70 |
| 88. HB-EGF | 89. HE4-extracellular |
| &nbsp; | &nbsp; |
| **Proteinase Inhibitor** | &nbsp; |
| 90. heparocyte nuclear factor 3 | 91. HER-2 |
| 92. HGF/Scatter factor | 93. hAP1 |
| 94. hAP2 | 95. HIF-1 |
| 96. human kalikrein 10 | 97. MLH1 |
| 98. hsp 27 | 99. human chorionic gonadotropin/CGA |
| **Human Extracellular** | **Protein S1-5** |
| 100. human Extracellular | &nbsp; |

### TABLE 1-continued

| 3. AKT 2 | 4. AKT 3 |
| 5. Aldehyde dehydrogenase IA1 | 6. Aldehyde dehydrogenase IA3 |
| 7. amphiregulin | 8. APC |
| 9. ARG | 10. ATM |
| 13. Be2 | 14. Bel-x |
| 15. BRR | 16. BCRP |
| 17. BRCA-1 | 18. BRCA-2 |
| 21. Cathepsin G | 22. Cathepsin L |
| 23. CD3 | 24. CD9 |
| 25. CD18 | 26. CD31 |
| 27. CD44 | 28. CD98 |
| 29. CD98/KAI-1 | 30. Cd25A |
| 31. Cd25B | 32. CGA |
| 33. COX2 | 34. CSF-1 |
| 35. CSF-1R/finu | 36. cIAP1 |
| 37. cIAP2 | 38. c-abl |
| 39. c-kit | 40. c-kit L |
| 41. c-met | 42. c-myc |
| 43. c-Ne | 44. cryptochrome1* |
| 45. c-Src | 46. Cyclin D1 |
| 47. CYP1B1 | 48. CYP2C9 |
| 49. Cytokinin 5* | 50. Cytokinin 17 |
| 51. Cytokinin 18 | 52. DAP-Kinase-1 |
| 53. DMR | 54. DIABLO |
| **Dehydroprydinase** | **Dehydrogenase** |
| 55. Dihydroprydinase | 56. EGF |
| &nbsp; | 57. ECadherin/CDH1 |
| 58. ELF 3* | 59. Endothelin |
| 60. Epigalactol | 61. ER-alpha |
| 62. ErbB-1 | 63. ErbB-2 |
| 64. ErbB-3 | 65. ErbB-4 |
| 66. ER-Beta | 67. Eukaryotic Translation Initiation Factor |
| &nbsp; | &nbsp; |
| 68. ELT4E | 69. farnesyl pyrophosphate synthetase |
| 70. FAS (CD95) | &nbsp; |
TABLE 1-continued

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 2-continued</td>
<td>Accession No.</td>
<td>Forward Primer</td>
<td>Reverse Primer</td>
<td>Amplicon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>SEQ ID NO.</td>
<td>SEQ ID NO.</td>
<td>SEQ ID NO.</td>
<td>SEQ ID NO.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------</td>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCB2</td>
<td>NM_000392</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCB3</td>
<td>NM_003786</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_005845</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABL1</td>
<td>NM_005157</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABL2</td>
<td>NM_005158</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTB</td>
<td>NM_001101</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT1</td>
<td>NM_005163</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT3</td>
<td>NM_005465</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALDH1</td>
<td>NM_000689</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALDH1A2</td>
<td>NM_000693</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>NM_000390</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AREG</td>
<td>NM_001657</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2M</td>
<td>NM_004040</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAK1</td>
<td>NM_001186</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAX</td>
<td>NM_004324</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCL2</td>
<td>NM_000633</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCL2L1</td>
<td>NM_001191</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIRC3</td>
<td>NM_001165</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIRC4</td>
<td>NM_001167</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1B4</td>
<td>NM_001160</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2A</td>
<td>NM_007295</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2A</td>
<td>NM_000059</td>
<td>70</td>
<td>71</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCND1</td>
<td>NM_001756</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH3</td>
<td>NM_000374</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH6</td>
<td>NM_001251</td>
<td>79</td>
<td>80</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC25A</td>
<td>NM_001789</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH1</td>
<td>NM_004360</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK1A</td>
<td>NM_003099</td>
<td>88</td>
<td>89</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK1B</td>
<td>NM_004064</td>
<td>91</td>
<td>92</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK2A</td>
<td>NM_000077</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP1B1</td>
<td>NM_000014</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEFR</td>
<td>NM_000791</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYPD</td>
<td>NM_000110</td>
<td>103</td>
<td>104</td>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>NM_001953</td>
<td>106</td>
<td>107</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>NM_005228</td>
<td>109</td>
<td>110</td>
<td>111</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIF4E</td>
<td>NM_001968</td>
<td>112</td>
<td>113</td>
<td>114</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*NCI 60 drug Sens/Resist Marker
In Cluster Defining tumor subclass
Jan. 19, 2002
### TABLE 2-continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession No.</th>
<th>Forward Primer SEQ ID NO.</th>
<th>Reverse Primer SEQ ID NO.</th>
<th>Amplicon SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERBB2</td>
<td>NM_004448</td>
<td>115</td>
<td>116</td>
<td>117</td>
</tr>
<tr>
<td>ERBB3</td>
<td>NM_001992</td>
<td>118</td>
<td>119</td>
<td>120</td>
</tr>
<tr>
<td>ESR1</td>
<td>NM_000125</td>
<td>121</td>
<td>122</td>
<td>123</td>
</tr>
<tr>
<td>ESR2</td>
<td>NM_001437</td>
<td>124</td>
<td>125</td>
<td>126</td>
</tr>
<tr>
<td>GAPD</td>
<td>NM_002046</td>
<td>127</td>
<td>128</td>
<td>129</td>
</tr>
<tr>
<td>GATA3</td>
<td>NM_002051</td>
<td>130</td>
<td>131</td>
<td>132</td>
</tr>
<tr>
<td>GRB7</td>
<td>NM_005310</td>
<td>133</td>
<td>134</td>
<td>135</td>
</tr>
<tr>
<td>GRO1</td>
<td>NM_001511</td>
<td>136</td>
<td>137</td>
<td>138</td>
</tr>
<tr>
<td>GSTP1</td>
<td>NM_000852</td>
<td>139</td>
<td>140</td>
<td>141</td>
</tr>
<tr>
<td>GUSB</td>
<td>NM_000181</td>
<td>142</td>
<td>143</td>
<td>144</td>
</tr>
<tr>
<td>hHGF</td>
<td>M29145</td>
<td>145</td>
<td>146</td>
<td>147</td>
</tr>
<tr>
<td>HNF3A</td>
<td>NM_004496</td>
<td>148</td>
<td>149</td>
<td>150</td>
</tr>
<tr>
<td>ID2</td>
<td>NM_002166</td>
<td>151</td>
<td>152</td>
<td>153</td>
</tr>
<tr>
<td>IGFBP3</td>
<td>NM_000598</td>
<td>154</td>
<td>155</td>
<td>156</td>
</tr>
<tr>
<td>ITGA7</td>
<td>NM_002206</td>
<td>160</td>
<td>161</td>
<td>162</td>
</tr>
<tr>
<td>ITGB2</td>
<td>NM_002211</td>
<td>163</td>
<td>164</td>
<td>165</td>
</tr>
<tr>
<td>KDR</td>
<td>NM_002253</td>
<td>166</td>
<td>167</td>
<td>168</td>
</tr>
<tr>
<td>KIT</td>
<td>NM_002222</td>
<td>169</td>
<td>170</td>
<td>171</td>
</tr>
<tr>
<td>KITLG</td>
<td>NM_000899</td>
<td>172</td>
<td>173</td>
<td>174</td>
</tr>
<tr>
<td>KRT17</td>
<td>NM_000422</td>
<td>175</td>
<td>176</td>
<td>177</td>
</tr>
<tr>
<td>KRT5</td>
<td>NM_000424</td>
<td>178</td>
<td>179</td>
<td>180</td>
</tr>
<tr>
<td>LPL</td>
<td>NM_000237</td>
<td>181</td>
<td>182</td>
<td>183</td>
</tr>
<tr>
<td>MET</td>
<td>NM_000245</td>
<td>184</td>
<td>185</td>
<td>186</td>
</tr>
<tr>
<td>NKI67</td>
<td>NM_002417</td>
<td>187</td>
<td>188</td>
<td>189</td>
</tr>
<tr>
<td>MVP</td>
<td>NM_017450</td>
<td>190</td>
<td>191</td>
<td>192</td>
</tr>
<tr>
<td>MYC</td>
<td>NM_002467</td>
<td>193</td>
<td>194</td>
<td>195</td>
</tr>
<tr>
<td>PDGFA</td>
<td>NM_002607</td>
<td>196</td>
<td>197</td>
<td>198</td>
</tr>
<tr>
<td>PDGFB</td>
<td>NM_002608</td>
<td>199</td>
<td>200</td>
<td>201</td>
</tr>
<tr>
<td>PDGFC</td>
<td>NM_016205</td>
<td>202</td>
<td>203</td>
<td>204</td>
</tr>
<tr>
<td>PDGFR</td>
<td>NM_006206</td>
<td>205</td>
<td>206</td>
<td>207</td>
</tr>
<tr>
<td>PDGFRB</td>
<td>NM_002609</td>
<td>208</td>
<td>209</td>
<td>210</td>
</tr>
<tr>
<td>PGK1</td>
<td>NM_000291</td>
<td>211</td>
<td>212</td>
<td>213</td>
</tr>
<tr>
<td>PGR</td>
<td>NM_000926</td>
<td>214</td>
<td>215</td>
<td>216</td>
</tr>
<tr>
<td>PIN1</td>
<td>NM_006221</td>
<td>217</td>
<td>218</td>
<td>219</td>
</tr>
<tr>
<td>PLAU</td>
<td>NM_002658</td>
<td>220</td>
<td>221</td>
<td>222</td>
</tr>
</tbody>
</table>

### TABLE 2-continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession No.</th>
<th>Forward Primer SEQ ID NO.</th>
<th>Reverse Primer SEQ ID NO.</th>
<th>Amplicon SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>NM_006347</td>
<td>223</td>
<td>224</td>
<td>225</td>
</tr>
<tr>
<td>P53</td>
<td>NM_000314</td>
<td>226</td>
<td>227</td>
<td>228</td>
</tr>
<tr>
<td>PTGS2</td>
<td>NM_009963</td>
<td>229</td>
<td>230</td>
<td>231</td>
</tr>
<tr>
<td>RBP4</td>
<td>NM_006744</td>
<td>232</td>
<td>233</td>
<td>234</td>
</tr>
<tr>
<td>RELA</td>
<td>NM_021975</td>
<td>235</td>
<td>236</td>
<td>237</td>
</tr>
<tr>
<td>RPL19</td>
<td>NM_000981</td>
<td>238</td>
<td>239</td>
<td>240</td>
</tr>
<tr>
<td>RPLD0</td>
<td>NM_001002</td>
<td>241</td>
<td>242</td>
<td>243</td>
</tr>
<tr>
<td>SCG5P1</td>
<td>NM_025208</td>
<td>244</td>
<td>245</td>
<td>246</td>
</tr>
<tr>
<td>SERPINE1</td>
<td>NM_00602</td>
<td>247</td>
<td>248</td>
<td>249</td>
</tr>
<tr>
<td>SLC19A1</td>
<td>NM_003056</td>
<td>250</td>
<td>251</td>
<td>252</td>
</tr>
<tr>
<td>TBP</td>
<td>NM_003194</td>
<td>253</td>
<td>254</td>
<td>255</td>
</tr>
<tr>
<td>TFF1</td>
<td>NM_003225</td>
<td>256</td>
<td>257</td>
<td>258</td>
</tr>
<tr>
<td>TFFC</td>
<td>NM_003234</td>
<td>259</td>
<td>260</td>
<td>261</td>
</tr>
<tr>
<td>TR1</td>
<td>NM_003258</td>
<td>262</td>
<td>263</td>
<td>264</td>
</tr>
<tr>
<td>THPRF6</td>
<td>NM_000433</td>
<td>265</td>
<td>266</td>
<td>267</td>
</tr>
<tr>
<td>THPS6</td>
<td>NM_00639</td>
<td>268</td>
<td>269</td>
<td>270</td>
</tr>
<tr>
<td>TOP2A</td>
<td>NM_001067</td>
<td>271</td>
<td>272</td>
<td>273</td>
</tr>
<tr>
<td>TOP2B</td>
<td>NM_001069</td>
<td>274</td>
<td>275</td>
<td>276</td>
</tr>
<tr>
<td>TP53</td>
<td>NM_000546</td>
<td>277</td>
<td>278</td>
<td>279</td>
</tr>
<tr>
<td>TYMS</td>
<td>NM_001071</td>
<td>280</td>
<td>281</td>
<td>282</td>
</tr>
<tr>
<td>VEGF</td>
<td>NM_003376</td>
<td>283</td>
<td>284</td>
<td>285</td>
</tr>
</tbody>
</table>

[0572]

### TABLE 3

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession No.</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AK055699</td>
<td>AK055699</td>
<td>206</td>
</tr>
<tr>
<td>BAG1</td>
<td>NM_004323</td>
<td>207</td>
</tr>
<tr>
<td>BBC3</td>
<td>NM_014417</td>
<td>208</td>
</tr>
<tr>
<td>Bc12</td>
<td>NM_000633</td>
<td>209</td>
</tr>
<tr>
<td>BRCA2</td>
<td>NM_000059</td>
<td>210</td>
</tr>
<tr>
<td>CA9</td>
<td>NM_001216</td>
<td>211</td>
</tr>
<tr>
<td>CCNB1</td>
<td>NM_031966</td>
<td>212</td>
</tr>
<tr>
<td>CDC25B</td>
<td>NM_021874</td>
<td>213</td>
</tr>
<tr>
<td>CEGP1</td>
<td>NM_020974</td>
<td>214</td>
</tr>
<tr>
<td>Chk1</td>
<td>NM_001274</td>
<td>215</td>
</tr>
<tr>
<td>Chk2</td>
<td>NM_007194</td>
<td>216</td>
</tr>
</tbody>
</table>

---

**[0572]**
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP3A4</td>
<td>NM_017460</td>
<td>297</td>
<td>PDGFβ</td>
<td>NM_002608</td>
<td>334</td>
</tr>
<tr>
<td>DIABLO</td>
<td>NM_019887</td>
<td>298</td>
<td>Upa</td>
<td>NM_002658</td>
<td>335</td>
</tr>
<tr>
<td>DPYD</td>
<td>NM_000110</td>
<td>299</td>
<td>TBP</td>
<td>NM_003194</td>
<td>336</td>
</tr>
<tr>
<td>EGFR</td>
<td>NM_005228</td>
<td>300</td>
<td>PDGFRa</td>
<td>NM_006206</td>
<td>337</td>
</tr>
<tr>
<td>EpCAM</td>
<td>NM_002354</td>
<td>301</td>
<td>Pin1</td>
<td>NM_006221</td>
<td>338</td>
</tr>
<tr>
<td>EPHX1</td>
<td>NM_000120</td>
<td>302</td>
<td>CYP</td>
<td>NM_006347</td>
<td>339</td>
</tr>
<tr>
<td>EstR1</td>
<td>NM_000125</td>
<td>303</td>
<td>RBP4</td>
<td>NM_006744</td>
<td>340</td>
</tr>
<tr>
<td>FOXM1</td>
<td>NM_021953</td>
<td>304</td>
<td>BRCA1</td>
<td>NM_007295</td>
<td>341</td>
</tr>
<tr>
<td>GATA3</td>
<td>NM_002051</td>
<td>305</td>
<td>APC</td>
<td>NM_000038</td>
<td>342</td>
</tr>
<tr>
<td>GSTM1</td>
<td>NM_000561</td>
<td>306</td>
<td>GUS</td>
<td>NM_000101</td>
<td>343</td>
</tr>
<tr>
<td>GSTM3</td>
<td>NM_000849</td>
<td>307</td>
<td>CD18</td>
<td>NM_000211</td>
<td>344</td>
</tr>
<tr>
<td>hENT1</td>
<td>NM_004955</td>
<td>308</td>
<td>PTEN</td>
<td>NM_000314</td>
<td>345</td>
</tr>
<tr>
<td>HIP1A</td>
<td>NM_001530</td>
<td>309</td>
<td>P53</td>
<td>NM_000546</td>
<td>346</td>
</tr>
<tr>
<td>HNF3A</td>
<td>NM_004496</td>
<td>310</td>
<td>ALDH1A3</td>
<td>NM_006935</td>
<td>347</td>
</tr>
<tr>
<td>ID1</td>
<td>NM_002165</td>
<td>311</td>
<td>GSTp</td>
<td>NM_000852</td>
<td>348</td>
</tr>
<tr>
<td>IGF1R</td>
<td>NM_000875</td>
<td>312</td>
<td>TOP2B</td>
<td>NM_001068</td>
<td>349</td>
</tr>
<tr>
<td>Ki-67</td>
<td>NM_002417</td>
<td>313</td>
<td>TS</td>
<td>NM_001071</td>
<td>350</td>
</tr>
<tr>
<td>NFkBp65</td>
<td>NM_021975</td>
<td>314</td>
<td>Bclx</td>
<td>NM_001191</td>
<td>351</td>
</tr>
<tr>
<td>NME1</td>
<td>NM_000269</td>
<td>315</td>
<td>AREG</td>
<td>NM_001657</td>
<td>352</td>
</tr>
<tr>
<td>p27</td>
<td>NM_004064</td>
<td>316</td>
<td>TP</td>
<td>NM_001953</td>
<td>353</td>
</tr>
<tr>
<td>PI3KC2A</td>
<td>NM_002645</td>
<td>317</td>
<td>EIF4E</td>
<td>NM_001968</td>
<td>354</td>
</tr>
<tr>
<td>PR</td>
<td>NM_000926</td>
<td>318</td>
<td>ErbB3</td>
<td>NM_001982</td>
<td>355</td>
</tr>
<tr>
<td>PRAME</td>
<td>NM_006115</td>
<td>319</td>
<td>EREG</td>
<td>NM_001432</td>
<td>356</td>
</tr>
<tr>
<td>Ps2</td>
<td>NM_003225</td>
<td>320</td>
<td>GCLC</td>
<td>NM_001498</td>
<td>357</td>
</tr>
<tr>
<td>RP66X1</td>
<td>NM_003161</td>
<td>321</td>
<td>CD9</td>
<td>NM_001769</td>
<td>358</td>
</tr>
<tr>
<td>src</td>
<td>NM_004383</td>
<td>322</td>
<td>HH-EGF</td>
<td>NM_001945</td>
<td>359</td>
</tr>
<tr>
<td>STK15</td>
<td>NM_003600</td>
<td>323</td>
<td>IGFBP2</td>
<td>NM_000597</td>
<td>360</td>
</tr>
<tr>
<td>Surv</td>
<td>NM_001168</td>
<td>324</td>
<td>CTSL</td>
<td>NM_001912</td>
<td>361</td>
</tr>
<tr>
<td>TRPC</td>
<td>NM_003234</td>
<td>325</td>
<td>PREP</td>
<td>NM_002726</td>
<td>362</td>
</tr>
<tr>
<td>TGFβ3</td>
<td>NM_003239</td>
<td>326</td>
<td>CYP3A4</td>
<td>NM_017460</td>
<td>363</td>
</tr>
<tr>
<td>TK1</td>
<td>NM_003250</td>
<td>327</td>
<td>ILT-2</td>
<td>NM_006669</td>
<td>364</td>
</tr>
<tr>
<td>VDR</td>
<td>NM_000376</td>
<td>328</td>
<td>MCM3</td>
<td>NM_002388</td>
<td>365</td>
</tr>
<tr>
<td>VEGFC</td>
<td>NM_005429</td>
<td>329</td>
<td>KRT19</td>
<td>NM_002276</td>
<td>366</td>
</tr>
<tr>
<td>WISP1</td>
<td>NM_003882</td>
<td>330</td>
<td>KRT18</td>
<td>NM_000224</td>
<td>367</td>
</tr>
<tr>
<td>X1AP</td>
<td>NM_001167</td>
<td>331</td>
<td>TIMP2</td>
<td>NM_003255</td>
<td>368</td>
</tr>
<tr>
<td>YB-1</td>
<td>NM_004559</td>
<td>332</td>
<td>BAD</td>
<td>NM_004322</td>
<td>369</td>
</tr>
<tr>
<td>ITGA7</td>
<td>NM_002206</td>
<td>333</td>
<td>CYP2C8</td>
<td>NM_0030878</td>
<td>370</td>
</tr>
<tr>
<td>Gene</td>
<td>Accession No.</td>
<td>Seq ID No.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCR3</td>
<td>NM_016434</td>
<td>371</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAUR</td>
<td>NM_002659</td>
<td>372</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI3KC2A</td>
<td>NM_002645</td>
<td>373</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGFP2</td>
<td>NM_002006</td>
<td>374</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-G</td>
<td>NM_002127</td>
<td>375</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIB1</td>
<td>NM_006534</td>
<td>376</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCP1</td>
<td>NM_002982</td>
<td>377</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig46653</td>
<td>Contig46653</td>
<td>378</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 3-continued**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession No.</th>
<th>Seq ID No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RhoC</td>
<td>NM_005167</td>
<td>379</td>
</tr>
<tr>
<td>DR5</td>
<td>NM_003042</td>
<td>380</td>
</tr>
<tr>
<td>RAD51C</td>
<td>NM_058216</td>
<td>381</td>
</tr>
<tr>
<td>BIN1</td>
<td>NM_004305</td>
<td>382</td>
</tr>
<tr>
<td>VDR</td>
<td>NM_000376</td>
<td>383</td>
</tr>
<tr>
<td>TERC</td>
<td>US6046</td>
<td>384</td>
</tr>
</tbody>
</table>

**SEQUENCE LISTING**

<160> NUMBER OF SEQ ID NOS: 304

<210> SEQ ID NO 1
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1
gtoccagag ccctcatcct

<210> SEQ ID NO 2
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2
ccgggctggt gttctocata

<210> SEQ ID NO 3
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3
gtoccagag ccctcatctg ttggtcag cattgtcgag acattgcct atgagacaa

cagcogqq

<210> SEQ ID NO 4
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4
tctcgtgcrr cytoastg

<210> SEQ ID NO 5
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 5

cgattgttctt tgctttctat gtg 23

<210> SEQ ID NO 6
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

tcgagttgcc ctgcaactct gtgtgtgctg tcag ggtcggtgtaca tgaatggcata ggtgatggcga tigaagiaccala gacgitaticag gtggcc.caca 60
tgsagagcsa agscaasctcg 79

<210> SEQ ID NO 7
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

aggggatgcctttgacacat 20

<210> SEQ ID NO 8
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

aaaaactgctgatggttgc 20

<210> SEQ ID NO 9
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

aggggatgccttggacatcgatggtttgtccccaatggcccattaaccatgggaacgacatggaaattgagtt 60
gtttt 65

<210> SEQ ID NO 10
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

tcatctctgctgatgtcttcctcct 22

<210> SEQ ID NO 11
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

ccggtgctgcgactcagcaca 20

<210> SEQ ID NO 12
<211> LENGTH: 91
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12
tcattctgga gatcttactc cttggcaga acctaggtcc ctctgtcctg gctggagctg 60
cattcatggt ctttgagtatt caaactcaacg g 91

<210> SEQ ID NO 13
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 13
agagcctggaa atctacaact 20

<210> SEQ ID NO 14
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 14
agagccttg gagaagaagt 20

<210> SEQ ID NO 15
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 15
agagcctgga atctacaact cgagtcctag tggttttcaca ctctgctct ctctctcagg 60
ggtct 66

<210> SEQ ID NO 16
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 16
gccagaga ggtcttgtaa cttca 24

<210> SEQ ID NO 17
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 17
gtttcaagg cttggtggat tt 22

<210> SEQ ID NO 18
<211> LENGTH: 94
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 18
gccagagaa gttgtatagaa cttcatgcag catgttgagca gttgatcc ctctgacggc 60
cctctttgg tgtaatccac casgcctttg amac 94
cgcagtgacg ctgagtatct g 21

<210> SEQ ID NO 20
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20
tgccoaaggc tactctcaact t 21

<210> SEQ ID NO 21
<211> LENGTH: 80
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21
cgcagtgacg ctgagtatct gtcgacagt ctaactcagt gacgacctct ggtgagaa 60
agtgagagta gcccctgggca 80

<210> SEQ ID NO 22
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22
cagcagatgt ggatcagcaa g 21

<210> SEQ ID NO 23
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 23
gcatattcg gtygagat 10

<210> SEQ ID NO 24
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 24
ccgagatgt ggtcagcaga gcaggtatg gacgagtcgg gcccctcact cgctcaccgc 60
asatgc 66

<210> SEQ ID NO 25
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25
cgcttcattg ggcgctgagat 20

<210> SEQ ID NO 26
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26
tcccggtaca ccaagttott 20
<210> SEQ ID NO 27
<211> LENGTH: 71
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 27

cgtctctag gcgtcagat tgtgtcagc cttgacactc tgtgactgga gagaagcgtg 60
gtgaccgg g 71

<210> SEQ ID NO 28
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 28

ttttccttcg cttgacat ctca ac 25

<210> SEQ ID NO 29
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 29

cagacatacg atctctccac ttga 24

<210> SEQ ID NO 30
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 30

ttttccttcg cttgacat ctacatcctg gagaagattg tgtacggtg gttcaagttg 60
agactctaat gttgg 75

<210> SEQ ID NO 31
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 31

gaagagaga aagagagatgt ctaca 25

<210> SEQ ID NO 32
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 32

cgccaggyag atccacaac 18

<210> SEQ ID NO 33
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 33

gaagagaga aagagagatgt ctacaaggc gtaagggcc caagcacgga ttttoagatt 60
gattcctcg ggcg 74
<210> SEQ ID NO 34
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 34

tggtgacat tgggccagga t

<210> SEQ ID NO 35
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 35
gasggygatc ttggtgatct ga

<210> SEQ ID NO 36
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 36
tggtgacat tgggccagga ttcgggccc acgtgggac agcaatctct tcctaaccttc
agatcaacaa gatggccttc

<210> SEQ ID NO 37
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37
ggcacgacag aatggttttc

<210> SEQ ID NO 38
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 38
accactcga ttggtttttg

<210> SEQ ID NO 39
<211> LENGTH: 69
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 39
ggcacgacag aatggttttc tocatagag tcacgggag ccaatggttct agaaacaat
cgagtggt

<210> SEQ ID NO 40
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40
tggtgatgaa atgctttctct ctggtg
<210> SEQ ID NO 41
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 41

ttggtacctg tcatctact cttctga 27

<210> SEQ ID NO 42
<211> LENGTH: 82
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42
ttgagtgaag atgcctctta ttagaaccc ggctcgcgga gcggactatg aactctcaga agagatgtgt acgaaccac aa 82

<210> SEQ ID NO 43
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43
gtctgcttc cvgtgtc 19

<210> SEQ ID NO 44
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44
cgtgagttaa cttgatct ttgga 24

<210> SEQ ID NO 45
<211> LENGTH: 93
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45
gtctgcttc cvgtgtcctag ctgtagctgc gctactctct ctgctgccc ttgagctat ccaggtact ccagagtttc actctc acg 93

<210> SEQ ID NO 46
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46
ccaattccac cattctacct 20

<210> SEQ ID NO 47
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 47
gggacatag accaccac 20

<210> SEQ ID NO 48
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

ccatccac cattcctacg gaggccagga cgctcggggt gttgggtatt gttgggtcat 60
gttcc

<210> SEQ ID NO 49
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

cggcgcgttga cacagact 18

<210> SEQ ID NO 50
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

ttgccgatcg aasacatgtc a 21

<210> SEQ ID NO 51
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 51

cggcgcgttga cacacgactccc ccgcgcaggg tcttctctct cagcagcact gacatgtttt 60
cgacgcgca

<210> SEQ ID NO 52
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 52

cagcagcagcac tgaatcctct tsgaga 25

<210> SEQ ID NO 53
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 53

cctgtgatct aagggctgatt ttgcc

<210> SEQ ID NO 54
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54

cagcagcagcagcac tgaatcctct tsgaga tctcctctct cagcagcaggt gaaaatgcc 60
cattcctctct tggagagcgcgttga
cggcgcgttga cacagact

<210> SEQ ID NO 55
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 55
cttttgtgga acctatggtg aaca 24

<210> SEQ ID NO: 56
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 56
cagcggttga agcgctctct 19

<210> SEQ ID NO: 57
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57
ccttttgtgga acctatggtg aacaatgcaag cagcgcagag cgcgaagggc caggacgcct 60
tcaacgctg 70

<210> SEQ ID NO: 58
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 58
ggatatttcc gttgctcttt tttc 24

<210> SEQ ID NO: 59
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 59
cttctcatca aagcagaaaa atctt 25

<210> SEQ ID NO: 60
<211> LENGTH: 86
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 60
ggatatttcc gttgctcttt tttcaactct ccataaaact cttgaaactc cagagcaaat 60
cgaatatttt ctgctttgat gagaag 86

<210> SEQ ID NO: 61
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61
gcagtggaac gacagggaa agt 23

<210> SEQ ID NO: 62
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
tgcgtggcacc tattttaaag a

<210> SEQ ID NO 63
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 63
gcagttggaa gacacaggaa agtaticca aantgcagat ttatcaacgg ctttttcttt
60
    gaaatagtg ccccgc
77

<210> SEQ ID NO 64
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 64
tgttttgatt cccgggotta
20

<210> SEQ ID NO 65
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 65
caaacgtgc agotctagca aaag
24

<210> SEQ ID NO 66
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 66
tgtttttgatt cccgggotta ccaggtgaga agtgagggag gagaaggca gtgcoccttt
60
tgctagagct gacaggtttg
80

<210> SEQ ID NO 67
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 67
tcaggggyct agaasctctgt
20

<210> SEQ ID NO 68
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 68
caatccagt tgtctgtgag
20

<210> SEQ ID NO 69
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69
tcaagggggct agaattctgt tgtatgggc ccttcaccaaa cagcccccaca gatcaactgg

aatgg

<210> SEQ ID NO: 70
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

agttctgct tgtgcaagatg

<210> SEQ ID NO: 71
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 71

aagtaagct gggtctgctg

<210> SEQ ID NO: 72
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 72

agttctgct tgtgcaagatg gtgcagagct ttatgaagca gtgaagatyg cagcagacc

agottaacct

<210> SEQ ID NO: 73
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73

gcatgttcgt ggcotctaaag a

<210> SEQ ID NO: 74
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74

cggttagat gcacgcttc tc

<210> SEQ ID NO: 75
<211> LENGTH: 69
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 75

gcatgttcgt ggcotctaaag atgaaggaga ccacccccc gactgocgag aagctgtgca
tctaccccg

<210> SEQ ID NO: 76
<211> LENGTH: 69
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76
agatgaagtg gaggacgtt 20

<210> SEQ ID NO 77
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 77
tgctctcgtc atcggaact g 21

<210> SEQ ID NO 78
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78
agatgaagtg gaggacgtt ttcacccggg ccagctgcca gcacagtttgc gcagttcag 60
agcgc 65

<210> SEQ ID NO 79
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 79
tgctctcacc cccggtgtg 18

<210> SEQ ID NO 80
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 80
cctccagacc ctggtgtg 19

<210> SEQ ID NO 81
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 81
tgctctcacc cccggtgtgc acctccagcc cccagttcag attcagttca ttgactacac 60
ccaggttga ggag 74

<210> SEQ ID NO 82
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 82
tcctgtcggc tacgctctt 20

<210> SEQ ID NO 83
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 83
cgtctagttg gcacagttct g 21
<210> SEQ ID NO: 91
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 91
cggtggacca cacagagtta a  21

<210> SEQ ID NO: 92
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 92
ggtctgctc ttcctgtgc  19

<210> SEQ ID NO: 93
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 93
cggtggacca cacagagtta aoccgaggact ttggagaqca ctgcaqagac atggaaqagg  60
cgagcc  66

gcgaaqgtc cctcagaca  19

<210> SEQ ID NO: 94
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 94
gcgaaqgtc cctcagaca  19
tctaatcttc cggagtttct ca  23

<210> SEQ ID NO: 95
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 95
gcgaaqgtc cctcagaca cccogattga aaagaaccga gaggctctga gaaacctcgg  60
gaaacttaga  70

gcgaaqgtc cctcagaca cccogattga aaagaaccga gaggctctga gaaacctcgg  60
gaaacttaga  70

gcgaaqgtc cctcagaca cccogattga aaagaaccga gaggctctga gaaacctcgg  60
gaaacttaga  70

gcgaaqgtc cctcagaca cccogattga aaagaaccga gaggctctga gaaacctcgg  60
gaaacttaga  70

cagagttgt ggcctgtcact at  22
<210> SEQ ID NO 98
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 98

gggastgttg tagcccasqa

<210> SEQ ID NO 99
<211> LENGTH: 71
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 99

ccaggtttgt gcgtgtaact atctctcatg ccagcaactgc caaagatctct gttctgggct
accacaattcc c

<210> SEQ ID NO 100
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 100

ttgctataatc taagttgttc tocaaga

<210> SEQ ID NO 101
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 101

gtggatggtc agctctgtg t ag

<210> SEQ ID NO 102
<211> LENGTH: 73
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 102

ttgctataact taagttgttc tocaagacccc caaactgtgc tc cacatcctct gctacagtga
gctgacttc cac

<210> SEQ ID NO 103
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 103

aggaogcaag gggttttg

<210> SEQ ID NO 104
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 104

gatgctcggc cagttcttac t

<210> SEQ ID NO 105
<211> LENGTH: 87
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 105

aggacgcag gagggttttg tcaagtgcga ctcgagactg taggcactgc catggccccct

60
gtgcctcga aggactcggc ggcacatc

87

<210> SEQ ID NO 106
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 106

catatgacg ccaagagtgt gaca

24

<210> SEQ ID NO 107
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 107

ccacagattt cttacgtgga atgg

24

<210> SEQ ID NO 108
<211> LENGTH: 82
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 108

catatgacg ccaagagtgt gacaagccac gttggacagcc tyccactcct cacagcctcc

60
atttcatga agasactcgtg gg

82

<210> SEQ ID NO 109
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 109

tgctgactga cttccagac

20

<210> SEQ ID NO 110
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 110

attggacag cttgatca

19

<210> SEQ ID NO 111
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 111

tgctgactga cttccagac cacotgggoa gttgcocaaa gttgtatoca agotgttcca

60
at

62

<210> SEQ ID NO 112
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 112

gatctaagat ggcccagtgtc gaa

<210> SEQ ID NO: 113
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 113

ttgcattcg ttttcctctc ttctg

<210> SEQ ID NO: 114
<211> LENGTH: 82
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 114

gatctaagat ggcccagtgtc gaaocggaa ccaccctac tcctaatctc ccgactacag

SEQUENCE: 115

aagagggaa aacggaatct aa

<210> SEQ ID NO: 115
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 116

cctctgcaaa gttgtcctaat

<210> SEQ ID NO: 116
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 117

cgggtgtgaga aatgcaggtaa gggggtgtggc gaggtgtgct atgtctgggg cattggagacac

SEQUENCE: 118

ttgccaggg

<210> SEQ ID NO: 118
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

SEQUENCE: 119

cgggtatgta atgcggaaata cac

<210> SEQ ID NO: 119
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 119

gaactgagac ccactgaaga aagg

<210> SEQ ID NO 120
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 120
cggittatgtc atgccagata cacacctcag aagttcctcgcgg aagggacccc

<210> SEQ ID NO 121
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 121
cgtggtgcc ccttatagcc

<210> SEQ ID NO 122
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 122
ggtcatgtgg gcatgttagg

<210> SEQ ID NO 123
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 123
cgctggtgcc ccttatagcc tcgtcttgga gatgctgagc gcocaccgcc tatactggcc caactagc

<210> SEQ ID NO 124
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 124
tgctccatcg caagttatca

<210> SEQ ID NO 125
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 125
tgtctcagcg atcttctgctc ac

<210> SEQ ID NO 126
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 126
tgggccatctg caacttatcca catctgtatg cggaacactca aaagagtcac tgggtgaag 60
cagatcgact agasca 76

<210> SEQ ID NO 127
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 127
catccatgc acatc tctggta togt 24

<210> SEQ ID NO 128
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 128
cagttctctg ggtgccagtga 21

<210> SEQ ID NO 129
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 129
catccatgc acatc tctggta gggactcatgcccc acatcactg 60
caccacagac actg 74

<210> SEQ ID NO 130
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 130
caaagagact catctggtgc 23

<210> SEQ ID NO 131
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 131
gagtcagact ggtctatcca cagatg 26

<210> SEQ ID NO 132
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 132
caaagagact catctggtgc ttcggactacacc acactgaa cttgcacccc atctgtgaat 60
asgctcttct gactc 75

<210> SEQ ID NO 133
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 133
ccatotgcat ccatcttgtt

<210> SEQ ID NO 134
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 134
ggccaccag gttatatctg

<210> SEQ ID NO 135
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 135
ccatotgcat ccatcttgtt tgtgtocccc acccttgaga agtgctctag ataataccct

<210> SEQ ID NO 136
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 136
ggtgccc cgtgatct ctgacacagt ac

<210> SEQ ID NO 137
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 137
tcaggaacg ccaagctcg a

<210> SEQ ID NO 138
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 138
cgaasagatg ctgacacagt cacaatccaa ctcagcagaa gggaggag gaagttcacttg
tggatctccc tga

<210> SEQ ID NO 139
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 139
gagacocctg tgcctccagaa

<210> SEQ ID NO 140
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 140
gttctagtc agaaagag atc
<210> SEQ ID NO 141
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 141

gagacocctgc tgtcccsagaa ccagggaggg acagcttcca tgtgggaga ccagatctcc
   60

ttcgcgtacc acaccc

<210> SEQ ID NO 142
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 142

ccacactcag tggcagctca
   20

<210> SEQ ID NO 143
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 143

ccacaggtg gttacgtctt
   20

<210> SEQ ID NO 144
<211> LENGTH: 73
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 144

ccacactcag agccagctca casgtttgag aaaacagccc gttacttga gcaagcgtga
   60
taccacocgg acgg

<210> SEQ ID NO 145
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 145

catcaaatgc cagccctgga gttc

<210> SEQ ID NO 146
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 146
	tacctgtagg ttttaacccc gatacg

<210> SEQ ID NO 147
<211> LENGTH: 85
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 147

catcaaatgc cagccctgga gttcagatg accacagaa ccagctttt tgtcctcag
   60
catcggsggt aaagacctac actgaa

<210> SEQ ID NO 148
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 148

ttcgcgtacc acaccc
<210> SEQ ID NO: 148
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 148
tccaggtgt taggaactgt gaga

<210> SEQ ID NO: 149
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 149
gcgtgtctgc gtacgagctg tt

<210> SEQ ID NO: 150
<211> LENGTH: 73
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 150
tccaggtgt taggaactgt gaga tgaa gaaa cggtgtctgc gtagtagctg gcggaggagaa gggcatgaaa ccaagcagctg gaacagctac
tacgacaca cgc

<210> SEQ ID NO: 151
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 151
aacagctgct acctcagact cas

<210> SEQ ID NO: 152
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 152
ggatttccat cttgctoacc tc

<210> SEQ ID NO: 153
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 153
aacagctgct acctcagact caaggagctg gtgccacga tcocccagaa caagaaggtg agcagaggg asatcc

<210> SEQ ID NO: 154
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 154
tccaggtgt tgcctcaagg a
<210> SEQ ID NO 155
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 155
cggacagagc gagctgactt 20

<210> SEQ ID NO 156
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 156
tcggagatgt tgtataaag aggttgaga tgtattggcc acccttcaag cctgcaaagt 60
cagtcgtctc tgtcgg 76

<210> SEQ ID NO 157
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 157
aagcacgagg tgtctga 17

<210> SEQ ID NO 158
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 158
tgcotttcct tgtgtgtat tato 24

<210> SEQ ID NO 159
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 159
aagcacgagg tgtctgtacc casgttccag cccctctcatt ccaagataat cattctcag 60
aaaggcga 68

<210> SEQ ID NO 160
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 160
cctattcacc tgtgtaacag ga 22

<210> SEQ ID NO 161
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 161
cggacctct aagtttaagc a 21

<210> SEQ ID NO 162
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 162
ccattcaccc tggtaacag gaccoccaagg accctgctcc cgcgaagtgc cttaacctag  60
eaatccggy  68

<210> SEQ ID NO 163
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 163
cgcaaggac ccacatgtct  20

<210> SEQ ID NO 164
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 164
ggtaattg tggacatctc aaga  24

<210> SEQ ID NO 165
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 165
cgcaaggac ccacatgtct gcccacactc ggcggcgaga catgcttgag ggaacagctct  60
tgaggatgtc aacaattac c  81

<210> SEQ ID NO 166
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 166
casacgtga catgtacgtg cta  23

<210> SEQ ID NO 167
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 167
gctgggtgc gcactcct  18

<210> SEQ ID NO 168
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 168
casacgtga catgtacgtg ctaagaattt gcccacactc ggcggcgaga atcactcaca ggtgatatgg  60
cgagttggag aaagttgagc cacaagac  88

<210> SEQ ID NO 169
<211> LENGTH: 25
-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 169

gagcgaacgt cttatggctt aatta

<210> SEQ ID NO 170
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 170

ggcaotcgcgt tgcagcat

<210> SEQ ID NO 171
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 171

gagcgaacct cttatggctt aattaagttca gatgcggcca tgaactgtcgc tgaataagatg

cctcaagcc ga tgcct

<210> SEQ ID NO 172
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 172

gtcccggggt gtcggttt

<210> SEQ ID NO 173
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 173

gatcaotca gatgtctgac aattg

<210> SEQ ID NO 174
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 174

gtcccggggt gtcggttttt gcgaactgtc tgggtgatca aatgagatgtt

tcaagacgt tcagcgatgct

<210> SEQ ID NO 175
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 175

cagagatggt ttcgggacca a

<210> SEQ ID NO 176
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
actctgacacc agctcaatctgt tg

cgagattgg ttcttacaga agacagagga actgaaccgc gaggtgccca ccacagctga
gcttgtgcag agt

tcaagtgagga aggsattgga

tgccatacag agagaaaca
tcaatgagga aggsattgga cccagcaaca ttttgtttgt ccaoagcagt gtttctcttg
gatagtccca

tggactgaaga gaaccagact ccactg

gtgtagccag cgtagacact

tggactgaaga gaaccagact ccactg

gtgtagccag cgtagacact
gtacagaga gaccaagact ccaagtccat tgtgtggtct tgtgtggtca gggctcgag

<210> SEQ ID NO 184
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 184
gacatcacc gttcctgccgg gctcagc

<210> SEQ ID NO 185
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 185
cogcatcgg acacatgttg

<210> SEQ ID NO 186
<211> LENGTH: 86
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 186
gacattcacc gttcctgccg cagctgccct ctgcgccacc cttggttcct tgtggtggtc
gctacagac atgtgtgggt tgtggtggtc

<210> SEQ ID NO 187
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 187
gttttggagg aatgttgttc ttca

<210> SEQ ID NO 188
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 188
ttttttaatc caatgatcgc ttaagg

<210> SEQ ID NO 189
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 189
gttttggagg aatgttgttc ttca tgtggtcag agaatgcagc aaaaagcaga ccctgataaat

<210> SEQ ID NO 190
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 190
gtctgcagc ccctgccgta agacggggtcgt gttttagag a
acgagaacga gggcatctat gt

<210> SEQ ID NO 191
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 191

gcatgtagtg gttccoaatc ac

<210> SEQ ID NO 192
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 192

gagagaacga gggcatctat gtcocaggag tcagaacccg aaaggtgcgc gctgtgattg

<210> SEQ ID NO 193
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 193
tcctcocaact gggasagact a

<210> SEQ ID NO 194
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 194
gaggtgtgct tgtctgtct ca

<210> SEQ ID NO 195
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 195
tcctcocaact gggasagact atctgtgtgc caagaggtgc aaaggtggaca gttcagagt

cctgagacag atcagcaaca accg

<210> SEQ ID NO 196
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 196
ttggtgtgct gcocctgtg

<210> SEQ ID NO 197
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 197
tgggtttctgt ccasacactg g
<210> SEQ ID NO 198
<211> LENGTH: 67
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 198
ttgttgtgt gocctgtgtc ctgctgtccg ctgactctcct ctcgtggcag tgtttggaca 60
gaacca 67

<210> SEQ ID NO 199
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 199
actgaaggg acocctggag 20

<210> SEQ ID NO 200
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 200
taastaacc cc tgtccacaca 20

<210> SEQ ID NO 201
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 201
actgaaggg acocctggag cctaggggca tccggcaggag aqgtgtggtg ccaggyttatt 60
ta 62

<210> SEQ ID NO 202
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 202
agttactaaa aasstccacag aggtcttt 28

<210> SEQ ID NO 203
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 203
gtgggtgtgt gatttggtc a 21

<210> SEQ ID NO 204
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 204
agttactaaa aasstccacag aggtcctttca gttgagcaca aagaccggtg tcagggatt 60
gcacaaatca ctcacgac 79
<210> SEQ ID NO 205
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 205

gggaacctcc aagagatgga

<210> SEQ ID NO 206
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 206

cctccacacc cttcccacac

<210> SEQ ID NO 207
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 207

gggaacctcc aagagatgga ctagtgcttg tgtggttctta ggggtctgga gctttttgga

<210> SEQ ID NO 208
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 208

aggtgatgta aag

<210> SEQ ID NO 209
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 209

tocccgtoscc astgacactg

<210> SEQ ID NO 210
<211> LENGTH: 90
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 210

aggtgctact ccacacgtc tgtgtgacac caagtgcacgc tgtgtccgc caggtgaga

<210> SEQ ID NO 211
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 211

agacgcaatt gcttgacac tccg
<210> SEQ ID NO 212
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 212
ctggccctc aacgctcttc a 21

<210> SEQ ID NO 213
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 213
agacacagt gtgtgagact tcaaatctct gtggggcag gatgttcctg tcttgaaagga 60
ctgtgtagcc ccag 74

<210> SEQ ID NO 214
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 214
gaaatacgct catcgttgat aaaaaa 26

tgacgctg acacaatt 19

<210> SEQ ID NO 215
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 215
tgacgctg acacaatt 19

<210> SEQ ID NO 216
<211> LENGTH: 78
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 216
aagaaatacgct catcgttgat aaaaaaca gaaaaactg cccagcctgt cgctttaga 60
agtagctgca ggtggca 78

<210> SEQ ID NO 217
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 217
gataaaagcc tacatcaga 20

<210> SEQ ID NO 218
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 218
tgaactgtga ggccagagac 20

<210> SEQ ID NO 219
<211> LENGTH: 68

-continued
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 219

```
gatcaacggc tacatccaga agatcaagtc gggagaggag gaccttgagt ctctggcctc
```

```
acagtca
```

60
68

<210> SEQ ID NO 220
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 220

```
gtggagtgcg cctgaagga
```

19

<210> SEQ ID NO 221
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 221

```
ctgcggatcc aggtaagaa
```

20

<210> SEQ ID NO 222
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 222

```
gtggagtgcg cctgaagga aagccagggc gtctacagag aagtctcacag ttcttacccc
```

```
ggatcggcag
```

60
70

<210> SEQ ID NO 223
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 223

```
tggaacctta tgtgtaggaa aqettga
```

27

<210> SEQ ID NO 224
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 224

```
cagtcgagc gcacaccagga ta
```

22

<210> SEQ ID NO 225
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 225

```
tggaacctta tgtgtaggaa agatggagaa tgtctccaca ggcoccaaca ataagccca
```

```
gcctgccgtg tgtccctcgcc agtg
```

84

<210> SEQ ID NO 226
<211> LENGTH: 25
<210> SEQ ID NO 227
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 227

tggctaaatg aagstgacaa tcatgt

<210> SEQ ID NO 228
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 228

tggctaaatg aagstgacaa tctgttga cgaatctact gttaagcttg aaagggacga a ctggtgaa tgaatgt

<210> SEQ ID NO 229
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 229

tctgcaagat tgggagctct cta

<210> SEQ ID NO 230
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 230

goagccttt tctatcaga a

<210> SEQ ID NO 231
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 231
	tctgcaagat tgggagctct ctaatgtgcc atggatgcgg tggagctgta toctgcctt

tctgtagaaa gacctggcc

<210> SEQ ID NO 232
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 232

cagacagctg acgggtcag tact

<210> SEQ ID NO 233
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 233

coggagaa agoagaga 10

<210> SEQ ID NO 234
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 234

acgcagctac acctatgatagtacctgaa actgatggc acctgtgt tg 60
acctacat cttcgatatt tcocgg 86

<210> SEQ ID NO 235
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 235
cctcogctat ggtttctat 19

<210> SEQ ID NO 236
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 236

cocagttcct ggacctgtgg at 22

<210> SEQ ID NO 237
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 237

cctcogctat ggtttctatg aggtgtgatg cctcgcggacctgctcactg acagtttcca 60
gaacotgaa 68

<210> SEQ ID NO 238
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 238

cacaagctg aagcagaca 20

<210> SEQ ID NO 239
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 239
gcgtgcttcc tgtgtttag a 21

<210> SEQ ID NO 240
<211> LENGTH: 85
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 240
-continued-

ccacAACgtg aagcGCagaca aggcccGcGaa gaGgcGctctg gcGaccGcagg ctgagGccGcG 60
ccaggtctaaG accaasGgagG cacGc 85

<210> SEQ ID NO 241
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 241
ccacctcataG atcaasGggtG acaa 24

<210> SEQ ID NO 242
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 242
tcGGaagtg gcGaggtGta atc 23

<210> SEQ ID NO 243
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 243
ccacctcataG atcaasGggtG acaaaGgaGtG cctGGcGcttG tcGGtgGaga Gggattacac 60
cGGCCGcctt GgtGa 75

<210> SEQ ID NO 244
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 244
taGcgGagcc GgtccGataGca 20

<210> SEQ ID NO 245
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 245
taacGcGtGg cGctcGcatt 20

<210> SEQ ID NO 246
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 246
taGcgGagcc GgtccGataGca tcGaaGgGaaG ctcGgGatGag GgtccGatG 60
gGtcGcGcGcG GtGa 74

<210> SEQ ID NO 247
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 247
-continued

cgcaacgtg ttctttctca
<210> Seq ID No: 248
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 248
tgctggttt ctcctctgtgt t

<210> Seq ID No: 249
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 249
cgcaacgtg ttctttctac cctatggtgt ggtgtctgggt ttggctatgc tcacagtgcac
aacagagga gaaacoacoac a

<210> Seq ID No: 250
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 250
tcaagaccat ctccttccc attgt

<210> Seq ID No: 251
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 251
GGATAGGAGA GTACACGAG TATAACT

<210> Seq ID No: 252
<211> LENGTH: 96
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 252
tcaagaccat ctccttccc attgtctcg g acgtgcggtgc cctggctcct cggccgcagc
agcaagttc cgtatctctc gtgtactccc tgtccc

<210> Seq ID No: 253
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 253
gccgagaacgc gcgataata

<210> Seq ID No: 254
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 254
cggtgtcttc ttatcctcatt gat
<210> SEQ ID NO 255
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 255
gccgaaacg cgaatataa toccaaaggg tttgtgcgg taatctgag gataagag  60
cacgy                  65

<210> SEQ ID NO 256
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 256
gctctcccg tgtgcaaat  19

<210> SEQ ID NO 257
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 257
cgtcgatgt ataggagag aagca  25

<210> SEQ ID NO 258
<211> LENGTH: 86
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 258
gcctctcccg tgtgcaaat agggctgctg tttcgacgac acggttcgtg gggtcccctg  60
gttcttatct cttataccsa tcgacgy  86

<210> SEQ ID NO 259
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 259
cagctgatgct cagccattcc taaccttg  27

<210> SEQ ID NO 260
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 260
cagctgactg ttacctgca ctacct  25

<210> SEQ ID NO 261
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 261
cagctgatg cagccattcc taacctttg tgtggagac cattgtcata taccoggttc  60
tagctggttc ggcaagtga tgccgataac agtctagtg  99
<210> SEQ ID NO 262
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 262

cacaggaaca acagcatott tc 22

<210> SEQ ID NO 263
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 263

agataagccc cttggcatcca 20

<210> SEQ ID NO 264
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 264

cacaggaaca acagcatott tcaccaagat gggtaggcacc aactttgcttg ggaactttgat 60
cccaggygt tttctt 75

<210> SEQ ID NO 265
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 265

ggatgtgctca acacccatgc t 21

<210> SEQ ID NO 266
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 266

ggcaattaaca cttttggacg attaa 24

<210> SEQ ID NO 267
<211> LENGTH: 91
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 267

ggatgtgctca acacccatgc tgggcatctg gaccottccta octotggttct ttcgtgctgt 60
tgtagatt tcggtcctaa aaactttaatgc t 91

<210> SEQ ID NO 268
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 268

gcactttggg attcttttcca tttt 24
<210> SEQ ID NO: 269
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 269

gcatgtcag agccctccac tgea 24

<210> SEQ ID NO: 270
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 270

gacctttggg attctttca ttgtgtattct ttgttacagg caccgagaat gtgtattcaga 60
gtggaggtct tcctacatgct 80

<210> SEQ ID NO: 271
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 271

aataccaggg ggasggtgtat 20

<210> SEQ ID NO: 272
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 272

gtacagattt tggccagagga 20

<210> SEQ ID NO: 273
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 273

aataccaggg ggasggtgtat gactttcata tggactttga ctcaagctgtg gctcctcggg 60

caaaaatgtac ac 72

<210> SEQ ID NO: 274
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 274

tgtggacatc tcccctcag a 21

<210> SEQ ID NO: 275
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 275

cctgcccaac cggttctgt 18

<210> SEQ ID NO: 276
<211> LENGTH: 66
tggccgatat tttccttcag accttcctac tggccacacc tcctgctgca caaacogttg

<210> SEQ ID NO 277
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 277

ctttgaacct tgtgcttgc

<210> SEQ ID NO 278
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 278

ccccgacc aagcasatg

<210> SEQ ID NO 279
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 279

ctttgaacct tgtgcttgc  taggtgctgg tcaagaagcc ccaggaattc cattttgttt

gtagcggg

<210> SEQ ID NO 280
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 280

gctcctgtgt gcctttgca

<210> SEQ ID NO 281
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 281

cgtgtatgtgc gcactatag

<210> SEQ ID NO 282
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 282

gcctggtgt gccttttaac atcgccacgt acgcctgtct cacgtatacg attgogoaca

tcacyg

<210> SEQ ID NO 283
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 283

ctgctgtctt gggtgcatatg 20

<210> SEQ ID NO 284
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 284

gcagcctggg accaccttg 18

<210> SEQ ID NO 285
<211> LENGTH: 71
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 285
tctgctgctt ggggtgcatatg gacgcttttg aactctgtctt acctccacca tgcaagttg 60
tccccagccc c

<210> SEQ ID NO 286
<211> LENGTH: 1947
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 286

cccccccc atatggttgt ctatcagc ctagatattc tagacagag ggatatcagat 60

gagagagaaa atgtgtagat attagctctc gtctattctg agggcagact aagtggctca 120
gcgcggcttt ttacctctct gctgtcagc cttggcaccga gctctcactcAccaactcccc 180

agtagcttgatta gtagagactat gttgtccattt atatgtttttt gatgtggag gaaacagtc 240
ataataactag atgatgagaa gaagtagagc cagtgatgg atgagagac taaaagagc 300
actcaagtgg aagttttttt ttttttttaga ttagagttttt gaagtagatg taaaataattt 360
tgtcaagcct tgtgattttt tttcttgacta aagttttttt tttcttgcttt cttgtgacta 420

atatatag ttttttttaga gttatatata atacctgttc aaggtgtgta aaaagttaac 480
ttttattcatt cgtgagttag atacaatata atataaatcc cttccagagt gttgacttgt 540

ttataattatgt ttatatgata gttatctgata tatttggacgt ctaatttgga aaaaatccg 600
taggtagca gacaaatggt tgttctgata ggattttttt atgttttttt tttcttgactg 660

gtgtagattt cttcctgacta ctggctgctag attgctgatg ggtttttttg 720

dcggtttt tttgagagc ctaaagatc gactacatag ttaaagagtt gggagtatc 780

cacctcgc ccctgtagaaa atgtgtcttg gctcataaaa tggagctcag ctagggca 840

dctgtagaca gacagactga atctgatac aagaataatc caataatcag aatgtagcct 900

aatcagtt actgttagtt ttacgtatgt ttagatattt tggggttgga aaggtagag 960

gagagagag aaggttcttt gtaagacct gctgtgctg atctagcttg ctaaagactt 1020

tacattttg aaggttacat cttgcaagaa toctataaag aatttagttac tgaatttac 1080

tttgaaatct agtagcactg gtcctcagct gactgttggtg aagggctatatgcttgctattt 1140

eyytgattag ggtgagatctt ggtggtaag gaaagaaaat tttctgttta aactccagaa 1200
agactttcag gcctgacctc ataggaggct atccatttta tcccctgagg tttcttcaac 1260
cctgctgttg gaggtggcgt tttcgatgct tcocagaggg attttttttt tttggggtgt 1320
aggggttgg cttcttcaat cactctttct ctaaaaaag gagaatggtg aatgtagcaoc 1380
taaaagctgc atagccgcat gaaagttgta tattygctc atccacagt accattttttg 1440
tgctgtatt cactccactg acatcttacc agtaacgcgt caggattgaat aagaasaaagc 1500
aaagtcgaca cccggagctc tctctgctgt ggtacaggg ttcagcttca caagttgaca 1560
gatcaccaca ccgccaggg agttccattc gacattttca ggtgctgtt gcgtgcgccg 1620
cattgtctgct cggagattgc aagttttaat tccctgtaaa gasattttca tttgctcaat 1680
cocagctgg gtcggagcg tccaatacgag aagttcctggt tctgcaagc 1740
actctggctt tgtcttaccc ttttactcct cttactctgt aatggttatt 1800
gttgaagg cccagtttgc tataaatgtc tctggcttcc atatatttttt ttaagctctc 1860
ccacacactg atgctgacat ggatttaatg ctaaactttt tgaataattg tcaagtttagt 1920
aaatcctgtg ataaaagtgttt tgaagttgt 1947

<210> SEQ ID NO: 287
<211> LENGTH: 1311
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 287
tagctgggag gggtgtaga aogocogcgt agocttcogc agoattggcg tcaaoaagtg 60
cgggctggcc ccggccggag gggggccggg gggcccggag cgggtcggttt 120
cocggcgcg ccggcgggcc aogcggcgc gcggagaacc gtcggccgcg 180
tggcttcggc ccggccgctc cggcgcgcgg gcccggggcg ccgcgcggcc 240
cgggctgagc gcggcgcgag cgccggccgg gcccggcgac gcggcggagc 300
gggtgagcgc ccggcgcggc ccggcgcggc ccggcgcgac gcggcgggcgc 360
tgggatcgg cggcgctgcc gggattggt gcggcgtcgg gcggcgcgtg 420
gggtgagcgc ccggcgcggc ccggcgcggc ccggcgcgag gcggcgcgac 480
tgggatcgg cggcgctgcc gggattggt gcggcgtcgg gcggcgcgtg 540
ttcggccccag cggccgcttc ccggggctcc cgggggctcc ggcttcggcct 600
aggtcttcatgg gggctgccag tcttttccag aatctatatt taaggggaa tctctgagaag 660
aaatcacaac acgcttggca gcaccttgaa gaaatggctt ccggctgggtc atgtaatttg 720
ccgggggaga cgggggagtat ggaatctgatc gctgtggcag ccgggagcgg 780
tgggattttc gctgtggcag gcgcagtag aatgtttcag aacagtattg aagggagag 840
aggtttttct gcggcgggt gcggcgggt ccggagttct gagaattct gcggcgggt 900
agggccactgc aaggggtcctg atgaagactg tgggagatc ggagacactg atcggcgc 960
aasattcag cggcagttgc ccgggaggg aagggctgtg ggggaggtt gggggagcttc 1020
tgtagggcgc tggcagctgt gccgggagat gctggcgtgag ccgggagttc 1080
cagccttcgg cccgggagat gggtgtaga aagggggtgt gggggagcttc 1140
gggccacactgc ctgctggatgt ccgggggtct ctttttccag agggtgtgg 1200
ggggagcttc cgggggagat gggtgtaga aagggggtgt gggggagcttc 1260
ctttgtgatt tgaagtaacg ttctttcttg tttttccaa aaaaaaaaaa

<210> SEQ ID NO 288
<211> LENGTH: 582
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 288

atgcccaccg ccagccagaa gggcccctcc ccagccccgg tagaggggct ggcggcggct

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

ggcgagccc ccggcgagcc cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

<210> SEQ ID NO 289
<211> LENGTH: 6030
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 289

gttggccgcc gttacttttc ctctggggaa tggagccac gcccggagga caggtaaga

taccaggag aatggctagc aatgactacca tattaacgct tggagccagag gactagagct

gagtgggga ggttgggggc cccagcccc gggggccag cccaggccgg gacttctcttcc

ctcgagccgc ggcacacgoc cccatccagc cggccgctcg ccggtggctg

gccggtaag cccagccgag cggccgagct cggcgtgtgc cagccgctcg ccggtggctg

aacctgtgcg cctctgtcgc cctctgtcgc cctctgtcgc cctctgtcgc cctctgtcgc

tggccagctg tggagggcag ttttgagagag cggagtgctc gtttggtctt gtttggtctt

ctttggttc gggcggtcag tggaggtgca gggagtccag ggagttgtc ggggctctgt

ggacaaacat ggcctggta agtatgcttg agatgctcag gggagctgct ggggctctgt

ggagctgct cttgtggag cttgtggag cttgtggag cttgtggag cttgtggag cttgtggag

tgataggtc gttggagtc cttttttttt tttttttttt tttttttttt tttttttttt

tgagatccg ccagccgagc aagcoagct cgtcgagct ccagccgagc aagcoagct cgtcgagct

tgctctctct tggctctctct tggctctctct tggctctctct tggctctctct tggctctctct

tggctctctct tggctctctct tggctctctct tggctctctct tggctctctct tggctctctct

tggctctctct tggctctctct tggctctctct tggctctctct tggctctctct tggctctctct
ttcaacaggg attcacaagc tatttgaaas atgtatatat attaagaggt caaaggggct 3480
aatttgtgctc ttagcttggg gtttggatgc cttcagttaa taacgaaa
3540
ttgccagcag cttcttgcccc cagaaatgta caggatgctg gotcactgct cttcagaact 3600
agtttagttg cgtttcctct tatttgtaaa acaatgtaga acgcataagc tgtatatgaa 3660
agctctcact cattttttcct tctctctct tttttttttt attatatctct attatatgctg 3720
gttgtggcag cagagaaaac cccctatctttt gttattgaa gggattcaca tcggctactt 3780
aatgttcttt ttcgatttta cttctctttt cttctcgttc cacttgccag 3840
ggtcacagtct atatagactc tttatctcattt tcaaatggtg gcataaggggt ctaaaaag 3900
cccaaagaaag acagaaaaaaatctgagcgct otcgagcccc ccaacatgtc gcctgcaca 3960
atctcgcagc aacagggcga ctatgagctc gcggcagctt tttcaacgtt tcggcttccc 4020
cgagagtttg gaaggggagca aaacttgctg taggtttaa agttgatata aagtccaca 4080
atacaggaaga acocctcaact tagtcctacaaa cagagccaaa cctcctcctc cccagcagc 4140
tacocctcttt ttggtcgactc tgggcttctac tttttttttt ttatattctat 4200
tgatctatcc ctataagttt acgtctgccg cagctggaaa aattaggaag tgtttaaaca 4260
tcgagagagcc tttatctctat gtaaatattttt cccgcaacat ccaaccatgt 4320
tctaggttctc atctctctctc tattaggggt gacagatgct cttctctctg ccaccttctg 4380
atgtaactctt caaatgatgct ctctcccctt atatattttt atatattttt atatatattt 4440
acetactatt taaaagcttt gcccttcttcata ctaccttacttattca 4500
acactctct ttcagcgaagaa aaaaaacagaa cgtctgagct caacgacctt ccagagctac 4560
tagggagtcc gttgaaattttc tatttctgac tttctcttgt gttcttttct cagocccag 4620
aaatggttgtt acocccctttt taaatctcaga aaaacttcat cttcagagttg ctaaagtttc 4680
cactcagata tttaatttta ctcaattcgg aotcttcaggt gtttttttttt ttaaaaattt 4740
atggagacaa ggcacattgcttg agggggttt gcgcaggttgga acaaaattttta aataataaac 4800
attacagaatt tggtaggtgg aagttgtaag tttcagagaa aacccaatct aaggtataga 4860
agggccgctag tggggtcagc tgtagagcct ccctgatacag tcctttgttgg caaatgcgaa 4920
acactcggatat gtttattgtg cacccctttg aagctctagc tcgtctctag aaaaattttaa 4980
acctgccttt gggcagcttc tttcagttgg tggatttttc aatattttatttattttgg 5040
aactaataaa ttgcaggttt taaactacag gtaatttaa gaaactccctg tgtgttgatag 5100
acacactcatg ttctaatctcttt tattataagtt caataccagaa aaaaatttttaa aaaaatttaa 5160
caaatgtagaat ctaagatgtac cgcagatcctt cggcttcaco cagtacatna 5220
tgatgtgcctg ttcgagtcgct caaacagctg caagctctgg ctaagctaggct 5280
ggaatgaaag aaaaaagatt tggacttaga tyccctctctg gcaaggtgtttg cagagctgtc 5340
cagaaaaaa atctcagaca aacatcctat ccaacacaaag gtggtttttc aatcagacgt 5400
gagccgagaag tggagagccattggag gcgtgttgcgg cagctctcagc acgacaaattt 5460
tgagcataat ctttatattg agactgtatag atgacactag aatatacctat caaatacaac 5520
ttttgaaaata tgtcggctttg gcccctagaa tagtccttcct tattatgtct ctacttcacaa 5580
ggattagattt gcacagttgggtg ttggctatgct tttttttttt tttttttttt 5640
gttattttg tttaatcttttatcttttttc acacaaggtgc gtttattttg ttgatttttc 5700
<210> SEQ ID NO 290
<211> LENGTH: 10867
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 290

```
ggggttatttt tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued aag caggca tatctggaac ttcto cagtg gottctt cat titcagggitat caaaaagttct 1740 atatto agaa taagagaatc acctaaagag ...

OO acaaaact ga atgtttctac toaa.gctotg caaaaagctd togaaactgtt tagtgatatt 396 O
gagaattata gttggaacag ttcgctcagag gtcacccca taagttttagg ttcaagtaaa 4020
tggcagtatt cttggttctt aagtgttaag ataaaccttc ataatgtaaa aacttgaag 4080
gaaaaataa ataatacgca actgataatta caaaaataaa ttgaaagtac tactggcaat 4140
tttgttgaag aataactgca aatacttacja aagaaactctg aataatgcaaa actaacactt 4200
actgagctca gtagaaccct ttcatatcttt gaatggtttct ggagagttctt cagaaaaat 4260
gatacttgct ttcactttaa gatggaacag ggtctgtcat ttcactgata gcaacacata 4320
tgtccttaat ttcgctcagaa gttttagaag gggcagcaacc ctcagattaa agaaagttttg 4380
tccgatattata ctttttgcggag aagtggaacag gtttgcagag cattcttgtt ctaacttca 4440
ataaaaaacc agtttaactag tcaaacaaaag gcgaaacctta aaaaagattt tggagtcttct 4500
gataattattt ttcgacgtcag aatgtagaa gtagtaagtc ttcgcaagaa gccttttaaat 4560
aaatatgaaa attttatctg tccagaaaca gcaagacccttgc aacttttttt ctttaactct 4620
gaataccatt ctgacataag aagaaacaaag atggcacttca taagtttaga gaaacacagc 4680
atgtaaacccaaactactg gaaagaaatgt ctccagagtgg ttacgagaa tcaactagtgt 4740
acctgtcagg gacaacccga actgtgatgaa aagatcgaag aactacctct gttggttttt 4800
catacagctg ggcggaaamat aatataacat ctaaaacctg aaaaagagag 4860
catttttgtt gaaaaagaca agtttactgt gnaaaccocca gttttgacca tttggtggca 4920
aagcctattca atggctgtta gcttcttattt taagttctga gacactctag 4980
ataacagctg ccccaaggtg taaagaaatg cagacattct tcaataataa taaaacacct 5040
gttctatag aagcattggt gcaccttaagg cttctattgctaatattaag tagaaacacat 5100
gaaaaatcca aacaaactaa aagarcttttt ttcgaagtaa aagtacatga aaatgtgaa 5160
aagaaagcog aamaagaac gcacaccatgcc tcaacttttga tcaacataac atgctccctta ttcaagccct 5220
gaaaaatccag cttcatgatttt ttaccaaatggt taagtggaa aaactctccgt gacgcaagct 5280
tcaactattct aacaaaaaaat atggctgaggg aaggaatat gtaggctca accagaagaa 5340
ataaatattgg cagattatgg ggaagataaatt ttagtgaaaa ataatttccag cagatcata 5400
gttgaaagtt cccaaactttc ttctttcaca aacacageta cttatatcaag tacaagaga 5460
agtctaaac gatttcaactt ccaatttaga aggtacataa acttctggtt caaatttattaat 5520
gaaaaatccg acattgtttcttattgaa gttcagataa atgatcaag tcaatctgctttaa 5580
actatttttt cccaaacttct tccaaactta aaagagctgaa atggcatacc aacaaactgta 5640
aatgaagaatt ttcgctcttg gcaagtcctgg aactagttctc caaccctgcaaa aataaaaaat 5700
gcagcataa aactgtgctgct atctaatgaat aataattttg aggtagggcc acctgctatt 5760
aggtatgcaac gttgctaaat ctttggtttct ctaacgtgaa caatattttca atggttgaca 5820
ataatttccg acagtcttct caaagtaaat aagggaaaca acgaagataa ataacaactt 5880
tccgtcaagaa aaattttgag atgtctcttg acttcacag gattcatttg cggatattttct 5940
cataatctcct tagattataa gtaagttgac acggattcaca ataggtttct tggctcattt 6000
cgagatgga aatattaaaca aacaaactttaa aatagtcgtg gattgtgacaag ggatttctaa 6060
ataatttactt tggagttatg ctttgtgaa cttagatatt gtaacttag ttagggaag 6120
cctctacttg cagcattccac tcaaatctct gtgggtttctt ttcacaggat cagttggaca 6180
tttgttcagg ttagcagata ttcctacaaaccgtaagcag aagtgttctc tgaatatagaa 6240
-continued

gatagttcagaagagcgtctttaaccagagaattcccattgtttaaaaactaggaaccatttaccaccag 6300
cctcgaagagacacgactgatacgctacggactcagacatgagaaag 6360
tttcaatagagcttatgtggatatcctttctctctggactcactgaaagttcccaaggaag 6420
cagtctccgatttttacatctcttcttcttcataacagagaggtcttttgaatcagag 6480
ttacaggaattgcctcatctctgtctcttcttttttcttcttcatttttttctttttttcttcttc 6540
aatctctctgtgttgtaaagaagccagctctgcttaaatgcttgaattttttaagacagaagtttcatgtcataagtaaagttttttaaagtgaagtttcatgtcataagtaaagttttttaaagtgaagttt 6600
acgtgcagttacagatgtaataactcatttcacttctcattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ttgacacctc ttgacgcc cc agaatctctt atggaanaa tttctcgtca cagtaacttg 8580
cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8640
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8700
cgcaccctgactgccat cgcagtctgcgctctgctggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8760
agagagacaaaagaaatcg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8820
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8880
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 8940
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9000
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9060
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9120
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9180
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9240
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9300
atagtgtag ccagcagcag caaatctcag tattcataac taccggttcag ataggaatt 9360
atttcgagc tatttcgagc aggcggcccc cattctccagcc gcaaatcttcctttaaattc gataacaattgctgcatccttgatcgcagagcctagcctgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9420
ctttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9480
acagacgtcg cccctctcgctctgccgcgctggccagcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9540
agagagacaaaagaaatcg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9600
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9660
gattctaattgccagcagagcctagcctgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9720
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 9780
cagacggtag cacctgccgagctgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9840
gtgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9900
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 9960
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10020
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10080
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10140
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10200
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10260
gtgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10320
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10380
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10440
tttgctgctag cttgagggag cctgctgcgtggcttg cttgcagcttgatca atcctgctcgttgactcttt cacaaagtca 10500
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10560
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10620
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10680
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10740
agctgtgagc aatccggttctg atcctgctcgttgactcttt cacaaagtca 10800
agccaggag ttcagaca cggctgggca aactagggaca ccccccatttc taccagaa 10860
aasaaaaggg ggagaaagaa atctttttaa ctctggagtc tacaactac aattattatt 10920
ttcacaactca caaaatgtgc atccaaactc aacactgaga aataatcttg ctttcaatt 10980
gaccta 10987

<210> SEQ ID NO 291
<211> LENGTH: 1552
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 292

ggcctctggc agcctgggtc ggatggccct gcaactcttc ggtgcctggt cggccctctg 60
agcccctggt gcttcatgtct gggctagctg cgggctgct gcttcatgtct ggaacccctg 120
cgccctcttg gtcgctgtct ctctccctct gtcgctgtct ggtgcctggt cggccctctg 180
tctctccctct cggctgctct cggctgctct cggctgctct cggctgctct cggctgctct 240
atgctgggatt ggtgcctggt cggccctctg ggcctctggc agcctgggtc ggatggccct 300
ggtgcctggt cggccctctg ggcctctggc agcctgggtc ggatggccct gcaactcttc 360
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 420
agcctgggtc ggatggccct gcaactcttc ggtgcctggt cggccctctg ggcctctggc 480
cgccctcttg gtcgctgtct ctctccctct gtcgctgtct ggtgcctggt cggccctctg 540
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 600
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 660
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 720
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 780
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 840
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 900
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 960
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1020
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1080
tctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1140
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1200
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1260
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1320
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1380
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1440
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1500
ctctccctct cggctgctct cggccctctg ggcctctggc agcctgggtc ggatggccct 1552
asgagcagggc cantaagggg gggcagcagtgc gggggttataa tctgaggcota gcgtggtcct 
  60
tctcgttgct ctgcygccgta aacgcctgtgg gttctgtctg gttgtcgctc ctggygyttg 
  120
tccgcctcg ctggtctgtg ccaccctgct gcggcctgct caggtgaaaga ggaagactcg 
  180
tgcctcggtg tccctcgggtg agtcctcggg tgtgggaacgc ctgagccacaagcactag 
  240
gcggccgaga aggctgcctcc tacggcgttccgtgcta actcgccgactggac 
  300
gagccacgct tgggggacact tggtaaccaca gtcgtaaacactgcctgaggc cagacatct 
  360
tggaagagg aaccaaaccc ttcagctactt gggaaaggcttgagttaaaaa actaacaaaaa 
  420
cctcttggata agtcatcctt gctcgtgccc gtcgcagctgc tctgcctggc ggcagc 
  480
gaaactagcg cagagcctgaa gcssctgaactg ttcgccttcgta ggcagttgag 
  540
gtgagatctg cctcctcag caaatggaac aacctctggat gttgcctgtgggagac 
  600
cgctgcctctctctcatcactttctgtt ctgcagatctt atcgatgatct actgcattttg 
  660
gtgctccaaaa accttccactgtgtaaaaactatgactttctct gatgaacactt 
  720
gagggacgc aacgatcagc ccctaaaacactcctgtgctg ggcagcctggg 
  780
agagccatct taaagcttgc gcctagtcag ctcctcatcttt aatcagggtg cccgtaag 
  840
accctgcct cgttcagctg catattctt ggcggtcttgg aacattttgg gttggg 
  900
aagatgtcgc aagttgctgct gcacatcgc cttctattt gcaatagcata cgggatc 
  960
tacctctcgc aaaatggaca gctggtcttctgtgtggctgc ggcacactac 
 1020
atcagacag tggaaatgaa gcattccaga ggttcttactct tgggtcgtggg cgggctcttc 
 1080
ccctttctgc ctcttcggag agctatctag atgggcgtg gtcctctctg gcaagctctac 
 1140
ctggctcaaat aocctagggg actactattg tggactactg accattgtgc cttcctctct 
 1200
tctcaaatgy aggccagctgg ctctctctct gactcttcgtca gctcttc 
 1260
acacaactct accacactg caagcttcatt actgactgactcag cctcccttgcagttactag 
 1320
ccactgttcct gatctgtgcttg atagctgttctgt gctctgttactg gctct 
 1380
aacaaatgc ctcctacgct gctctgtaag actaatcactc taccctcaga ggatcttcgc 
 1440
tcactctct cttaccttact gcgctgcgtgg caagttgactaatgtgccttc 
 1500
actactattt actacttaat tggaacctctgc tggctctctgt gaaaaaaa aaaaaaaaaaa 
 1560

aaaaaaaaaa aaaaaaaaaa

1578
-continued

agcggcagcc gcctgacgca cctatccttg ttcgcacggg cattcgaatc tttccttgct 480
ttcgatctc ccctaatcttc tgatgcaggt ctctgcatgg atctcccacag cctatatgac 540
cccccacatgt ggagcagaca gtttggacag gatccatcag cagcccaacgc gatcattcga 600
aaccagcagt ttgctttcag cggcttccag ttctgctcgg tgaagactgtc gggccacaccg 660
cctctgttcc ggcacatcag caaactccac ggcggcagcc gcggaggagaa gcagagaggg 720
ggcagtgag ctcgcagcag ctcgtgggga gacacagagaa agtcgctctt ctcaggaacgc 780
ggggcccagcc ccctgcttggag agagaggg ggtcgttggg tgggctctct gcgtatggtag 840
gtcctgcttcc cccgctctgg ggtcttggct tcaagatctct tcaagatctcc gcctaaaacctg 900
aacacactcc tacccctccac gtcggtgggca cggcaggggaa gcggccagga agcctttgcc 960
cagagacccgtgcggccccc cggacggatgt cttcctccag tggagacgga gatggaagttg 1020
gagggactttc ggcctggtgg cctacaattc ttaactctgac gcgcctggaa ggggctactc 1080
gagggactttc ggcctggtgg cctacaattc ttaactctgac gcgcctggaa ggggctactc 1140
cccggcagct tgggagacgc actsattgagc cccgctctcg aagcccttgga aagcagaggg 1200
gaaagacgct cttcatctca cagcagtcgc caggggcctc tccgctctcc gtcctgtgcc 1260
tgcgctgtag ccctgccgct ccccaagggg gttgagcagc gcggccacag gcggccacagc 1320
gttggagaatt cggcagacgc ggcaggtgcc cttcctggag cggcagacgc ggcaggtgcc 1380
cccacaagcgc ggcgttttcct cctaaagctca ggtgtttcag atgactgtca gaaactcttg 1440
gacaagtcacc gcggagacgct gattggagct tacatctaggg cttctccgct atgactcaagta 1500
gcgagaca cccgctttcc caagcctctc tcctggaacaa cttggtggcg cctacagcctg 1560
ggcagtttc gggacagatg cggataattt gtgatgttag acggagacag ccccttgcga 1620
ttggagggc gcgacagacc agctgccttg acctgtgccc tggacggcga cggcagacgc 1680
ttcacactga agacgcccat gcggcagctgt agcgtggaca agagagctat cctatatttc 1740
cactggaat ttcctcttcag gcgtgggccc gcgaatgtcc gttctcagac gcggccacagc 1800
cgtgtgtgca aagcatccgc cagctcatac tacatcggta ctgtcatctct gaaaggccg 1860	taacaggagt tcttcctccc gcacccgcaac ttcgtgtgcaacc cccagctactt ccggcccagt 1920
aaccattgct cctttcagag ttagatcagc aacccctgcac tccagacgtg cagctggggt 1980
ccggccacg ccctggcaggg gctggagcgc ggcctggtgg cccgctctccg cctgtccttg 2040
agttgctgact cctcccttggc ctttcgcggg gtcagcagc gtctggctg gcgtactcgtt 2100
gctgagggc cggctggtgg cccgctctccg cctgtccttg ggtcagctgt ccggccacg 2160
cctggtggc cgcgcagatg ttcctgcggt cctggtgggc ctttccatcag tttgctgtgccc 2220
ccccaccgct gcagggccag gttggagct ctctgctgtg cttatccttg tccagttctga cagacatgta 2280
ggcagatattt cttggtgcttc aacgcggtct tttccctggt tggagttcag ggtcagctgtt 2340
tccgagcttc ttttgcacgc cttccctggt tggagttcag ggtcagctgtt 2400
cctgtcgcct gcagggccag gttggagct ctctgctgtg cttatccttg tccagttctga cagacatgta 2460
ccagatcctcc tttcctggtt ctttccatcag ttttccatcag tttgctgtgccc 2520
tgcctcttc ttttccatcag ttttccatcag ttttccatcag ttttccatcag 2580
agctctaccttc cttttccctc TTCCTGCTGT GCCGCCACGG CTTGACCAAG 2640
cacccagag cagccttcctga gacgtttgag ctttccttggt cttcctcagg ttaactctcg 2700
-continued

ggccacagtt tttgtgcccc aagaaggagat tttattatcc tttggggttc ccagggcaag 2760
gytttaagcc gccctaatcag gctctgtggg aagccacgcc ccctctgtgt gaacccctgag 2820
gccagctgcc cccagacggct ctgggctgctt gatggagag gatggaaggt tggatggatg 2880
ggcgcagctgc cctggacgag cggccgcttg tggcagcata cccagccaaa ccaggtgga 2940
gcggtgtggt tgcagtagcag cctgcagcag gaaatatag tggctcttct tggctggcaca 3000
aatatttaac acctgggctt tggagctatt cagagaaaaa tgcacaagaa gcagcataac 3060
cacggagc ccacccctct gcattctaat ctcagagtt ttggcccggt gccagcggct tggctttgag 3120
gccctgtgct gtcctctgtg aggccccttg gttcctttag ctacgagcag cgtcagcaca 3180

aaaaaaaaaa aaaaaaa 3195

<210> SEQ ID NO: 294
<211> LENGTH: 3737
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 294
ggcgcgcgc cccacctcctc cccgggcccc cccagccccg cccagctcgg cccagctcgg 60
cgcgcagcgc ctcagcctcgc atggggtctg cggccgcccc cccgcctcgcc gcgggtcgtg 120
cgggctctgct gctgtgccttc cctgcgtgct gctgtgccttc gcgggccttg ggcgtgtctg 180
ccgggccttg ccaccgctgct ccaggtgctg ctgctctctgc ccaggtgctg 240
tagctgcga gtaggcctgc ccctgctgca aacaacccac ttctacacag tgcctctgca 300
agcctgtgct cccaggggag gcagccgaggt tgacagcagtt cagtgctgtc gaaagatgcc 360
tttgtgggt ttgggtcctc gtaatctggt cacttatggc caattactg tgcaccccttc 420
tttgtgggt ttgtgtgtct ctcacagcct ataatgtgct tgggtgtgac gatgtgtctg 480
agcagcaatgg cgccgtgcag cattctcctg ttcacgctag ggaggtgtat gatggtgctg 540
gcagaggggcc gtttctcctg cggctgcaaat cgcagccctg ctttacccg cggagag 600
gccgtgctg ctacgacattc gacgacggt ctgagacat tcgagatac ccagagaggg 660
gccagctgc cccagctgct ccagagtcgct cttgacggtg cccagagag ccaggagctc 720
tttgacgtc tattgctgctc aaacgtgctg gcggctgctg ttcctctctg gcggctgctg 780
gcaggtggt gcagcagcg ccagctgcag cccagtgcag gagcagcag ccagctgcag 840
agcagaggg cctgcctcag gattgatgag agacagacac ccccatctgtg ggtggtgctg 900
atattttggc gaaagcggct gttctgtggc cttcagacac gaggctgtgc 960
agcagctgc cttcagagag ctcctcctgt ttcctcctgt ccttcctctg gatctcactc 1020
tctcagttgc ttgggaagct tggatctgctt gcgtctctgc ccaggtgctg aagatgac 1080
gctgcagac ttcgtcgctg gttgctgctg cggctgctg aaagggctt 1140
aatattttgc gattggaccg tttggtatgtg cgggctgctg aagatgac 1200
gcgcggctgc cttcgcagac cccagctgcag ctcctcctgt ttcctcctct gcctcctct 1260
tctcagttgc ttcagcagag ctcctcctgt ttcctcctct gcctcctct gcctcctct 1320
gctgcagac ttcgtcgctg cccagctgcag ctcctcctgt ttcctcctct gcctcctct 1380
agcagctgc cttcagagag ctcctcctgt ttcctcctct gcctcctct gcctcctct 1440
ccagctgcag ctcctcctgt ttcagcagag ctcctcctct gcctcctct gcctcctct 1500
gctaccttg gctaccttc ttctcagatg tcacccacat caggacaagt gtaactttta 1560
agcataagta aggcaagtgt agttgaaaa atcgtagcct gtccccggag ggtotcgacg 1620
cagcactacc agacaagacg acgtcagtaa aagagcctt cctgctacgta aacctcactat 1680
gcaagctcgg cagcssagtt cccagggccc tctgcccagcc acgcacccct aaccgatgtg 1740
ttcacctaatt ttcgcgcgtg tagcagcactt acacagaggt ggtacagctct ttcgctgamc 1800
tcgcgtcgt cgttaacagta aacgagaacgc ggctcgcgtaa gcggatcagc acgatcaaga 1860
agcgctccca cagggagcag tttacacccg acgtctctgg catgacactc gcagctggtta 1920
aaaaagctcc cagcacactt gcgacgaggc cagagctccct tagtcgctggc cagccgcatg 1980
cagaaaaacc atgtgcatct ctcagggcttg gcacatatta tagtcggtcag cggagaagct 2040
gacttttagc tccaaatgga aaccttcaaa atcgggagg acaaatgact tgcgacccat 2100
gcaacgagtc acgaatcctt gggccctgta gagccctaga aagtgcgatg aagtcgagta 2160
gtgggaggct cgtgacacct ggaggtatt tctcagatgt ctcccagctt tgcgacccct 2220
gtgcacccgg cagcacccca cccacgccat gcagctgcctc ccaaggtgccg 2280
gcctgccca caacacatcg gcgctgcttt cccccgagga atcgcttcat caggtgct 2340
gttgacgctc aatccttac aacaccaocca tctcacgctgt tccccgctgg caatggtgaa 2400
cataaccagc tggattcagg aaaaaatttt gttgattcctg ccagagaccat actccagctg 2460
caatcctgcc aaaaatcagc tggatatcctg gcagctgcctc ccaaggtgccg 2520
gagattccaa tgggtcagg aaccttcocaa acattccacag caaattaaca acacaaaaag 2580
agtgcaggtc gcaccacgct ccaacccccca aagggcgcat cctgtcgtgct ccaccggtga 2640
ttcocctggc catagagcag cacagtggggc cactatggtg gatgagggaa acoottocat 2700
cccactcttg cagcactatc aacacccatc ctcagagactg ttcgggctc gagctccaccct 2760
ccagctacaa gcagcggttg agtacagtta aagtcacacta aagggcagc gcaagaggtt 2820
tccagggcct atacctgcca tctgagggc cctcgggggga cctctcctgg gacacaatgg 2880
gagagccagc gctctacatt gcaccaagtct gcagagttt gataatcag gtaaagctaa 2940
tcaagctctt gttcgtagtc gccggccccc tccagctcag aacaggtgccg 3000
agctccgaga agtttctcagc atagttcctt gcacctttcctt gttgctggtctt 3060
cttcagaccc ttcacccaga cctgctgtt cgtgctgctt ctctgcagga aacacgctg 3120
agggagcctg cttcctctcg ctctgcagga cagggctcct ttcgctgctt 3180
cctctcactg cagcagcatt aagtttctat tttttctgcct cagcagcatt 3240
gacactttctg tttttttctt cagcagcatt agaattcctt agagctgccg cttgctggtctt 3300
ctcagctctc cttgctggtctt ctctgcagga cagggctcct ttcgctgctt 3360
ttcgctggtctt ctctgcagga cagggctcct ttcgctgctt ctctgcagga cagggctcct 3420
ttcgctggtctt ctctgcagga cagggctcct ttcgctgctt ctctgcagga cagggctcct 3480
ttcgctggtctt ctctgcagga cagggctcct ttcgctgctt ctctgcagga cagggctcct 3540
ttcgctggtctt ctctgcagga cagggctcct ttcgctgctt ctctgcagga cagggctcct 3600
ctcagctctc cttcctctcg ctctgcagga cagggctcct ttcgctgctt ctctgcagga 3660
ttcgctggtctt ctctgcagga cagggctcct ttcgctgctt ctctgcagga cagggctcct 3720
agcacttcttg gagacat 3787
<210> SEQ ID NO: 295
<211> LENGTH: 2042
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 295

ggggcc gttcgcggga aagcttttgt cttccacctc atcatacaca caatattaat 60
cotcgggycg cttgacaggg cagaatacta acctcttggt tgtcttgggag tggcgattgt 120
gatttacag acacaaatgc ccaggtgtctc gttgagcgat ggcacgtgccc ttttggtgas 180
actcggacct gttgcaacac ctgggagagag tgttcttggg agaatctcata cttcgttgtga 240
ataqagtaco tcgaagagacgt tgggacagtga agatgtgagaa tgtgaagogt gccgtatagct 300
gtccgaaat tattaagaaag gagaattgtga tcgtatcaca atcataatc gaaatatgtag 360
taaatgtcctg tgttacaag aagaacagga atacgaataat tttatttctg ggtatttgta 420
gtggagggaga gtttcttgga gagataagag tcgaacagag cagctgtcga ccaagatgctc 480
agagattccttc caaacagctg atggcaggggg tgttttattc gctatgattc ggaatagctc 540
acagagatata taacccagga aacattcttg tgtgtgaaag ggtatgcaatc aacacttgac 600
acctcttgctc ggccgcactgt ccctgatgtg cgtgtgcttg caggtcattt gagatagagc 660
gtggattctt cccctgcttc gttcgcagag cccctgcttc aagatagacct cccctgcttc 720
cagttcttgctcc ctgtagctgt ggtggggatt tgtctttgat gttggggact ggtggagatgc 780
ggacccgac cacagtggcc tggactctag gatcgtgactt gcggagcagtt gcatggagctc 840
tccaccccttg gaaataactctct ctatgccttg gttggagagac aatcgaatct 900
agacacaaaag atccactacg acataaaaaa gatagatggt tacaacaaac 960
ccttcagaaaagggcctgacac tctatattc gcgtgggcg gatcggcatac 1020
gtggattctctc taagcattc taaccaacct tgtgctttct cccattacaa aeagtttt 1080
gtgggaccac gttgagtcg actaggctg gctacgagcc cggccaggtg cttcctttat 1140
gggtaccaag cccctcata attgataat ctggcagag cgggtccttgg tttcagagca 1200
catcgctcga tctctcttctt taagctctgc agttactctgg ccaccccagga tctocaga 1260
acccctggca ggcctggca aaaaagatga cagggtttct tacaatattg tggagcagac 1320
aatatttcata atgggttga aagactctgtg aagatgggg cttacatgtg aagagaaaattg 1380
gtatacttta gttactata tcaacactctgt atagagagaa ctaacactctg ttctcttatc 1440
tgaataatgt aaaaaatgttat ttttttttctt aaggggtattg 1500
gatggagac ttaagagcgac ttctcttgcag ttaaggggct caggtggtgtt tgtgagagac 1560
gccgagggc ttgtgcctct gcagacctgt cggccagctgg gcgggggttg atctgtggga 1620
atatactctgt ctaattcttct cctagagag aatattctgt cccgacactc 1680
gcataatactgccttctc agttgtctgttat taaaagggaa atgtctgctgcttctctttttc 1740
tgttttggct ggctacccag cttacttat cccacttttg tggacacagcct cggccaggg 1800
agttttattg agatttttttt attttttttttt ggtgtatttgg gttgatgtttt atgtggaaa 1860
cactctgtgg aaaaacaggt ctctgggagc tggatctttcc agcttttattt cagagctattc 1920
tcttttttat ccacacaaaa ctgggtttaacttac gaaataaaaaaata cttggtttt 1990
aatatttcata tgaacactata aagacaaact tttgcttacaaaa aaaaaaaaaaaaaaa 2040
<210> SEQ ID NO: 296
<211> LENGTH: 2547
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 296

```
ttacaaggt acaagctctct gctacagggg gcaagagggg tttatatagc atcatoccoa c 60
aggtgcgat cgcgctcagc gaaagctctct tctgggtctg gctcctctac gaaagctctct gaaagctctct 120
ggcggcggc gagggcggttt tgcgggttcg tcggggctg 180
aaagagggaa gtcgggtttg tctgggttcg tcggggctg 240

gctggcaag gagggcggc gagggcggttt gggggcggc gaaagctctct tctgggttcg tcggggctg gaaagctctct 300
ttcacaagag gggcggggtg gttttgctct tggggtcggc cgcctgggttt gggggcggc gaaagctctct tctgggttcg tcggggctg gaaagctctct 360
cttgggtctg gagggcggc gaaagctctct tctgggttcg tcggggctg 420
cgcgttcggc ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 480
ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 540
ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 600
ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 660
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 720
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 780
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 840
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 900
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 960
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 1020
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 1080
tgggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct ggggggagct gctctctctct 1140
```

```
cttggaacct ctggagttta ggaggttattc tttttacttg ccttaagggc tatcaccctt 2040
ttcctgagca tggagcactca ggtgcctgtg cagatcagat caccaagtgc gcatacaact 2100
tccctctctga agtcttggcc caagagctcag agaagacgct cttcttgcct tcagagactg 2160
tggtctgtca tccaaagctca cgttatctga cctagaagac ctttacacac ccgtgccttc 2220
agagactgac ccgtagagac aagtttacttc acctctgtgc ttgcagaaat gatccacagtc 2280
tcttacccca ggttctgtcc cagcctttctca ctatcgaaga ggcgacctcg gcagagggag 2340
cgaggggtgc cggagagctc caacgcgcag agtgggctcg tgcgttgcgt gcagctgctgt 2400
ggttgyacca cggcagaaat ctacacccct tccctctctct ctgcgyagtc 2460
tggttttttta tagttttatat ttaaatattg gcataaatttg cttttttaca gtccttgatg 2520
taactctaca acacccatag acacagc 2547

<210> SEQ ID NO: 297
<211> LENGTH: 2760
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 297

cactgtgctgg cagggccagg aagctccatg ccaatagccc agcaagagcg ccacacagcg 60
tgaagagagactcagagc aagagatagag ttagagagag aagagatgcct ctcataacccag 120
actggygcat ggasaccctgg cttctctctg cttccctgct tgcgtctctg tatctcttg 180
gaacacatt acatggcaacct tttaaagac ttggaaacct ccagggccac cctctcccttt 240
tttggggaaatt ttgttgctcg tcatcattag gcaggtgctgc gagctcagttgc 300
aaagaagatgg aaaaaggtgg ggccccctag atggcaccac gcgggtgggg gctctcagac 360
actcctgtac gattcaccacag tgcctgatgca aagcagatttt tcgggggttgtgtg acaaaaccgga 420
ggggtgctttgg cctccctcttg ttttggagaa tggccacctt ctttcgttctgtg gacagagat 480
ggaagagatt ccagatcagt cctgccttac caatcagcag ttggaagacg tctacacggt 540
ttcctctctac gtcctacagatt gggagagtgg tcggagaaaa tttgagggg gcagagcagg 600
cagggccaggg cttccagctcg ccagagcagct ttggggcctca cgcgctggat gtcctcaact 660
gccacttatt tagtgctgac atgctgccttct caaaacaccc ccagaccccc ctttgagaaa 720
acccaccagaa gtttttaagc ttgatatttt tcgatccttg ctttcctctca ataccgcttt 780
ttcctctctct ctctttccatt cttgaaagtt taatagcttc tgggttttca agagaagtt 840
caaatcctttaa aagaaacct gcagagagga ttgcaagcca gtagctccag gataacaaaa 900
gccacaggt ggattctcc cagtcgacag tggactctca gaaaccttca gacaoctaaga 960
caccaagaggg cttggccagt tgcggccagat tcagcttctt tttttttttt atctttctgtg 1020
gctagtcag cccagcctgc gttttctctt ctctttagta tcgtacctgg ctcactcccgt 1080
atcctgacga gacccctccc gagaaaaatt atcgcttttt ataccaatgg gcacacccca 1140
ccttgagtatg ctcttctgac ctggatctct ccgagctggt ggtgagagaa cgcggctcagat 1200
tacccacagt tggctgacat cttccctccttttttctgtgct cggagagaaaa gagttcttgg acaataggg 1260
tgctctctat caagaggggt gttgctgtgt cttccacta gcctctctcc cctgagccac 1320
agagctggac aggctgtcag agttctctcc caataaggcc ccataagagga 1380
acatagacgg ttaatcatac acaccctttt ggaattgacc cagaaacctgc atggcctagt 1440
-continued

ggttgtgtct catgacactg aaacctgtcct tatacagact ccttcacagt tctcctctca 1500
aaccctggca agaaacacag atctcctctga aatctatgct cgggtggctt ctttaaccag 1560
aaaaaccctg tgtctctcaac gttgagcaca aaatgtagga cacgtagaaat 1620
tcctctagcc ttcttcacag aatctgtgcc ccagtgcac acggagactca 1680
aatctcttg ctggactagag ctgacaaatg acgtgggtct ccacacattaa 1740
taacagggga tctctgacat gatagtgcag ctcctattgt cttgatgtac agtttacactt 1800
gggaatataa agagggtgac caaactcaagtg tggaggtcta gatttggtcct cttctgctct 1860
cacggaccct atctctcctcc ccacaggttac caccctttaac tctctgtaacat 1920
agaactacaac ttttaactaatg atctcttcgac caaactattta atgaaactatat 1980
agtggctcata acatctgcat tataatcata tgtttctcc tgcagtactc ataatttataa 2040
ttgttaacta atataacacaa agttaactctg gtttactgacac gatacctaa 2100
gagaactata tagaactgaa tgaacacag caaagataa ttttgccctt gtaatctac 2160
tgtggcgctct gggtccttct cgaactaagta attttgattat taactctatgg gaaagttaat 2220
cacattggac tttgccccat gtttagagact attaacctata gtttaattatat gcttttttgtt 2280
atgacaacac gttgctccac ccctgtaacc tgcgattgct ggggtcctgag cgggggggtat 2340
cgctctgagct caggggtcct aagacagct gctctagact gtgacacccc catctcact 2400
aaccactactac aactacagct ggctccgttg agctgctgct acgtctacat cacagggtcc 2460
tgaggtcag cacaatgcct aaacccttgag ggggtgtgag aatgtagctg agattgcaacc 2520
acgctccac aggctgggtg agagctgagac tcctctctct gtttcttatgg aaaaaataac 2580
cagctcact aagcagctc gtctctctcct aatatctgggg gacagcagaa acagtgttcct ctatctctctt 2640
ttttcagccc atctctctctc ctatctttttgg gacagcagaa acagtgtttcc ctatctctctt 2700
tgctctctc cccactccct cagactcctc gtagcaagta acacattata aaaaaaaatc 2760
gattgggt 27680

<210> SEQ ID NO 298
<211> LENGTH: 1358
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 298

gccgtcgcgg cgctgacacac tgccggctct gccagttgg gttgaggtca ggtgtaacttc 60
attctctgaag tacagacact gttgttggct tctctgtgtg cttctatctta agaggtgcttg 120
tttttccgaac tggataacac gatgctgacaa aagctgtcagc atggcttttg gttcataacct 180
gtggtgcgct cctctgtcgc acgaatacagag ctctactcctc ccagctttgct gaggactcttttctg 240
gcggagagac ggctgctctt gacagagatg acacgttaccc ttgctctgcc gacagagatg 300
tgggtggagtt gctggtggct cgtatttatg tatacggagtt tatacggagtt tatacggagtt 360
cgccaaacat ccaatagttt cttggccaaat gctctctgag agggaagtagg aagtgtgcccc 420
ggtgatcta gacagcagac atgagctgag ctcctactcc caaagatgcct tgaagctgct 480
aaccctcttg aggctcagag cttgtctctt cagatgacag gcaagagctg catataacatc 540
tgcgtcgaag caggctctac ttaaccacag gatacacttt ccagctgctga aagtgcgctgt 600
ggagagagtt cagagagctg ccgggaaagc agaagacacag ctgggagagct cagagagatga 660
agagtoctgt cagaaaaacg aggaggaagg ggaggagagc gctgagctgg agcaggggcc  720
cctactgctg gaggattgaag gctgcaagc aactgcctgg ttctoccaact cagtgagga  780
agagagggca gatgcgaacc tcggccaggg ttggctagact gtctgacac cggagaaggg  840
cggcagtctgc tcggcctgac acatcagcga gagccccttg gtcgctgtga gccgctccttg  900
tggtotcagg ctgctgctgg aactgattct tattccttgg gcactgcaccc ctgtttaaca  960
ttcatacaca ctgctgctag cttgcttcag cccatccctttt cttcatacaca ccacacacat  1020
ttgctccatc ggtggcagaga gaggatctct ttttctctga cccatcaacgtt caacaactgt  1080
ttatsactggt ttcctctgta ttatcctagt cagaaaatgt tcaattcatg ttctccctctg  1140
tgggagggca gttgctacgg ggggaggaga acotgctcaag tttgtcagat tcatcctgtat  1200
gattatcctg ttgatgcctct gatactgcc aactgccttg tcgggcaacg ctagaggcaag  1260
gatggaacat ctttactata ctaagaaag ggtgcagaga ttgtagctctg tgtgctcaca  1320
cgaggactata ctcacacagc gttttagaaaa aactggcaga cttggccagcg gattcag  1385
<210> SEQ ID NO: 299
<211> LENGTH: 4407
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 299

```
agtttgttcg gacacagcaaa gatgaaactct ggaatgtgaa tgaagatgaa gatcagatgg
1380
tcaaattcga aagcaggagtgt gatcagatgtg ctctttgatt gatttcgagt tatcatcaaaag
1440
taaagaga gcttcggtct ataataatga taacagatgag gctccgaaag atagatcag
1500
aaaaatcgtc caactgctt gacctggtat ttgcggtgtgc ttcctgtcatt gctttggcata
1560
acactaacgt ggaatgggttg aatgatggga agcagacctc ttcgatcatt cacaactaag
1620
tacagtoaca atatggagct tccttttctg ccagagctgta actacccctt ttttcaactt
1680
cattgactc ggtgagcatt atgtgagaaa tgtcggaggt gaagtttata aatccottttt
1740
cttggtctag cggctccacc ggcagacaa ctcctcagct ctagagactt tttgaagctg
1800
gatgaggggtt tgcctcaacc aaaaaccttct cctctgtataa ggcattttga acaaatggtt
1860
ccccccgtat cttccgggaa aaccocctgtg gcoccaaga tcgtccccag caaagattcttg
1920
tcttgagtac atgagaaaa gccgtcataa tctggtctca cagttcactg tgctttctccag
1980
atacaaggg tgtactccaag gacacatctg tggagctgag cattagttggc agttcacaata
2040
aaaaactgtg gaccaagact ctagagattc ttcgagcagat gccoctggagt
2100
taatatttac atgctccactg ggcattggag aagagagataa ggcctcgctct tgcgtcggag
2160
atgagactgt ggtgaggcgcc aatcccgccg cctttggccag gttcgacccg atttcttttt
2220
tttgcacagct gacaacacaa tgcagctaatt tggcagctgtg ccaagaggtg
2280
ggctgctcaag tcgcagctca ggcagccctc cctgtggagag taaaccattttg
2340
atggcagcagc gtcggaccaaga gttgggattg ccagaggaacc taatatagtggttggtggtcg
2400
ggcagcagcat cagacttcaat tgcggattg gtcgacccctt gcttgctcctg
2460
atgcttctgc aggctcagtt ggaatggattt aatccagcgct cattttcttg cgtcgctcgct
2520
atgctttggtc ttcctctgctt caggtagcta gtcggcttccg gtcctgaaggattcagttctg
2580
tcgactaggat gttccagttcg ctcacagcct cttctcttcat tggatggcgcc ccggatggcag
2640
aagactggga gacacagagc ccaagactctg tgcctcagoc gaaaggggaa ccagttcaccg
2700
gtcatttgca aagcagactg caagttggg gacattatcg ccagacggaa
2760
aagactaatg aagcagacac caagttgggt ataagtcctg tgcagctggg aatcattctg
2820
ttaagagaa cttttttacc cccaaaagct cttaccctta caataagactg ttaatagggac
2880
aagactttggc gtaccctggga acatggggtg attggagcc gcttgagcgaa gttgttgctgta
2940
tatgtgatga agaatttctg atcacttcag gtaaatctca ctatgccactg satgtttcgtg
3000
gctccgacaggt ttcataagtt caacagcagaa cccagctgcg cccacaaacc gacacgcta
3060
caggtggtac tgtggtctgc agtgggtgct cttgctgctg ctcacagcctc cttgggtctga
3120
ggcacacac gtcctgccag ccagatatgt gtcgatcagc gttgggaatt
3180
gttgatgtyg gacoagagctc gttgaccaatt cgtctcggct cattctccag atctactcata
3240
atcgctatcc ctgtagcagc ctttccttaccc attaactcca atataccttta ttaaataaca
3300
aatgtgattc tccacatttc atttgtactg ttcacagcttt ctctctgctg ccgctgctcc
3360
aatatgtgga ataatataaag ataatcttttt tctgtggaga gtagttaaat aacggtgtttgg
3420
cagccatttg gatgctctct gcgagttgtgc cttggtgaga aatctttctt tttggtggtcc
3480
atattgtgta cagttcctca aatcttgctg gctgtgccttc cttcatttgat ctcactattg
3540
agttgaaatc agatattttga acaacagact ctttacctaca acaacagataataacgag
3600
<210> SEQ ID NO 300
<211> LENGTH: 5532
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 300

gccgccctgc gcccgcgccc cgacgctgac cgggcgccgc cgccgcccag acggcgcag 60
agggcacttc gcggcgtgctc gggcgtgctc ggcggcgcac ggcctctgcg ccaaccccg 120
caacccgcct ggccggccag ccggggggtc cggggggggg cggggggggg cggggggggg 180
geagcagatgc gcggcgtgggg gggccggcgg cggggggttc cgacgcggttc gcggggttgg 240
tggcggcggc ggcgggtgcc ggagggagg aacgtctgag ggacggcag ccggcgctg 300
aacggcaggt gcgccttcggt gcggcgccgg gcggcgccgg cgacggttcc gggggggggg 360
gggtgtgcct tggcgtgctc gcggtgtgcct ggtggtggtc gggggggggg gggggggggg 420
tttaagcct cccggaggt cgggtgtgtg gggggggggg gggggggggg gggggggggg 480
attctctggt aacggccctg ggcggggttc cggggtttg gggggggggg gggggggggg 540
ttggcggtct tcggtcctg gggggggggg gggggggggg gggggggggg gggggggggg 600
atttttcagg cggggttcgg gggggggggg gggggggggg gggggggggg gggggggggg 660
gtggacgctc ggacgacggt ccggggtggc gggggggggg gggggggggg gggggggggg 720
ggctctccg cggggggttc gggggggggg gggggggggg gggggggggg gggggggggg 780
tgggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 840
tggcggggt cggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 900
ggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 960
acgggacagt ccggggtggc gggggggggg gggggggggg gggggggggg gggggggggg 1020
agtggcggc ggcccggggt cggggggggg gggggggggg gggggggggg gggggggggg 1080
tggcttcgg cggggggttc gggggggggg gggggggggg gggggggggg gggggggggg 1140
ggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1200
-continued

ggaatagata ttcggtgaaatt taaaagacct taatccctaa catatacagaa tatttacacat 1260
tccaaaaact gaaccccttg cagggtgat tccaccattc tcgggttggc atttaggggt 1320
gactocctca cacatactcc ttccttggat cccagaggaac tggatattctt gaaaaaccgta 1380
aagggaaatcc caggggttttc gctgattcag ggtggctgtg acacaggaac gacccctctac 1440
gctttgaga acctagaatt ctaacgcggg aggagcaagc aacagttcga gttttctttt 1500
gctagtccca gctggaacat aacatcctgg ggtttacgct ccccaagagga gataagctat 1560

gagtctgacg taattttcagg aaaaacaaat tttggtctag tcaatacact aaaaactggaa 1620
aacaactttg ccgacctccgg tccaaacaaac aaaaattata gcaacagaggg tgaacacagc 1680
tgcaagggca caggccaggt ctggcccacg ttggttcccg ccaagggtgg cttgggggcag 1740
gagggcgagg aatgggcttc ttgcccagat gtcagccgag ccagggatag ccctggccag 1800
tgcaagcttc tggagggga gcacagggag tttggtgaga aaccctgatgc cataacgctg 1860
cacccagaat gctcctccca ggcctgtaac ataaacctgc cagggagggg acacagacac 1920
tgtataccag tgcagccata ccttgcgcgg ccccaactcgcc tcaacgctctg cccgggaggag 1980
gctatgyggag aaaaacacac cccggcctcg agtaagcacg acggccgcca tgtgtgcacc 2040
cgtcgtcata caaaggctgc ctaagggctgc aagggcagcc ctgggttggcag ttggccacg 2100
aatggggcga agatccggcc cacctgcaact cgaggggtagg ggcggcctct gctggtgtcg 2160

gtgtggcccct ccgggctttt cttccttctt cggagggccc acctgtctcg ggcggcagcc 2220
cggccggagc gttgagggga ggggctggtc ggtgagccttc ttcaccaagag tggagacag 2280
ccccagggag ctttctcttg gacggttgcac caaacgagat caaaacagct caaagctttc 2340
ggctcggcag ctgcctggag ggtgatgaga gcctagctgca tcagccagct gggagagatt 2400
aaattcctcg ctgctctcaaa gtaatattgg gcggacacat tccgaccagc caacggaaga 2460
atctgtgatc agcgcagaat gttggcgcag tggccacaac cccagcggtyg cgctgctgct 2520

ggcgtatcgc tccctccgcag cgtgcaactc ataaggcgcag ttcagccgctt ggtgcctcct 2580
cgctgaggcc gctgggggca ccaagccact atttscctgt cagtcctctgc ccgggccctg 2640
gacacagag ccagagcttg gaaaaacagc ccagactgctg tgtgagctgtt gttgggtgtc 2700
gcacccctcg tggggcggcag cagagggagga taccctgttg ctcacccagc aaggtgacatc 2760
ccacccagcc tgggtgggga cggagggaga taccctgttg ctcacccagc aaggtgacatc 2820
aaggtgatcgc cttggactac aatattttacctg agatccgata cccagccggc ttgggtctcg 2880
agcgcgggag ctcgggggctt gggagagcg cctggttggc agatcctgtg gggagagcag 2940
cctgcaacgc gcgcctccct ccaagggcag aaagggaga cccagccgcc ggtggttctgg 3000
ttgctctcaag tgcgcagag tgggtggccc cggagttggtg ccctgggtcag ccggggtagc 3060
ccatctcggc tggggttggg cgtggtggtg ctcagccgag tgggggtgtt gggagagcg 3120
tgggttctcc agctactcatg cgggggctgtg cttggtggtg ctcagccgag tgggggtgtt 3180
cagctggtgag ccagccgcctg gcggtggtg ctcagccgag tgggggtgtt gggagagcg 3240
ccacagagct gcgcctccct ccaagggcag aaagggaga cccagccgcc ggtggttctgg 3300
agtgaacacg cacacacactc ccaaggtggtg tggggttggg gtaattggtg gcggggtagc 3360
ccccctcaggg ctcggggcag cttggactac aatattttacctg agatccgata cccagccggc 3420
gacacagag ccagagcttg gaaaaacagc ccagactgctg tgtgagctgtt gttgggtgtc 3480
<210> SEQ ID NO: 301
<211> LENGTH: 1528
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 302

cggcgagcgg caaattcagga cgcgttcgyc gcgacccctc gctgctgccc tcggagcagcg 60
gacccctcttg aacccacgac cgcggtgccc gcgacccctc gctgctgccc tcggagcagcg 120
acgaccctggc gcgggtccag gcgggtccag gcgggtccag gcgggtccag gcgggtccag 180
ggcggtcggcg caggttcgcttg cgttcggcgc gcgacccctc gctgctgccc tcggagcagcg 240
cgcgttcggc gcgggtccag gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 300
tataacgatgctg gcgtttggtt gcggagccgc gcgggtccag gcgggtccag gcgacccctc 360
ggtgctgcgc gcgtttggtt gcggagccgc gcgggtccag gcgggtccag gcgacccctc 420
agcaacattc ggacgggttgg gcgggtccag gcgggtccag gcgacccctc gctgctgccc 480
ggcggtcggcg caggttcgcttg cgttcggcgc gcgacccctc gctgctgccc tcggagcagcg 540
tttggttcgat gcggagccgc gcgggtccag gcgggtccag gcgacccctc gctgctgccc 600
catactttgatg cucagctggc tcgacacac ggacgggttgg gcggagccgc gcgggtccag 660
tttggttcgat gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 720
ccagttatg gcgtttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 780
aaaataacgg ggtcgtttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 840
aggtgagatg gcgtttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 900
tttggttcgat gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 960
goaggtgtatg gcgggtccag gcgacccctc gctgctgccc tcggagcagcg gcgtttggtt 1020
tttggttcgat gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 1080
gataagacgg acgggttggc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 1140
tataagacgg acgggttggc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 1200
acagcctggc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg gcgtttggtt 1260
tttggttcgat gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 1320
gtttttattgc gcggttttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 1380
acacctggc gcggttttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 1440
tttggttcgat gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 1500

gaggtgagatg gcgtttggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc 1528

<210> SEQ ID NO: 302
<211> LENGTH: 1856
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 302
tgaaactgtggc aacctgttgac aacctgttgac aacctgttgac aacctgttgac aacctgttgac 300
acgtgtcttgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg gcgtttggtt 360
tgcggtggtt gcggagccgc gcgggtccag gcgacccctc gctgctgccc tcggagcagcg 420
Continued

gcctttccaa ggtggaacg tcagatgagg agatcacaag cttcacaacg agatcagtaa
480
agttcyytttc ccoccacaccc tyyggagacaa gttgctcaca ccyygctgcc aactcacaact
540
aactgaagaa agttcatcgc tacctgcgga agataatctga cttggagaaag caggttgagg
600
ttccacaag cttcaccatgc tttacactga agatgaaagg gttggagccct cactcctcc
660
agttgaagcc ccocacagct gcggccgagc aataaccggaa gcocctgcgt gcgyctgcgg
720
gctggccgg cccttttccg cagttttcata gatcatcoco ccttctgatgc gacoccaaga
780
acocatcagct gcagcctcag caacgctttt gcacctctag cctctttcaca cctgctgctcag
840
gctttctgac gatccctcgc ccgggagctg ccaggaggctg ctaaccctgg ctgctgcttgc
900
acaagctgtg gctagagctg gggctccacc aatttcattc aacaccgggg gcacgggggtc
960
cctggtctgc cactctagct gcggagactg gctggccgga cctggagggg ctgctgctcga
1020
aacatggttttg ggttttacgc aacacttctgc cctggaccc gccttgggac agcgcttgctg
1080
gggtgtctgc gttgccctgc cagggagagc tgtggcttcag gcacccctgc gacacagacg
1140
tactttccac cctgcagcag gcagcctgccc acagctccag ccctgcaggg cagocctgga
1200
cctggtgcc cttgcgtgac tgccttcttc gcggctgctg gctgtcatag cttggagaca
1260
cttccacccg gccaactcg gacttccag gcctcgaggg ccaggagctc gaaacgctctc
1320
tctcgctgga gcggccgctg aaccactgca tgcctactgc gacacacggc aaccactcaat
1380
ttccttcggcc gcacccgctg gcacccgcgg ggcagcctgc gcagacagagc gcagacgacg
1440
ggatgaaagtc ctatggccgt gccgctgcc gctgctgctgc cttggagaga tttggacagta
1500
tgtcacaagct gcgtgtagttc agatcacaag cttcacaacg agatcagtaa
1560
gggtgccaccc tgggtggggtt gaggagcag gcagtgtcgcc cgggcacatgc cggcaagctc
1620
tgcttggtct gcgggtgtaggg tggcaacacc ccctctccac ccctgggagca gttggagccct
1680
agtggctccgc aggcttctcc tggggagatg aaccctttttc tggagaaaaat gttggctgcc
1740
gtgcctgtgc ccaggcggg ggcacagctgcc aaccctccag ccctcgccacc tacaactcaca
1800
acacacacct gcggctgagct aacggacaatg cttggagatg aaagctattt acctta
1856

<210> SEQ ID NO: 303
<211> LENGTH: 6450
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 303

gatggtgccc tggagtaggg tttattaacca tgtccgagggca actgagacaat gctgcaccg
60
cttcgacgc cttgtaagcc catttaagct cggagaccca gcacctttaag ttggaaggcc
120
ggagcagcc gcagcctcggg agggttggct gcggctgaga cctgcgggtat gcgggactcc
180
gcggccctgc gctgcggcag cggctcgggg cggccagcag gcgagcactg cgcgtgtcgg
240
cggggctgcc gctgcggcag cttcgcgcgc ggcggtgtcc gttggctccc gcgggctgcc
300
gtttctggc ccctggagcc gcgcgtccag cctgggagacgc gctggccgga cctggagggg
360
atgacgata cctcggacac cccagccgat ggtatggagcc ttcgctgatca gatccagagg
420
aggatgttgc aggctccgag atccagcgtcc cttgaggaac gcagctgtccg cggctcgggg
480
gggttgtaag cttcgcgcgc cggcgcgtgc gttgcacac gcggggctgc cggcgcgtgc
540
gtggctcacc ccggccgggg ccgggagacgc cggcgagctc gctggagggg cgggcgcgtgc
600
-continued

ggcoccggtg ctggagagtc ggcgtctggc tcaccagcggc tcgggggattt ccaccaaccc 660
aacaaggtgt ctccagacgc gcgtgatgta ctcgaacccg cgccagccggt gtcgcccttc 720
cctcaagcccc acgccccaca ggtgcctctac taccggagaa acagcccaag ccggtcacaag 780
gtygctcgagg cggcgcgtgc gcacccctac agggccacatt cgatattcgc acggcaggt 840
ggcagagaaa gatttggcag taccgaatgc aagggcaagtt tgggtatggat atctgcaag 900
gacytctgct acctgtgcaagt gtgcattga gtagctcag ctgactacatt tgggtctttg 960
tctgtgaggt gtcgctaggg cttotcctag agaagctatc aaggactataa cgaattatag 1020
tgtcgaaccc ccaaccagtg caccattgta aaaaacggga ggaagagcct gcagcgcgctc 1080
cggtctccga aatgtcaagaa atggtgaatgg atgaaaggtgc ggtatcaagaa aagaccaaga 1140
ggggggaaga tttggataac ccagcgcagc agatagtgatg gggggccgag gggggatgtg 1200
ggtttctgtgg gacacattag atgtgccaaa ctttgggccaa gcggctccaat gcat accaaagc 1260
ttataaaga aagcctgctgg cttgctccatg acgccagcag cagatggctgg tgcctcttgg 1320
gattgctgac ccaccatataa cttatcctac gatgtactta cccagcccct cagttgaagt 1380
tgtatgaggg gctttcagtc caccctgcca gcagacggag tgcctcaaact gatcaactctgg 1440
gcggagaggg tggcagctgg tgggtatggg agcctcaattc atctgcggaca ctttgcaga 1500	tgtgctgtgc tagagcatcct gatgtatggg atgtgctgctgc gttcattgga gcaccagcttg 1560
aatcactcgt tttgcctcttg cttgctccatg gcagagccac agggaaatct tgttagggcagc 1620
atgtggaga ctttcgacat gttgctggct acatcaactct gtttcggatg gatgaacttg 1680
caggggaggg atggttggct cctccacactt attatgttttc tttaattcttg aatgcaccaa 1740
tttgcctca gcaccctggaa gtcctcggaa gagaagagcata caccacag tctccccgagc 1800
aagtcctcag accctttgtgc ccaccctgtgc gcaccagccg gctccatgcc gcagacccgg 1860
caccaggggg tggccagctg cttccttcatc atctccatca tcagggcaa gataaaccac 1920
ggtatggagc atctgtcagc atgagatggc aacagctgtgc ccocctcota tgcocctctg 1980
cgggagagct gcaggcctag acggagcact gcacccacat cgcggggagg gcgtcagctg 2040
gggggcaggg accesacgca cttggccact ggcgcgtctta cttcctgcga cttccttgca 2100
aagttacca tcaccggggag gcggagggct ttccttcgcac gatgttggaga gcttctgcagc 2160
tccaccaggg tctccatataa cccttgcaca tttaaccctc atcagcacc acttattgcca 2220
aatctctctc cttcgtcaca cttccggtagc cttcccacatc ccattgcctt cctagatgctg 2280
ggcttcactt tgcgtctgtgc agtctttttg gcaccacctt cttgctctgtg tgtggacaccg 2340
cccaggttgac tcaccagctta aacagctttgc ttcocctcttg ctctagttact 2400
aagctgtgaggg atcccctgctag cttccttcacag atctagccag cttatggggt ggcgtcaca 2460
taatccagtgc cttttaagct acctgtgact accagggctg gcagagcata catctttctct 2520
tctgataacgc acctttaaat gcgtctagac atacagcataa gccagagcatt caagagccttat 2580
ccctttacctgc ggctcctgag tcacaggttct tttatagcaac cttctctgtat 2640
ttcatctagct actgacatttt tataagacctt aagcctttta tataatgctgata 2700
gcagagtatttg ggtttgttctc ccacacagct gacactaaccgg ggcaccacagc 2760
gggaggggac ccctctgtgt gggccagctg atttttacct ctgtaaaccg cagctccagc 2820
gttttgcgaa gtcgctggct gcggatgtgct gcaggtttgc tccagtttaa cagcttccacac 2880
atggacctat gaggacgacg aagttgactc tagttaatgc tc00cttatag gaggataagt 2940
tcttatttt ttttttatt tttgtttaac aaagaaga0 ot00ct000 cgaactctga 3000
gtaaggtcag ot0ccagacc tcctcagtgg gc0actctac tttgatcttc cg0gcttg 3060
ttgtgcttc acg00gtgga acgtcctct tgggtgacgg atggctgagaa taaatgttag 3120
ttssaaggg cgagggcgcct gttgtgtcct ttaa00cttg ggctgtgacg tgaacatgtc 3180
tggycagga tgttgtggcc tatactagaa caaqagggaa agtatyccgc aataaagtgaa 3240
caggttctag cac0ccag0 ac0c0ttcag gggccctcga caggtgtcag ctacctatgg 3300
acacct0ctg cagggcccgct atg0cttctttt gggttgccct tgttatcctgg ctgtttcctca 3360
gctttatctc att0ccagtc tgtggccttt g0g0aaaaag cag0cttca0a gttgct0aca 3420
gctgtgtc0c tatataaggg ccac0ccacg gggccgcttg aaagggcttg ggccgctgcg 3480
ttgtctact ctgtgtactg ggggctcctgt cagacctagt atggcttcttg tgtgtttcaga 3540
ataactc0aa atac00gtttt cgtgttgcca aaaaaaatc cc000c0ttt cc000c0ccc 3600
gttctctac g00ct00ccc ctg00ccctgc att000cttc atttttttttt actacag0c 3660
taaaaaga0 aggtc0cttc cagccccag gc0c0cttcc ctg000ctcttgt atctctctag 3720
cacaatttg attgtc0ctt tttttttctaa ccc0000aga aa00cttttt ttttctttgct 3780
ac0ccttttct ctgatcttga aac0ccatgt agaggtgtc0c tctgtgttag ccaatg0c00c 3840
aggtccttgg ctggctttct tttgttgttgaa aagggggtt tgttatctttt tctttctttt 3900
cgtgtgtc0c cggatgatt tc0000aacg cc00000c00 tgggaggtgc aaaaaaaatttt 3960
a00000tct gc0ctc0ct tttttttttt gct000c00 tttttttttt tttttctttttt ttacact0ttc a0c0gttc0t a0c0gtt0ata 4020
gcttatagc atacttttt ctggtc0tgt ctgttttttg c0000000aa gc0000c000 cc00000000 4080
acrrrrotact atatattttctt ttttttttt0 g0000000atc attttttttt ttttctttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt tttttttttt
cgcctttgcc tagctttgcct taagtattct ataagctcat ctgctagcgca ttatggaggg 5220
cagcgtctc tgcgtgctga acacatgtca atctggtgcc aaatcttacc ccactaagca 5280
gttcctggt tggatttgg atcactatat cagtcagctt atgttttaac gagcttcgcg 5340
cactgcctt cgcgtgctga acacatgtca atctggtgcc aaatcttacc ccactaagca 5400
ttaaggtgct ctcctttgcgg atcatttata cttgaaaggg gttgatttgg atcactatat 5460
ttcagcagct cctctgcttc gttgactctg attccaccca ttctctcatt ccaatctttc 5520
attatgaga tggacttgcg tgaactggagg tgcacactaa ccaactagta atgctatatta 5580
ttcacgctgtgccatgcgtgcca gggagctttg tgttcatggtt cttggtctgc 5640
agtcttccgg gccgtagggc acacactgct gatgctgtcc gtatgctgtc ctctgtggttac 5700
actctgctgct atcactatat cttgaaaggg gttgatttgg atcactatat cttggtctgc 5760
ttcagcagct cctctgcttc gttgactctg attccaccca ttctctcatt ccaatctttc 5820
attatgaga tggacttgcg tgaactggagg tgcacactaa ccaactagta atgctatatta 5880
ttcacgctgtgccatgcgtgcca gggagctttg tgttcatggtt cttggtctgc 5940
ttcagcagct cctctgcttc gttgactctg attccaccca ttctctcatt ccaatctttc 6000
gtctgctgtgccatgcgtgcca gggagctttg tgttcatggtt cttggtctgc 6060
gttcctggt ctcgtgctga agaacacgac ccctggcgct tgcagagagc gttatggtgt 6120
attagctgca cagactgttc atctggtgcc aaatcttacc ccactaagca 6180
catgctttt cttgacactc ataacttcaa gttgagccgt cttgttttat gggaaagggc 6240
tccagtgcga aatctggtttc gatgctgcga ttcgactgac agagcttttc 6300
gatgtgacgt ggtttttgcag caggtgtgtc tttgattcgg aaacacttgg taacacttgg 6360
tttgacactg caagagagc gttgactctg attccaccca ttctctcatt ccaatctttc 6420
cattgctgtgccatgcgtgcca gggagctttg tgttcatggtt cttggtctgc 6480
<210> SEQ ID NO: 304
<211> LENGTH: 3336
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (0)...(0)
<223> OTHER INFORMATION: n = A, T, C or G
<400> SEQUENCE: 304
cgggcgcgcg tgaactgctt gaggccccag ccctgtttgc ataagagcgg gttgaaacgc 60
agatctgtta tgcagctgg cccgcttgg cccgctttga tggcagctgc gggaggtggc 120
cctcgagttta aatctggtctc cttgagggag ccactaagca atctggtgc 180
cgtcttggtt caaactgctg agagctttgc gttgactctg attccaccca ttctctcatt 240
tttgacactc ggtgacttgc atctggtgtc ccctgtttgc ataagagcgg gttgaaacgc 300
atccacctgc tttgactttg atctggtgtc ccctgtttgc ataagagcgg gttgaaacgc 360
agagctttgc gttgacttgc atctggtgtc ccctgtttgc ataagagcgg gttgaaacgc 420
cctcaagct ttcctgtttg ccctgtttgc ataagagcgg gttgaaacgc 480
cctcgagtttg ccctgtttgc ataagagcgg gttgaaacgc 540
gcctggtgc gggagctttgc atctggtgtc ccctgtttgc ataagagcgg gttgaaacgc 600
aacatagct taccaacat ccagttgcttc ggaaagatga gttctgatgy actggygctcc 660
cgcagctaca agcagagat ggaggaagaag ggagaatttc acctgaggaac gcgcaggaat 720
aaggttgagg agccttgcag acacatcgct tcctggcaga actctgtgtgc tgaogagcca 780
cctcctctcct ataagccacgt gacacttacc ggaacttacc gctcaacctg gaagcacgtg 840
actttaaag acaatctacag tggagttagag gacacatttct ctaactttaaa gcocattgcc 900
aagcagcgtag ggagaaactc atacggcaac aoccttttcc tcaacgcacat gtggtcggc 960
gacagctcg cccgtaaacc ggtcttccc tggacattc aacccagctg caaogcgtac 1020
ttgacattc acacgctgtg taagccactg gcacccggtg gttccacacc atcactaaac 1080
tgggtctaag agccagacaco ctggcagttg atgggctagt gtcgtaatgc 1140
gacctttccc cttggggcaag ggaagagatg aagccactgc cacccagcttg cagtcttatc 1200
cctgtatcata tccagtctgc ggtgaaccag taaggtgattg tgcacggtc ctttgaagggg 1260
ccactttccc ccctggtgtcct cctctggtgc tggagagtgtgc tgcagatattc cttaaggggt 1320
cagctttccc cccagctgtt gacggtgctag gggggagggt cttctccttc ttctgcaaga 1380
ccagyyaag aaggaacttc cctcttgggga agggggttatt cctctttctgt ccacggtcag 1440
acacttatgc agccagagat ccctctggag ggaagagact gcaacttcgag gccacactac 1500
aaggtggaga gcccctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 1560
ctctctact ccctctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 1620
agcgggotta gttcgaacaa cccgctgtgc tggagagatc tgggagattc acacagggag 1680
agcggggtgc gggggcagt cctctcctct gcagagagtg ccctctccggc gcctcctctt caaagggag 1740
ccagcagact ccctctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 1800
ccagcagact ccctctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 1860
ccacagacat tgggctgctaca aagcagtggc atcgctgcag tgggaagagc gacggtgctt 1920
ccacagacat tgggctgctaca aagcagtggc atcgctgcag tgggaagagc gacggtgctt 1980
agcacaagaa acacgcocca gggggtgtac gacagcaggat ccctggtgctg tgagcttggt 2040
gaccctgcag ccctctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 2100
agtcagacat cttgctgctaca gacagatcag cttgctgctaca gacagatcag cttgctgctaca 2160
gaccctgcag ccctctctct gcagagagtg ccctctccggc gcctcctctt caaagggag 2220
ctacgacagat tggagagatc tgggagattc acacagggag 2280
ctacgacagat tggagagatc tgggagattc acacagggag 2340
ccagggcaag aaagggatac gcggcggcggc gacagagagtg ccctctctct gcagagagtg 2400
gctctgttcc ttcctgcttc cctctggtgc tggagagatc tgggagattc acacagggag 2460
cggcggcggc gacagagagtg ccctctctct gcagagagtg ccctctctct gcagagagtg 2520
cggcggcggc gacagagagtg ccctctctct gcagagagtg ccctctctct gcagagagtg 2580
cggcggcggc gacagagagtg ccctctctct gcagagagtg ccctctctct gcagagagtg 2640
tctactacct gcggcaggat ctctctctct gcagagagtg ccctctctct gcagagagtg 2700
tgcagcagatt tggagagatc tgggagattc acacagggag 2760
tgcagcagatt tggagagatc tgggagattc acacagggag 2820
ccggttgctgctaca aagcagtggc atcgctgcag tgggaagagc gacggtgctt 2880
-continued

acatggatatct tgggttcttc actgcaggga cccagcaag tggatctgcc tgccagagtc 2940
ccttttyccct cttcctyccca ctctcccctt tttccacctc agctttctcttg caagaaagaa 3000
tctctgtaaa aaagcttctt tgtatggtgct cagaggtttga atttggtgtag gggaggtgga 3060
tgctcaactgc gcccagttgtg gttgcctcga tctgcgtctct tagatgcttc tctgtatcctg 3120
tcccactata taccagggag actgcagagct acgagaatac aggtgagagg ttagaatgagcc 3180
cgaagggcct cctgactctgc ctggctccctgt tagctggccc ctcagctttgy ccaagagccaa 3240
cctcaggtcgc cagctyccagc cagctgtttgct agcagctttg agaaccactta ctactcactaa 3300
aatacgcaggg tggacccnaa aaaaacaaa aaaaaaa 3336

<210> SEQ ID NO: 305
<211> LENGTH: 2365
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 305
tccagcctt ccctcccctcccc cccagagaagc aaaaattttaa acgacccccgc accttcggac 60
ggcagagcc cccagacgct cccggcgac ccgctccctc tcctccgctcg ggttccgggc 120
cctgagcgag gcgcggcagc cgcagcggcc gcggctctgtt ggcgcgctcc cgccgcgc 180
gtgccgaccc caccaccccg ccgccgctcc ca cggccaccgc ccccacccggg 240
cctcgccct tctctcaggg acgcgggcgg gtcgcctgcgtt cggagagg gtagtgctct 300
ttttaaacag cggctctcag gcacacagcgt ccggccctcc tccagctgcg cggctccccg 360
cgcggctcg aygytccctg cagcaccaccc gcggccctcc ggtgcgccccg cgccctctcg 420
tcctggatcc ctaccccctgc tggccggcgg ccaaccgctgg ggaagcgcac gcacccgctc 480
cctctggac ccctcgtcctc ttctcagaggc gtcctccctcc caccggctcc cggggccccct 540
tcccctcact cccgccccct cgtctccttc tctgtggggt gggcggtgggg gaacacacggc cccacccct 600
tcctccccttc cccggccccc tgccttcccg ccgctccgag cggcccccctcc ctcaccgcccg 660
ggcgtcgaggc gcggcgcgcttc ccggggaggt caagagagtgc tctagatcggc agtggcgccct 720
gccacacccgtgctcggcg gagctctgctt ggtttccgctgc ctctccgctc agcgagcgcc cctgctgggtg 780
agccgctcctcg tggccgagac acaccacatt caaccccctgg cctcactcctgc cgggaacgtcag 840
tccgcgtcag tccctccctc ggcacctgtct ggccccgctt ccccccctgcc tggagccct 900
tctccagccct ccagcttctct gcggcctccc ccggcgcgtt gcacccgctcc tggagctc 960
gacccagctcg tcggccggag atggggccag aacacatgct tgcacggtcg ggggctctca 1020
tccacaaaatg caagccacaga cccggccctcc cattacggcc aagcgacaagc tgcctcagcg 1080
cggggagcc gcagggcgcgg tgcgtcaggt gcctggctcctg tccacacacat cccagggcagc 1140
gaatcgcaag gcggccctgg cctctgctct gcctggttcgg tgcctcagag tcctacactgctc 1200
tccagagccc tctccctgtc acagggaagc ctcctccgcttc aagacaccagaa aactgctcg 1260
cacccacaccag atggcgccag ccagcgtgggg ccaggtcctgg gcacccaccc cccagcggctg 1320
gtctcaaccc gtcgctcctg ccagcagcctt gcctgggtcctt ctggctcagc ctgcgggccgc 1380
caccctagcgcc cagctgctgg tccgagccag gcctgctggg tccgagcgcc gcctgctggc 1440
tgccacacgcc cccctcccctgc tctgtccttg gcgttcgctt ccctcttgctgctc 1500
acagggcccc cccggcagagt ccgctcggct cctctcctctg tcatttttcg cggagcaagt 1560
atcatgaagcc tttaacgcca tggtatatag ttttggaagg caaaaagcaaa aatttgtttt
1620
ggcaccttgc aaggaagctc acgtggygtg ctgtggttca accactgtaat ctggacccca
1680
tctgtgataa agccatcctg acatoaatcc cctatattaac agggctcota tgcgtgtaaa
1740
asasasasas atcatgacat gctataactc tataatgtaa gasesacactg caactgctctt
1800
tatgtgcatc gggtacgtgt aagggcctgga ggtgocasag aagtttaaag aatatggtgag
1860
aatagttgtga aaaaaaaca aagaaaaaaaag acgtatattt acagtcacaa caactgccagtt
1920
ttggttctcc toctgagcaca caggtttggg agtggatatt agcggattaaaa aaaaaa
aaaggagaaa aaaaaaag aaaaaaaggg tcgggtaatc aagaaactga aaggtttggccc
2040
cctctgcaac gggatataca gttcgggga acaatactgta atgttcacac gttgcccattg
2100
aggttcatca aagagcccttt tactggcctta cttccttttg gacaacgttccc ctgaaatgtg
2160
tgctttcgtg taaaacctca aagaacacaa aagaaaaag aagttattta ttgggttctaa
2220
ttataccagc caaatggagta taaaaattta ttatactgctta gctctaaaga aagttgctttc
2280
tgcttttacct gttcaactta tttccatcct toacaatattt cgggttgaata aactagattta
2340
catcaagttg gcaaaaaaaa aaaaaa
2365
<210> SEQ ID NO 306
<211> LENGTH: 1117
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 306
gcccaaaccc gcacadctgca cagataactc ggtctcgggg gccggtgcccac
60
gccctccagc tgtctctggg ataacaacaag tcaagctatg aggaaaaaga gtaacagatg
120
ggggagcgcct cagttgtgctg cagaaacagag gtcggtgttgg aaaaaatactc ggtggccctg
180
gggtttcctaa acctgcttctt cttgagctat ggggttcaca agataacca aagaaaaagc
240
acctgcttc acatggcgcg caagcaacac ctgctttgctg gagacagaaa gggaaaaaatt
300
tgtgtggacag ttggcttaag ccaacacagct gaaacgtggg catagatgctc
360
tcacacaacg aagttactaaa aagttactggg gagaacacctt aaacgtttttt
420
aagctctcct cagagttcttt ggggaacgagg cagagttcttt ggaagaaaaa gataacttttt
480
tatagctttt tcgctatata tcgcctttgac cttcactagtatttgaccc tatttggagc caactgctcg
540
gacgcttccc cagagctgga ggggtctctat tctccttatt ggggtctttg ggaagaaaaa gataacttttt
600
gcactatac aggggcgcctg ctctctctaa aagaaaaag cggcgtgctg ggtggccctg
660
gggagcaagt aaggggtgga aagggagggg ggagagaaggg aaagaaaaag aagaaaaag
720
gcaggccgct ggcctctctg cttcctctgct tccctcttct tccctcttct ccccccaac
780
tcgtggttat cccctctttt actactaagaa ctttacactc ctttacactc cctctctctc
840
caccttcagc gcagaggcttc tggcctgccag caactctatc tctctctatg gaacagccaa
900
tccctctct tccctctctc gcacccgcag cgcgtgcttc tctcctcttgt tttgtgtggc
960
tcgtggtat tttacgctggg cagtgcctgtgct tggccgagc gtcggtggccag
1020
tcgtggtat aaggggactg ggagagaaggg ggagagagaaaa gacagacacaa
1080
gcctggttc acctgatgga aagttactggg gagaacacctt aaacgtttttt
1117
ctcgagaagcc cgctccacctg ccgctgagct ctctctgtct tctcgggtac tyggatatcc  
60
gtggtgtgc gcacgacact cgctctgcttc tggagtttac ggataccttc tatgagaga  
120
aagcctaccc gtggctggaa ggctctgact atgatcgaag ccaatgcgtct gatgtgaat  
180
tccagctgca ccggccctgc cctactctgc gctacgcttg ggtggaggg aacaagatca  
240
cccagagcgg cgctctgcc tgcctacccg cgggtacagc cggacacggt cagcagcctg  
300
aagaaaaaatt gatggagtgc gacctagcag gagaagagtct aagagatggtt ccagcaaac  
360
tgataagctg ctgtaagcgg cttcaccacag aaaaacgttg ggcctcagtac tggagaagcc  
420
taacctggaac actgcagcgc ttcctctgct ttcggttgaa atctccatgg tttgcgcggg  
480
aagtcagcgg cttggtggat ttcctcctgt atgatcctag ggtgaagagc ccgcttacag  
540
aagccccacg cggcgctggag ttccccaacc ttgaaacatt gatgctcctg tttgcgcggg  
600
tgagaaaaat ccgctcctgc ttcacagctg atcaagtttcg caagagccgct cctcaacaca  
660
agatgccccg gtggagggcg aagcagcatt gttgagctcg agggccatct ggcgagtcgt  
720
ctctctctca ctcctgcttc aagggctcag ggccttctcg tggctctctct caagacagag  
780
cggatgctgc gacatgcaccc ttctgcttcct gttgcgcggg taaagccaa aagggaaaaa  
840
ggatgctgc gacatgcaccc ttctgcttcct gttgcgcggg taaagccaa aagggaaaaa  
900
atctccaccc gtagcgcaggg ggtcgcgatct tgcctccacag ggcctcactg gggyttctcg  
960
aagaaaagccg tttcagctgt tggagagagc agggctgtaa aaggggcatt ggctccctgt  
1020
aagactcggg gacagcagcg cgagcagcctc tggcttctcc tggctcctca cggggctgct  
1080
uuacacgccc ttggagtttga ggagagatgg ggaagatttt ggtccttatt ggtctgctcc  
1140
ggcctccag cagcagcttc ggtgctgatc ggagagatgg ggaagatttt ggtccttatt  
1200
tgacaactat taactaatgg ttcacagccg tttacgcccc aacgaaaa acagcagctg  
1260
aatggt  
1266
---continued---

gggttgttctg atacactgcc tctctgtgaa gggtgcagctg gtgtgcctgc cctcttttgt 600
catccacagt atcaagactg tgcctttaaa ttcatttgtt ggcatactgc agggactctg 660
gttgggtctg gctggccccct tgcctgccag atcagacagc cctgactcga gaacctctta 720
cctagcgggc ttctttgcct tctggtgcct cttgctcttg ctgtccagtt aagccgactc 780
atgaaactg gcttggccct acctttatca cagctggtgc gtatactttt tgcacactc 840
catttactcg gccgtgcccag cgttggaatt ctacgctac taccgccagc tcaagcttga 900
aggacccggg ggcacaggag ccaagtgtga cctctattgc aagggagagg agccaagacg 960
agggccagcg gacatttggc ctttacgctc cagctctcac cccccaaatg aagccacccc 1020
tataacaggg atatattaag aatatcctttt cctggttcct tcttacgctt toacattcag 1080
tatctctctt gctggatcgc cagctctgtc tctgtgagtt aagccacaggc tccgggagac 1140
cagacgctgg ggaacctgact tcaatttcgt gcctcgttgt tgtttctctc caaagttaaa 1200
cgttggtcgc cgcagctgtc cagcttgcct cggtaagcct gcggcttgctgt gcacattcag 1260
gccgacggc cagagctggtg cggtgctggt gcgttacgct cgtctgtggct gcacattcag 1320
gccgacggc cagagctggtg cggtgctggt gcgttacgct cgtctgtggct gcacattcag 1380
tgctctggtc tctcagcgct ggcctccgtc atggccctcg gcgccagaga 1440
agtgacagca gtgtggccag agacgagcac aagctcttcg gcccttttct cctctttttt 1500
tctgtcagc cgggttcttt tctctttttg gcctgggcca atgggtcctg aagggagagg 1560
cagagagtct gctctgcttc ctgcttcgct gcctgctctc gcctcttcttc gcctgctctc 1620
atctctgtg gtcgcctggtt ttttttcttt atgcctccttc gcctcttcgg gcggcttgct 1680
gggcogggg atcagagacg cggggaggga ggttgctggt gcacagtgac gacattgctg 1740
gttgggctct cagcgctgct gccgcgggtcg tgcctgctgc atttccgctg gttttctgcc 1800
tctctggtct gcaatcctcg ctgctgtgac gcggacctga ggtgtccctgg ctggagaca 1860
acgctgtctg ctgctgtctg gcgtcgtcgt gcgtcgtacg cagctctgtg 1920

tggtctgggt gtggctggg acgtgcgtcg ctggctgtac ctggctgtac 1980
tctctgctgc gcctctgcct ctgctgtttt ttcagcctcc cctccacact ccccagcggc 2040
agttttttacc catctgctgac ctctattcct tcagccgta tgccatcatt taaaatata 2100

tggtcagcag cccagctggtc cttcgagctg acctgatttt ttcatttttt 2160

tt 2162
gaactgcag tttttatatga gcttgcctac cagttgcccc tcocacataa tgtgagttcg
420
catctgtgta aggctotgtg gagaggttt acacacegcct attggctgtg gaggaacct
480
tggtgatcgc tgtgatgatga tatgaaagat gacagtaaag cacagacgcag aatgcttttatt
540
ttgagaacgt tgtatgtgttgt ctacgatcgc atggtgctat tgttaaacatt
600
tatatggaa gcaacgacta actaggtattt actatacttg acacaagttg
660
tttgatattta ctctacaatg tgcaccagag gcaagcagaa acatggctac acacagaatt
720
ggctgctgaa aasaggtgta aagacaaac agacacgcgag gttttttctt cagaatgaa
780
tgtgccctaa cctcagccag aagactcttg aacaaactg ctggcacaatg gaggattg
840
cactgccgag gccacactga cgtatatgat acacacatg acaacaacctg tgcgggtat
900
aagacacccg ctcagctctg atggctgtgtg atttggaccc cccctctcct cccctcctaa
960
attgacactt gttatagacta gaaacacttc ctaccagcag acgctcgag agatgaattt
1020
tttttttatt aagaaaggat taccgatag atggtgtatg agccacagaa aacttttagcc
1080
cgctcaattt atgatattta tctatgtgctt gcctctgtat acttgtcagc aactcctcctat
1140
gatagttta ctaaaagcag atagccaca ggcagctaca ggtgctcttg caaagagggt
1200
ngagatgctc gggtgcaaat acacacactt gctaatattg aacacacagaa tcctcaccct
1260
cagctgcttg tctgtgtgaa ttaacctggtg aatgctgatta ttcgccactc cttgatatttc
1320
tctctccaa aaccacagcg tgcctatcag cccggtttgc cttcagatatt gaaactaggct
1380
cagcttacct caaagagttgc atcgaagaa atcaagactt atctttcagaa acacgaaagt
1440
gcacgatgy ctctctctct cggcgcgctc gccgctgtct gcacactcct atctttctat
1500
ctccggcaga accacaaagc aactgtagac caacaccttg agaaggatcc attataataat
1560
gactgatacg tcctctcacc ccagacaaaa atttggtattg cattgctctca
1620
attcctatgg ctcagcagctg acatctctcg gataagccag atggagctctg atctcaatc
1680
aatgcttcat ctaaaacgct cggcctgata cccacacacag cttggaggatct gatgctcacc
1740
ctagtcag acataaagtc agtttctgta atgggaagag atggagttgtc gatggtcacc
1800
cttcattgc gcgctgcata tttttttgtg atggctcttg atcggctcaat tcgaaacagt
1860
ctgaggattg taaaattaaa attattatg gacaacagaa acataa tttttattct
1920
cagacagc atatttgcatt ggaaggtttgattgctctata ttctcattgaa ttagggctct
1980
cagtagggt tcctctcact tcggctcacc ttagaagcc agtcggcag tcctgataac
2040
gcagctttca aacacagctct cagctgattc caataacag ggaggtctgt
2100
aagccacca ctacccactg cacoactgat gattttaaac caagcatcaca aacacagtct
2160
gagacatac aatattgatg tgcctctccaa tccttccacc caatatcagag gaactactaatc
2220
agcggcacc atcagactgct tagagatcc caagctggyg cagctctcacc aacagagcg
2280
gggaagagtg tctagagaca gacagaaaaa tttcttcccag gcctttttatg tgcgggtat
2340
ctggctttgg ctcacagacag tctgatcgatc agggaagag tatactccac gatagtctgtc
2400
tgcgcgaca gctccagagaa gcaagatgac ggtgtaggtc ttcgagatctt caagactata
2460
ggagtgcacac catatttaca gacgacagac gctccagcag ctcagacatc actttttttg
2520
aataggtgtg aaggactgaa tctatgcttg cagcatgcca tggcagccaaa gacaaatatt
2580
ttaaaaacgct tttttgtgaa gcaacagaa acataa tttttattct
2640
cGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
2700
GGGCTTGG AAACGAGG GACCGGAGG GCAGGTGAG
2760
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
2820
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
2880
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
2940
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3000
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3060
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3120
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3180
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3240
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3300
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3360
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3420
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3480
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3540
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3600
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3660
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3720
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3780
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3840
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3900
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
3960
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4020
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4080
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4140
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4200
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4260
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4320
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4380
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4440
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4500
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4560
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4620
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4680
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4740
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4800
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4860
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4920
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
4980
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
5040
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
5100
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
5160
cGGGAGGTG CATCGAGG AGCCTGCT CAGCAGGGA ACGGCTGAG
5220
cagcagcgt gcggagaactc catcggcgcac tcggtgctct tcggtacgct gttcgtcaag 780
gtggcagct ccccgcccaaa gcggggaag ggcgtacact gcgggtgcaca ccgagcgtc 840
ggcacaggt ttggcagaaag tgggtcacttg ggcggcagga agggggtcagaa aaatgcagaa 900
cagcggggggc ccggcgccgg ggcggggaagg ggccggcgcc ccagggccgg 960
cctgagccgg ccgagggccc ctcggtgcggct ccgagccaaa gcgcaggtgccc gcccggcg 1020
cgggtgtgc acgggagaca cgcggcgcga cgcgggcggc cgcggcgccc cgcggcgccc 1080
dagcccaaga ctctggtcagac ccagggccgg cgcggcgccgc cgcggcgccc cgcggcgccc 1140
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1200
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1260
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1320
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1380
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1440
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1500
tccgaggtgc ctcggtgcggct ccagggccgg cgcggcgccc cgcggcgccc cgcggcgccc 1560
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1620
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1680
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1740
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1800
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1860
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1920
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 1980
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2040
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2100
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2160
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2220
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2280
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2340
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2400
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2460
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2520
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2580
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2640
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2700
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2760
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2820
acgcagtttc ccgggacgca aagcggaggg agcggagggg gggagagggg ggagagggg 2872

<210> SEQ ID NO: 311
<211> LENGTH: 926
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 311

```plaintext
    ggggccacgt ccagctgtgg gcgccgctgg ccagctgtgg gcgccgctgg gcgccgctgg
       60
    ccagagagcc cagacagacg catcagagcc cagacagacg catcagagcc cagacagacg
       120
    ggcgcagtgc cagctgtgg gcgccgctgg ccagctgtgg gcgccgctgg gcgccgctgg
       180
    ccagagagcc cagacagacg catcagagcc cagacagacg catcagagcc cagacagacg
       240
    ccagagagcc cagacagacg catcagagcc cagacagacg catcagagcc cagacagacg
       300
    gcaggtgtgc aagcagaagc cagcagccac gccagcagcc cagcagccac gccagcagcc
       360
    actgcggcgct gcggcgct gcggcgct gcggcgct gcggcgct gcggcgct gcggcgct gcggcgct
       420
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       480
    atgcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       540
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       600
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       660
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       720
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       780
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       840
    gcagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc cagcagccac gccagcagcc
       900
    aaaaataaa aaaaaa aaaaaa aaaaaa aaaaaa aaaaaa aaaaaa aaaaaa
       960
```
-continued

gagtgcctct

gagatgtaaga

ggtctttccta

tctctcactc

tggcttgtaa

tgtgtaaga

ttttttttccc

tcttttattt

ttacactccttt

dgatgatcctt

tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ctggtgttg ctcgccaaag gccagccacc ctcgctcatc tgcgactgtg gacaaggggc 3300
gatgcaaaa gttatctgcc gttctctggg ccacaatag aatataacct ggtgtcagaa 3360
cctgcacgcc tgaacagatt gattcagtgt gcggccagaa tgcgcaacgg ctagggcatac 3420
cctgcacgcc atagcttgtg cccgcaagacc cttgctgccc gcgattgcgt gttgacgcga 3480
gatttcacg tcacacatcc agatgtttgt atgaagcggag atatctatga gacagacctat 3540
tacggaagaag gaggcaaaag gttggtgcgc ggctgctgga tgcgcttgga gtcgctoaag 3600
gatggcctct tccagccactt ctgcgagcct gttcctccgg ggtggtgcct tcgggagatc 3660
gccagaagcg cgcagggccc ctacagcgctgct tcgcacggc agcagagacct tcgcgctgtcc 3720
atggagggcg gctctgctag ccgccagtcg acatgtctgt aacatgtttt tgaactgagat 3780
cgtatttcgt ggccttataa cccgaagactg aggtctctcc tcctgcctgg gatcagcagc 3840
atgaaaggg agatggagagc tcggttccgg ggtgccttct ttatactacag gcgagagaa 3900
aagctgcgag ccgcggagac ggtggccttg cacgacagaa acatggagag cgtcctccct 3960
gacgagtcgg ctcgctccgtc ctcgccgca ctgcagcaga gacactcagg accaagggcg 4020
gagagctgcc cggtgtgtgg cggatgctgc ctcgctgcga gttgagcaga gagaagcct 4080
tcgccacca tggcagggg gggagcagac cccggcgctc cgcgcgtgc cccagccctg 4140
acctgctct cttgtgctgc ccagcttttg caagagctct gcggcggcgc cccgagcggg 4200
 tggtgagggc gagagagcct tttacaaacct cttcagcagc ctcctgtggtg cgcctgctct 4260
cctgagttct gcccgggttg ccggcggacct acaggtcttc tgcgatccaa caacattggg 4320
agtctctct tttcatactag caagcagctt tttatactcc gcccacaacc ttcaacggca 4380
tgcgtattta agccatctg catcgaacat ttaagcaggt agagcaactg agaagacgagc 4440
ttcgacaatt gcgctccctg ttcgtgctga ttcctccctg ttccctgctc gcctgcgcag 4500
gagaaaaatg tcgcctcagag ccagatgtgg aagctctctt tatcaagttg agaagttgagc 4560
tgctcctccg ccgccccacc aacccctagc aacccgctat agacgccgtt gcttccatcag 4620
aagaaaaagtg ggaacattga attttatcat ttatacttca ctcttttagg gacatgaat 4680
ttagagaggg cccctgtgca ttgacagcct ttacatcattt acagctgtcg aatatttgcca 4740
aatcttgaga cttctttcct catcggcccc gcggctggcc tggttgctgg gcggagacgg 4800
tggtacagc agctggttgg tccatttttgc agagaagctgg gcgacacac atcgctctctc 4860
gactggccct gtgtctcgtgc tcgaagccag agggcagtcgt ctgcacacag ttctgactag 4920
attatcatcc gggagacagt gacaaacag tgcctttcct cagtgaaggt ggggaagac 4980
tgcagccgc 4989

<210> SEQ ID NO 313
<211> LENGTH: 12515
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 313
tctacggccgg gccagagcct cggagagcg ctgccctgcc ggcagctctgg gggcaacatt 60
gacgccaggg ggcgctgcag gctccggctc gcctccgctc gcctccagct ggcagctcttt 120
aatttgcttg tgcctgccct ccacagatct ctcgctctcc gcggagagcct ccctctctct 180
actactagt tagaagagtg gcggccaagc gacgcttggct cactcactaa gagcgcgtggct 240
-continued

cgacggctcc caacttccccc tgagcctcag cacctgtctg tttggaaggg gttattgagtg 300
tgacactcct atcagccttc ctctgttgctg aaacaacat tycaaatcty aatactctga 360
gacagagcga aattatacaata attccagttc ccaatctacc acaacaaatac atgggtctgt 420
tatagtgctc ccctccacggtg cagactcgaa actatttttc atctgctccct 480
cagctgtgaa aatgaagaatg ttcgaataggg aaggaagctca aagtaattt ccagaaaatc 540
acgagtaacg gagacgcaac gctgctgctc aagatcagc ttctctctctg aacottgtga 600

gacacgctaa gatccctaga cctatctcaaa aactactctga ggaasagttct caggaatcttc 660
tcagcttacct atcagccttca ctaaggaag aagtctggctca gatgcttca aagacagtctg 720
tgacagggca aacatctaca ttcacagcctct aagacagtctg gacagctaa gcaagactgc 780
agctgatcct ctttccctggg atttttaaaga aatctcagcct cggctaaattag tgtccgatt 840
tgcgaagatg aagctgatcct ccatcagcctct aagctgctgctc aaagaagctac aaaaattcga 900
tacctttttg aagctgattga aacgctgtaa gacacagagt gtagtaataact cacaagaaga 960
aaatactctca cactattgtga gaaacacttg gttacacatt gattccgca caggaatcttc 1020
aagctgtctg ggttcactacg gggacacaca actgctgggt ctcgcttaagtt cagcacaaga 1080
atcggcgcct gacgcgtca cgtggagcctg gctggtcctg cctgacacacg acgcctgtcg 1140
gacacggagg aagggagacg acgtggagtct tgtcagcaatt ccaagcgagct ctgtgggccc 1200
cagcttctcc cctttctgcttg cagctaaacat ggacacccct gcataaatctt cacaacaca 1260
aaattctca cacaacaca ca aagacactct aactgtatcact aaccttgataa gagaactgtg 1320
gactcttgtt aamattgagag cttgcaacgc tctgtcaataa cctctctctcc caagacggt 1380
ttcacattca aatcagcact cagttatattc tcaagatggc gotacactctt cctacactgg 1440
agaaatccct ttcctttctca ccaaggggacc ttccttccat gttggtggaag tttgcctttc 1500
ggacacgtca aatcagcactt acgtctttttt aacctgttgg ctcacatcag ttcgagggas 1560
gtcacacaa gatcctctca cgaagcctgcg aatgagagttg actacacgctt gacaggtgtcg 1620
cctcttgtat cttctgtatt gttcgtttta ttcacattcag ttttgggttct caattaatga 1680
agttgggctg atcctttctgca aaagagcgttc ttcgctttttt gttggtgaccc caagacgttg 1740
actattttcat gaaaaacctgg cttcataaac gccttacaaa aggggagaag ccacacacac 1800
agaaagtctc ctggttacagc gtaaatcctcct gcgcttaaatg aacactctctc gggacagcc 1860
tcacaactca ggacacagag agtcaggtctc aaaaaatcct gttgaagattc gagcacaag 1920
cctctttttgg aagagcgttc ctcagcttctcg tggaaagatc cagggtcctca gttgatcac 1980
cgctgagcgg tcaacagagc ccctctgttg cagctgttggct ccagagcttcg cctctgttcc 2040
gagacgggga gacagagggc caaacctctgtgtg ctaagaggtg aagacatact 2100
acagagatttc ttcagatgtc tcctttttgtg gcacgctgta aagagatggtg 2160
agtcttcgca aacactctctc caggttacag cacaacatac aacactctctc gcaacactct 2220
agctctaaaa cttcagccttc aagttgtctct gtcacagcg gcaagagcgtc gttgacttcc 2280
aaagagcctcttggtgctg ctcagccttc aatgctcattg ccagacggttc ctctctctctc 2340
aaccattact gttgcctccat ctcagccttc aatgctcattg ccagacggttc ctctctctctc 2400
agctctctctctc accactttct ccacaaccacg aagttcccttc aagagaagctct ttcaggtct 2460
agcttgaagt ttcagaacgc cagttagga gcaacacagcag tgcacagcctg aatgtccttc 2520
---continued

cgcattctca aatccaga a attgcttg g aaaaacgttta ccaagaagtct gttccaga

agaactcctg cttcctcctcat gcaagcata ttggagact gggctctca gtcgacaga

atggagcaaa cagccatcag ataaaagctc tggcagccct ccccttaagac ggagaatcctt

tacagagact gggagactag tatttgcact tttgagcct catgacaaga

acaaatcca gaaattctg gatgtcctctcat aaaaagct gagaagagca caaagcagt

agaagccgaga aagggaaata agaagcctc tgggagcact atagagctcctt

cagctgaggc aaggggctg gaaagagct ggcagaagcc ggccaagacc

cagcagccctt gatgagcag tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

tggaatcagc gtcagcagct tttttctctt catgagcctt gcagtcaggt

tagagact gggagagctc ggggttggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc ggggtggtc

tggaacatc tattttgtt cagccgctcc ccaagacttg ccaagacttg ccaagacttg ccaagacttg

---
caaacacagt cacacaccca aaccacagt aagtgtgag aaaaacactt aagcatttat 4860
ggaactcga gtygcaaac agaagctag acgcagccg caagacgtg 4920
acaacactct aaggaagg ccagagctot agagaacctt ggtggtttta aagagccttt 4980
ccgacacaga gtygcaactt acgtacacct gaccacgatt cacaactcgc aagtagctgt 5040
cacatcctca cacacacgac ctaacaaaa cccacacagc tcacagctac ggctacagc 5100
atccttgggg aaatgtggcgg tgaataaagga gtctctacga gttgccacgc tcacacagc 5160
atgagcagat accacaccca cacacagcag ggaacacgga gatgtcgaag gcatgaagc 5220
atttatgag tcacctcaac acagttctgac ctcacgacca agtctcaact cgcaacagag 5280
gcaagtcgag acctctacag gaaagctgca acgttccgaa gactctggccg gctttgctcag 5340
gtctctacag acacacacag actataag tataaggttaa ttgatagttta ggaacacacg 5400
atctctacag gttcagacgc cagactctag ctaccacccca cccgacgca ggtggttaca ggaagacat 5460
caagacagct ctcgacacag cagacagcag aagtagctca ttcggtttagg aagacagcag 5520
ggcacacgca ggcagggcct tcggacaccc caacacgca ggtggttaca ggaagacat 5580
cacacgttt tggagcacag cagtgccgag aaccagaaatt taccctggcg ggaagacctt 5640
caacagcagtc ctgcttgaat ccgcgtaagc gacggtccct aaactgctgc tgaatggcct 5700
cagagcttt ttcgacacag ctaagacagc gacggtccct aaccactgta gatgtggctg 5760
aaaactacg tggagcacag cagtgccgag aaccagaaatt taccctggcg ggaagacctt 5820
agggccaaag aagacacagc aagagacagc ctgagagaaa agaaaacctg gaaatctacc 5880
actacacccg tccgagagcc aagccagctc ccgcgtccta gcgcctgtag gttggttaca 5940
agacactacc acctttggtt ggaactacgt ggaacacact gcgttcagat gaaatctacc 6000
tgacagccag acgcgccccc aaccacacag aagagcagaat gaggctctct aagatctggc 6060
tggtctcaaa ggtcttccag agacacacag tccacactgc gtaacactga gacgactcag 6120
actacacccg tccgagagcc aagccagctc ccgcgtccta gcgcctgtag gttggttaca 6180
caacacacag ctcaagatct cttcgagaga gtagactgtg aagagacagc tccacactgc 6240
cgccacagc acacagcact ccagacacac gacgactcag aagatctggc 6300
taggagacag aaccagcagt ttaagagct tcgaacactg atgtgtgacc aagagacagc 6360
tgacacttgg atggtggtgt ggcgacacag acttaagagga ggcccccatt cactacagc 6420
cctgagcggc tccaacagoc tctttccagac accgacccag aacgactgg tccacactgc 6480
tgacaacaaact acacacagg atctggtttc tccacacca ggcacaaatt ggcccccacc 6540
agggagagag aggcccccaaaaaacacctt ggggaaaagg gtagatgtg ggaagacagc 6600
agccttcgag cgcgcctcag ccgcgctggag gcggacacagc gcgcgctggag atgcacagc 6660
aggtctaccc gttcgagagg aaccacacag aacgacactg ggcgacacag caggtgtcag 6720
tgatagcagc acgcgcagcag atctggccct aaactgtaag aagatctggc 6780
tggttgggaag ggtctttccag acacagcagt atgcacagc acgcgcagcag atctggccct 6840
acgccaaaaa atgcacagcag ctcgagttgc atacagcagc acgcgcagcag atctggacacg 6900
cacaagagac gctcaacacag ggtctttccag aacgacagcact gcggagactc aacacacagt 6960
cggagacag acgcacagcag tgggagcct tgtggacacc cccacacagc cggagagctc 7020
tgacacagc acgcacagcag tgggagcct tgtggacacc cccacacagc cggagagctc 7080
-continued

ttaactggcc agcacaagat ggcccaacac ttcttaagga aaggccccag gttctagaaga 7140
ccctgctggc tccaagacct tttccagac aaccagcact gacaagcoca cgactgatga 7200
gaaactacc aaaaatctct gcacatcttc acacacagac ccagtcggaca ccccaacgaag 7260
ccacccagca cggcrocacac gaaacctcag gaaacccgac gtcgaggagga aatattttac 7320
acctagggaa cgaactcag cagcaagcag acgcatggac ccccaacac ccagcttaag 7380
tgattagaa aatacacta catttggga aactcacttg ctgaaaccttg aactgctagg 7440
aaatttaaat ggcacaacga gacagccaca gactccctag gaaagcgctg aggtctcaga 7500
gcactcagtg cgtctcaaga aaccctcccc cccacccgct cgcactccag aacactaaaa 7560
tgatgacaca atcagcagta tatactgtaa atctcacaac ccagacgtcat toacaactct 7620
aaacactcc aaccagcagc tcaagatctt cctggtgaaa ggctgcatga aagagagccc 7680
cctaagcctg atgaaagctt cagcagcttt ccgagagcat tggactgagc aacagagccc 7740
aaccagagct atgtaccagc tccaacgagt taagatgct ctaatgcagc toactgagcc 7800
aaacgcaagt gtaactgcgt gccggtgacc gtggaagac ttgataagga aagcgcctgc 7860
tcctagagcc cttggttacat tcaagagtct cttttcgcag cccccgttaca cttgagagtc 7920
aatgagatct gccaaaacct cccaaatatcc ctgcamatct cccccacccg aacatnccca 7980
cactgcccag aaggcccaca gccctccccc gacagctccc ccagagtaga taanagtgga 8040
gcttaccca ggtggtcagc tcaacagcaac atcagcagga aagcaccaca ccccaacaca 8100
aacacagacc ggtgctagagc gcaatcagaa attgagaaat cgctgaaagc aagacacaaa 8160
cccctgtaga gaggccccca gggccagagcg gcaacacgct cttctggagaa ggcgccacc 8220
cctgcgaaat cttgagcgtc tcaagagct ttcgtaacag ctcagcactac tcaacactct 8280
actgactgcg gccgacaagct cttaataacc gtcgactctt cccccnctctt aagttgtaga 8340
ccacacagca aacacacagtc gccatctcag gacagcgttg ccgagatgct aagtaaaaaa 8400
agagcttcag gtcgagcact ctcaacacac atcagggagg aacagctgggt cagacaaaaa 8460
aaccagagtt ggaagtaaag gcaacaaacg attgagggaa tttgoaacaag acagacggcc 8520
tccacagcgc agtgtaactgc gcagcggagc aacgccacgg acaattgcccc ca8580
agccataagc gcaccagctg gttcagaaag ccacagcaca gttcacoctgy aaagaataat 8640
gactgtgcg aacaaccaat cttatccttg caatcactca cccagactag aacacacgcc 8700
aacacgtcga aagacagcgg ccagacccgg tgcacccgaa gtaagagttc aaggcggctc 8760
gttagcttg gccagcagcg cacaaccagc agggagaccc aagccaccag ccaacagagc 8820
ggtggtgcag ggcacaccgg cagacatcatt taccacacct gcacagccgag aagtgacgcc 8880
aagagagatga tttgaggact ccacacacagc aagagacctt cagagaaagg ccacacccct 8940
gcacaacgct gcacagcttc aagacgtcgc tcaacaccca gccacacctg aggaagctgc 9000
aaatgcgtct ggttagatct ttaacagcgc tcaacagcaac aacccgctaca gttgaacacc 9060
tctaaaaata tccagacagag tttctcgggc cccttaagta gacaacccgg gacagcttgt 9120
aaccagcaca gacccgcttga aatacacaag ccaacaaacc acctctcctgc cccccactgc 9180
ccttcaaggg gcaggtgcagc aagatggaga attcagcagg cggcaacggg acaagacggc 9240
gccgaccccag gcagaattt tggaggaggt gcagacccagc aagacagcagc gttggtcctc 9300
ccagggagca ggcacactac ccaccccgct gccatccttg aagagaatgt tgaggacctc 9360
gtggaatatt cttgtttagta gaaatactct tttagagttca gctctaacca gaaatcttgcc 11700
tgaatgatgt cagcactctt tttcaacctg tgaatgatact ttttcaaga gcaocgcag 11760
ggtgggttttc attttcaagc gotctggatq atgggtaaaa aatgttcaatt taagggctac 11820
ccccgtggtat aatacgcgaa cacccactcc cttctgtgact ggggagggga 11880
gagacctggc aacaactgcc cttccttggct ggtgcacgtg attttagaa aataccccac 11940
cactcctcc cagctggtgq aacaacctct cagcttcag spcncatgta attcccacccg 12000
aatagactcc ttacactacr tttactcttt ctgcttcocca gggcctgtgcc 12060
gatctgtgtc tcacccactcc tctggttttt cttaacctctta aaccctcccc 12120
gtcccocactc acaccctccc aatacttcct ctgctgccctc tggcctcggt ctaagggag 12180
acctttccgc cccgctcccct cctcttgtac aatacctgcct cgtgtccgct 12240
ccttcgctc tgcggttccc tccaccccat cgtgctctgc cagctgtgq cagcttctct 12300
gttgcctgg aatacctgtq tacctcttttg tagagtagct ctggttaaaa ctaagccact 12360
aataggtggct tctggtttgg tgaaggttttt ttcacctcgag aagccgacct 12420
gatctctgg gcacggccac aacccgtgag aagaatatttg 12480
attcctcctg ttaagaactgc cccagttgac aacccgcgtg aagctccat gatacgtgtg 12540
attcctctgg cttaaaaaaa aaaaaaa aaaaaa 12515

<210> SEQ ID NO: 314
<211> LENGTH: 2444
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 314

ggcagcgggg gggcgccgggt cgcagctgag cccgctgggt ggacagcaoty tcctcctcct 60
ttcocggcgc cggcagagcgg cggcagctgtt ggcctagtgt ggcggtgggc 120
ggcagtggg cgtccggctc ggagagagga gcaagagata caacagacc caacccacca 180
tcctacgaca tgctacttcga gcacggagag cgtcgctgct ctcctgtgcc aacaagggcc 240
cctcactgct gcctcactcg caacagctgg tagagagagca cttggcggcg ggtctctgtg 300
agctgtgcgt ctgcgcggctcg cgccttcgc aacgtcctgg gaaacctgggg attccagtgt 360
tcagagaggg gcctcgttgt gcacagttctc gcctacgcag cttccagcaat atcaagcct 420
tccaccttcc tgtcgttccac gcaggggctg aactcagctct ggcgtgcttgt gcctctggct 480
tccaggtggc agtcggccaa cgcctcgctc cttcgctgtc gctcttctcc 540
atcctgcttc gcacagtgag gcocccacca cgcggagcgt ctgctgctgt gcagttgaaac 600
gaaccctcgg cgctctgctt gcgggtgatg acgtcttgct actgtggcag aagctgcaag 660
aacaggccat ctggggtgctt ttcaggggag cgcggagcgt gcctctggct 720
aacggtggct gacccgacaag gtcggcattc tgttccggac cctcctcctc gcagacccga 780
gctcggggc ccctcttggt gtctcagctgt gcctcagctgt gcgcggagcc 840
gttacctgg gcacccagcgc taccctgtcg agaacatggta ctttggtgttg accctgctag 900
aacccttcag gacactcagc acctcgctgg gctccttgcct ctttggtggag 960
cccgggcccc cccccggcctc cttccgagac tggcgttggc ctcctctcctc tcagctgttg 1020
tccctgccat gcggcgcctcc ctgtctccct gcctctgttc ctgctcagcc ctcaactcg 1080
attggcttct cacatcggtg gttctctctg gcgaagatca cccagctgag gcctctgtgc 1140
ggagctccc cactttgct gcgcggcgtc ccacggtggc gcacggggtgc 1200
cagctctgcgc acggttgcc tggctgact gcctctgat cccggctgct 1260
tgcggccag cgtccggtgc gcggcggct gcgccggtg gacgctggc 1320
tcagctgctg gcgttctgc aaggtggagc gcgttctgc acagctggc 1380
tgggtctgc tcctgctgct gcgtttctgc aaggtggagc gcgttctgc 1440
gtactagtg gcggccggtc acacggtgc gcctctgat cccggctgct 1500
cacactgtgc ggcggggtgc gcgttctgc aaggtggagc gcgttctgc 1560
ggctgctgct gcgttctgc aaggtggagc gcgttctgc acagctggc 1620
tgcggccag cgtccggtgc gcggcggct gcgccggtg gacgctggc 1680
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 1740
aacggggcc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 1800
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 1860
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 1920
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 1980
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2040
atccttggt cttcttggt gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2100
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2160
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2220
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2280
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2340
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2400
tcggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 2444

<210> SEQ ID NO: 315
<211> LENGTH: 732
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 315
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 60
aacggggcc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 120
ccagatggg gcggggtgc gcgttctgc aaggtggagc gcgttctgc 180
ttctctggt gcggggtgc gcgttctgc aaggtggagc gcgttctgc 240
gtggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 300
gtggggtgc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 360
agggggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 420
agggggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 480
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 540
tggtggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 600
agggggggc gcggggtgc gcgttctgc aaggtggagc gcgttctgc 660
ttctctggt gcggggtgc gcgttctgc aaggtggagc gcgttctgc 720
---continued---

cttccttcca gc                                  732
<210> SEQ ID NO 316
<211> LENGTH: 2422
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 316

gtgcacotccc ctctcaacgc catattgggc cacataaaaa ayygggtctg tcttttoggg 60
gtttttctt ccccttccc tgcocctcct tgcacgctgc ctgctgcaact cgaacgctggc 120
aaggttagga gacgcggctg gttgcgggga ccggcgggtct tcgacccggcc cagactcggg 180
cgggttttgg caccctcctgc gttgtctcgg ttcctctcct tcctctcgct ccctctcggc 240
agccatattgc atcgccggag actccggggc ccggccgggg ctctcccccga gcccctgcgc 300
gctctcaggg ctgctgctcg gctgctgctgc tttgctgctc gggagctgca 360
ggtgggttc cggagctgctg tggggcggcg taggggctgc tcttttcttt gcttttttgt 420
ttttgttgtgc tgtcgcagag cggcgtgtgcg agacccggtg gaaagatggc aaacgctgca 480
gtcacatcg ccggcgtcgct ccggtgcag ctgctggtgc gcggcgggcg 540
ccccggtgc gcgcacccgc ctgctgctcg tgggcagcctc aagacacatgc cggggcttg 600
gggacgccag cggccgacat gcgagcggcg agcgccgagc actgggactt cgtctttccg 660
aatccacacc cccactgcgc cagcgcagct tggacggggt tggagacggc cagccgctcgcc 720
agaacctctc acacacccgc ggcgcgcccc aaguacggtt gcagctgctc gcggcgggag 780
agcggctgct ccgcggctcg cgcgctctacat cttctgctctacatatctcctcc 840
gggacaccg cttgatgctg cccacaccgc gcacggcgtgc gcggctggctc 900
ggccgtatgc cccagagcgt gcacgagctg cttctcctac tcaaaaaccc 960
agacaccag gcacggagac cattctttcc gcagctcccg ccaatcgccgc ttctgtgagg 1020
cagccacccg aagaagctgg ctctagagac cggcagctgc aaccagctcg aatttaaag 1080
atgccccctt gtttatacgg tatacatactg ttggtgctag ccaaggaag atatacataa 1140
atatataaa tacatctcgc tcgctctctg gactggccat cctgtataa acaatgaaa 1200
caacaccaga atacactcaaa aatttaccg actccccact gctgctcctc taagaaagc 1260
gggtgtgctg ttctatcatc aggcttttcc tttttctttc cttttctact ctggtgtttaa 1320
ttttcgctgt gtaatggcg gcaaatcact cttggggaag gcgggcgagg gttggcgtg 1380
ggacatgcgc tggcggcggct gttgagagtc tttttttcct ctctgtctgtg aggtataa 1440
ttaacgtccat atggacagac gcataattgg gcagattttg ttatggttct cttaataattg 1500
gtaatccoc cccctagactc cggcctaccc gcctgttctta gcctctctct tttcataag 1560
cattcctgggc atagctcctt gtcgctctg gctccctctt cttggaggct cccactttctg 1620
ggatgattgct cagccgctcg ctcctccctc tggggttctt ttcacatgct ccccagctttc 1680
cccgtacacc accatccagg taccgactct cttccccctc tttttttttt cttctctatc 1740
attatactc aagagctcct tcgccccctt aagttctgtt atacagctct ggaataacaa 1800
gagtttgaag tgtaaaaat cggccctctg cttcttttttt tattttttgac aaaaatctct 1860
cattttctcc tcccttccgg ctcgctacac cagctcacc tattctctca ttctcctata 1920
ttttaaaga ctttgtagct actcccacagt aaaaatgaga atattaattttta atttaagc 1980
tactotgtcc atttatcacg aggaagtgct ttttttttaa ggaaggttca tggagaga
2040
agccaaactcg taagataaag gaaaygata ctatactctt aacaggttct taactgtcg
2100
tgtattggaa aacacattga agtgtcactq tgtacataac tctgtaaaaa cactgaaaaa
2160
tttctctac ttttttatgt tttttctct aacgatata caagccaaag
2220
tggccatttt tgtgctattg taatagttg gttggttga aatggtttttt ccotottttg
2280
taaataata tgtgctattt taaaaagtt gcaattgaco cagatgataat ttttttgtaa
2340
gttggaana ggtcgcaatt ataggtagac atataagta caataagagaa aacccttca
2400
gctaaaaaa aaaaanaaaa aa
2422
<210> SEQ ID NO 317
<211> LENGTH: 5061
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 317

atggctgaca ttataattgaa cagccaggtctt aacagaagtggc ctcttttcac caagccagac
60
acaagagcga aagattgtca ccaagagaa gcatgtcaga tggaagcaga ggctttagca
120
aaactgcaaa agtacctgag ctacagagag ttcttctggt ctgaagcagc
180
aacacagaaaa acacaaaaca ttttactccg ttcattgatgt ttctgttga
240
tcagatccct aaaaaaaaag ttacagatt gttgataaga acgttccacg aagtgaacct
300
gagaaactt tgcgatgtag tagttttagc actaaaaaaa cacoctattt acaggttact
360
ccttcttga gccctccttt ttcagcagac cttattttta gccactctat tcagagagga
420
cagtttcccc cgctgtcctt ttcagcagac cttattttta gccactctat tcagagagga
480
cagtttcccc cgctgtcctt ttcagcagac cttattttta gccactctat tcagagagga
540
ttacagaca ctataatttt aagtttttagc gcagatcctt cctattttttt atacotctttt
600
aacatggca acocctctcc tcocacagga aagttctctt ttttctttgc actagttactgtt
660
aatttggca ccaaactttt ttgcaaaaaa gtatgtcact cagatgacaatttttttttttttttttgg
720
aagtcagag ctatcttttg gataactgcat ctaaaactcag gcctaattca agtttctctca
780
aagtcagag ctatcttttg gataactgcat ctaaaactcag gcctaattca agtttctctca
840
tggtttaaag tggactggtt aagccacttgag gagaagagag atgtttCCAagg aagtcagag
900
aattttgaga cctttcttgag ttttgctgaa gttgagctctt gaaactctgg aagccacttgag
960
ncagagaatt aatgtctttttt gattttgc aggttctgtg aagccacttgag
1020
attggaacaa ctacaagggg caaaaogcag aagttctttt ctaaaaagaa caaactctttt
1080
actagattt cggatgcgtct cttctctgtta cagagaactttc cttgctttt ctaaaaagaa caaactctttt
1140
atgctgcttt tgttctgact cattctazed tttgaagcct aatctcctca ttaaatcctc
1200
cgcaaacaco caggtcattt gtttaagcga gtcaccgcgg aagaagacat atgctgcttt
1260
actagattt cggatgcgtct cattctazed tttgaagcct aatctcctcacaaggttttttc agtcaccgcgg aagaagacat atgctgcttt
1320
tgatgtgttg gtttactttt agaatagttc ccaataagtgc ccttggagcttgt cagaggcaagttgaggtgta
1380
 gagtagaatc aatgtggtgtc tgcagggct baseaggcaagtttggagcttgt cagaggcaagttgaggtgta
1440
cgtgcaaga atctattgtgg atgcagcagg gactttttaga aataagttcacaaggttttttc agtcaccgcgg aagaagacat atgctgcttt
1500
aagagataa gactatacact ttcgacottc agtggagtct gttcaaatgct tggcaagcaac
1560
--continued

```plaintext
agcttcacaa ggatcgggc ttcttttgtg ctgaacctctg atatggcata tgcattatat
3900
ggctgtcaca agccacacat togttttcag tgytttggg acctctgttg toagctcag
3960
aatattgataa gaagacgacg aaaaatcccc cttaacccct ccctactgtg gatcctccca
4020
ggctgtaccgg acactcagat ttataacagt ttggagtgtgc acctcagacc
4080
catactgcag acgcaagcgc tacaaatttcc tttaataggc ttatttgaact gattttgga
4140
agatctggcca caagttttaa cttcttcatt ccacaccccgt ctgagcttctg ttttcttgt
4200
tctcttcctg atgatgcc gctatcctcca ttttacacca aaacatcacc ctcttaccca
4260
gatgtgtcga ctaaagagtct cttgtagttt cacatcacg agaatacagc cccagctaaa
4320
cattatatg atgtacccgca aatgttttgtt gaagacacga ctaaacacta actttgcttc
4380
cgacacattg ttgaaattcc ggaacctcaca aaaaagcctg aaaaattttc ttcaccttgg
4440
agatccctcg ctttctcaac tggagacagt aatactcactt ttcctatttt ttcctttttg
4500
gcgcacaaag aaaaaggtta aataagcttt caacatagtt tgtgatgctg tctacacgat
4560
tgatggagatt tgtgatccctt tttgacctttt ttcaccccct ttcctctgta tggagaagct
4620
gagcggatctc atgactgctgtag gatggagcat gcttcttagtc ctatcagcag ccacataag
4680
gggtgtgtta ctaaatccat cttctcagc gcttctgctt gctcagcag cagttccac
4740
atctcaagata tgttaacttc gatggaagggt gaccccaacta catttgccat aacaaaccct
4800
citcggtagaag acacaactca atcacaagga ataacacaaaa tttcagagaa aacgaggat
4860
cgacacattca atgaagacgt gatgtagactg gatataagca aaaaaaccct aagacacga
4920
gaacacccata cactgtgtct cagctcagaa tttcctggcg cagaaattttt cttcttgggt
4980
gtacccctgt ctggtgacag ttttacccag aagaaagaga cgtttaaacct gttgtagctg
5040
actcgccagc cttgagccag a
5100
<210> SEQ ID NO: 318
<211> LENGTH: 3014
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 318

tcgtacaagc cccgctccct cccgccgcca cccagagggct ggagatcctct ccgtccagc 60
cacattcacc accaaactttct ccttctcctc gcctcatatat ccggacacccc cctctctcct 120
tctctctctct ctttcctctt gagaaggggt gggagaaaaag ggaggtcctag ttggcttacc 180
tgagctgaag ccaaaggttc cccggtctcc ccaagtgagg ggagggccgc cccttcccacgt 240
ggctggtacc ccactgttgt ctcgcacagc ccagagctcc ttcgggaggg gcaagccctct 300
gacacccctg ccttcttttt cggcctaccc ttcctctcct ccggggctac ttctctctctct 360
gcccctggaga gacagacgcc ccctcgcagca aaaaagccag cggcaagctg ctggctcgg 4200
cggtgagggc gcatatccag gggtgaaggg ccttggaggg gcaagttctg cccacaccgg 4800
tctttgagggca agacctccgg gcattgggaa ccaagttcttg gcaaccaatgg tggggtccctc 540
agctctccgg cccgctccac cccggccgac ccgggtcgcag ttcacccagc cctggtgctct 600
gtttgtccgt cacaaccccg ggctgccccct gcccaccagg ggggtgggtt 660
ccccctctgg aagccgggtct gcgggtccctt tggaagacct ccgggggcgg cggcttccag 720
taataggtgct cccgggggct gcctgccctc ctggaacagct tctctcctct cccttgagag 780
```
cctcaaatgg ttcgagggcc cacagagagc gtctccgag gcggctgcgg tggaatgta
840
ggagaggtt ggtctggtat cggagagct tcgagggct ctctcgagag gcacaactcg
900
gggtttggtt gcggcgaggg ttcgagggag acgcggcgtc gtcggcgcgg ggaggcgcgg
960
gaggggctc gcggcgcttc ccagagaaag ttcgaggttt ccaggtgccc gcggctgcct
1020
ggtgagcgcgg gcggcgaggg ttcgagggct ctgagcactga ccagttgtggg
1080
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
1140
gtggagttt gcggatgtta ccagagggg cggagagct tcgagggct ctctcgagag gcacaactcg
1200
ggagaggtt ggtctggtat cggagagct tcgagggct ctctcgagag gcacaactcg
1260
gggtttggtt gcggcgaggg ttcgagggag acgcggcgtc gtcggcgcgg ggaggcgcgg
1320
ggagaggtt ggtctggtat cggagagct tcgagggct ctctcgagag gcacaactcg
1380
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
1440
ggagaggtt ggtctggtat cggagagct tcgagggct ctctcgagag gcacaactcg
1500
gggtttggtt gcggcgaggg ttcgagggag acgcggcgtc gtcggcgcgg ggaggcgcgg
1560
gaggggctc gcggcgcttc ccagagaaag ttcgaggttt ccaggtgccc gcggctgcct
1620
ggagaggtt ggtctggtat cggagagct tcgagggct ctctcgagag gcacaactcg
1680
gggtttggtt gcggcgaggg ttcgagggag acgcggcgtc gtcggcgcgg ggaggcgcgg
1740
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
1800
caaaacgtg gcggcgaggg ttcgagggag acgcggcgtc gtcggcgcgg ggaggcgcgg
1860
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
1920
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
1980
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2040
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2100
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2160
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2220
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2280
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2340
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2400
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2460
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2520
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2580
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2640
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2700
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2760
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2820
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2880
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
2940
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
3000
ttttcattcct gttgctacct tcgggctcag ttcgaggttt ccaggtgccc gcggctgcct
3014
<210> SEQ ID NO: 319
<211> LENGTH: 2140
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 319

gctoagggt acaectccoc cggacagca aagccggcct gcaagccctg aacaocccct 60
cggacaccc caacgccctt ccaagctgtg acaagtcaca gcaacttccg ggtgggtgta 120
cactctctag gaaacaacat tttgattatt accctcagac ttggctgga ccaaagtgaact 180
gaggtaaaga ctaaccaagct tggaggctcg tggcccgca ctaagctgcg tctaaatgga 240
aoggaaaggt tttgaggggt ccactcgag cagatacatc acaatcagat gttggaccag 300
cccccgagct cttgccggcag tggagggcgaa gacgctgctg aaggtgtagg cctgctgcat 360
tggccgctg agtttgctcg caacgggagct ccctgcgcacgt ctttccctcg cagcttctga 420
cagggagac ccagccagccc tgaaggaat ggtggccgctg ctggccttcg ccctccctccc 480
tctgcggagg ctctgtgaggg caacacatct ttccctggag accttccaaag ctgctttcga 540
tgtgacttctgt gttctctcttg cccaggaggt tctcccgaggg ggttggaacat ttcagagctgt 600
ggatttacgc aagactcttc atcagcatttc ttgacgctgta tttgctgga caagccgcag 660
tctcttcacca cttcccgacgc cgaacagccg ctcagctcag acaagacacag gaaagcattga 720
tggtgagcag acgccgctcg aagcagccct cttccgctga gaggctcctg ttagctctgt 780
cctactagga ggttgctgttg atgagaattg ctcatccttc aggagaggaat tgaagggaa 840
gaaatcgtt cctgctggag cttgcaagat cttttcgtgc cttgacgctg ccctgctgga 900
tatgcagagt cttctgaaaa tgtgtaagct gcaatttatt ggaagttgtg aagctagtgg 960
taacctcgag ctctccacgt tgggcaatct ttctcttcac ggctcgcaga tgtttatctct 1020
gcgttagacgt ctcctctccc naacactctgc atcctccatc atttccacgg aagaagagaa 1080
goagcatatgc ccacgcctca cttctcagtt ccttacttgc cagtcgtctc agctatctca 1140
tgtgccctct ttatctcttc ttgaggcgct ctttgatcag tggccctgg acgtagctga 1200
cctgtgacgt cccttcacta cagaacctct ccggcttttg gaagggatcct tgtgtatctct 1260
gtocagagt cccagcgcctg gtcgtgtcat aagctgttggg gtaagctgcgct 1320
cgatgtgacct cccagccctct ttcagactct ctggctgacga gctttataacct 1380
cctgttcttc gatgagctgt gcgctacggt tgtcatgcct cttcgccctgct ctgtctctct 1440
gagccacgct tccacgctctt ctcgctcctcc tgcctctctt ctcgctctct 1500
cctagcagct ctcctctctgc acctctcggt gctgagcaat ctgacacagc tgtgtatactc 1560
tgtctccctg gaggcttatg agcactccag tgcagctctc caacgtggaga ggtttggtacta 1620
tctgctgtcc agtgctcgtg agtgcgtggt tgaagtctggg ccggccagca tgcctctggct 1680
tagtccggac ccctgggtctc aagtggtggct tctgctctgcct cagacactct cgccttcttcg 1740
gtggccgggt tctctccttc actagcggtg tggccctctc ttcagtcctctctctctct 1800
ttgccgcaact cagggcaggct tcgctgctac tgggttcaga caagccagcc acaattctcg 1860
agcatttcg tggagcagtt ggaacacat tgcctctcgag gaaaccacat tcagagaat 1920
gttcttgag ccaaaacag gcacagtggg gttcgctgacct cttgctgctt gggagatatt 1980
gttacttctgc gggggactac cttctacag ttagaagtctttactgacct cttaaagagga 2040
gattctggtc tgggaagtc atgtgaggt tatacctggt gtagactgtt gtaaagaacc 2100
ytgtaaataa aagagagagc aagtggaagc aaaaaaaaaa aaaaaaaaaa 2149

SEQ ID NO: 320
LENGTH: 540
TYPE: DNA
ORGANISM: Homo sapiens

SEQUENCE: 320
atccotgacg cgggtctgcc tttggagcaag aaggggacca atggccaccaca tggagaacaa 60
ggggatcgc goctgctgctc tgggtggccat gttggggccct gcagccctgg cggagcccaa 120
gacagagyac ctgaagctgg cccgctgtaa aagcagaaat tcgggttttct cgggtgtaac 180
gcccctccag tgtcgaacacta agggctgctg tttggagacgc acggctgcgtg ggctgcccctg 240
gtggtttwał cotcataacac tggagcctcc tccagaaagag gatggtgtaaat tttggacact 300
tcggagaagga gttcgctgtca ccctgaaaggg gttcgctttcc cggacacggt attggccaaa 360
gagcatcgct gcccctccag cggagcactc cggagcctgcct cggagcactg tgggtggcct 420
cacaacaagc atggactgtc ctggacctga ctgcctaaaa tttggctaaa aatattaaag 480
gatcagatat aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 540

SEQ ID NO: 321
LENGTH: 2346
TYPE: DNA
ORGANISM: Homo sapiens

SEQUENCE: 321
gcagagagcc cgggaggggt cgggrocctg cgggrocagaa aaggggagga cgggagttac 60
cggcgcggcg cagtcgagca cgggagctgg cggagctgtg cggagatctg tcgagctagc 120
cggtgagccgc cggagcctgg cgggagctgg cggagccccgg cggagctgtg tcgagctagc 180
gggagctgac accctgcgag aaggggagaag gttggacaag cttgcagaaa cttgcaggg 240
ttggaaacat ccagacactg tgtaacagca gggcagcagaa aatgaaagaa cttgcagaaa 300
gaggctacttc cggagagactg tsaagggggc tsgagaagga ttttcaagct cgggaggtta 360
aagggagcaaa atatatctgg aatgaaggtcg ttaaaaaggg ccataagtaa ccattaatgtga 420
agagagctga aagatagctag ctaaaattcag aatctataagct tggacactgc cggagacactc 480
cattcctcgc gctgtgtgacctt ctttttggaa ctttttggaa ctttttggaa ctttttggaa 540
gggctacttc cggagagactg aagatagctag ctaaaattcag aatctataagct 600
gggagctgac cgggaggggt cgggrocctg cgggrocagaa aaggggagga 660
cggcaacacg cgggaggggt cgggrocctg cgggrocagaa aaggggagga 720
acataagggtgc ggagagctgg cggagagctgtg cggagatctg tcgagctagc 840
cggagagctgg cggagagctgg cggagagctgg cggagagctgg cggagagctgg cggagagctgg 960
tggacactgc cgggaggggt cgggrocctg cgggrocagaa aaggggagga 1080
gggctacttc cgggaggggt cgggrocctg cgggrocagaa aaggggagga 1140
cttgtaaat ctgaagagga tgaatctcag ttgattaaca aagttaacag tcaaaaacct 1200
gtggacagcc cgatagcttc aaccttcagtt gaaatgcaca atcaagtttt ctcgggtttt 1260
aacatattgtc tctcccctgt aacctgaaat gtgaagaaag aagtcttcttc tgaacccaaa 1320
atcagctcct ctcggatgg tatgctgagc ccaagaacat ctcggagccc agtcaaaaat 1380
tctccgaggg atttctgggg aagaggtgct ttcggcagca cagcaaatcc tcagacact 1440
gtggataaca cactgaaacc aatgggctaa gcagagatgg atgagcacaat gatggyggaa 1500
gcatggaac cactttcaaat acaaacacct aacctcggcc cattacaaaa ccaagcttttt 1560
ccctagcttc cccaaaggggg agagcagctg ctcattcatt tattgcacag caactgctttt 1620
aatggaatta aggaaaaag gttggagaggg aagatgtgtg aaactctgca aagttggaaca 1680
agactccaaa tggcagttcc agagatgcaaa tcagcattaca tagaacacctt ccgacaccagg 1740
aaaatccaaa gttggatatta aaataaatct caagctgca aaaaaaaact taaagccaaa 1800
tagttattgtc gaacctctag gcacatactt tasattgatc ctcggcacat cttctcctac 1860
tctcatcag aatctctagt gtgtgactct aacaactcgac ttaactaggc aagatgtgagc 1920
tctgtaacct ctgtaaagtct atgtaaagag aaggtatcct tttcattagg caagtcacaa 1980
tgcctataaa tactgcataa ttaaggccaa ttcgctgca aagtctgctga caaacctttctc 2040
agggaaaaag ggacagcaca aacaaaagct ttcatttct gtagtggatt tgcacaaaaa 2100
aacctgcatt tttttttatt tattaatata tttttttact atagttttaag ttctgggtctct 2160
tatgtaaaac atccaaacat tttaaattcg aaattatcg ttgggtgaga aaaagccaga 2220
cactattcct tctcctctgt gtgtgacatt tttaaaaatac ctgatcattt cttggcgagc 2280
cgggtccgt cgggagggact tggggtgagc ogggttttta ttctcagtaa cccagctggcg 2340
gagcct 2346

<210> SEQ ID NO 322
<211> LENGTH: 2420
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 322
tcgggaggg ccogccgagc ogcgcgacac gttcacaatat ggaacttacag tggcgggtgct 60
cggagacga ccttcgcgcgg gcctgcgcgg gccttcaggt ttgcctccccc gcgctccccc 120
ccgccccttt cctcctctcc gcgaccggcg ctcgctgcgg ctcgccccctgt 180
cctgctggct cgctgctgcttc cctcctctcc gccttccgcac ccgcttccccc gctgtctccc 240
tcgccgaccc ctgctgcgcc ccctctctccc cgctgccgct gcgggcggct gcgggcggct 300
gatgcgcgcat gcggggagct gcggcccgcg gcggcccgcg gcggcccgcg gcggcccgcg 360
aagtgcacag tgggctttac tgcgactcgg gcgcgcgcaggt gcgcgcgcaggt gcgcgcgcaggt 420
aataacgcgc gcgtcgccat tgcgaaacag gcgttgctac gcgttgctac gcgttgctac 480
ttggtgcgac gcctgcctcc gcctgcctcc gcctgcctcc gcctgcctcc gcctgcctcc 540
ggcctccaag gcgcgacgct gcgctgcctct gcgttgggag gcggtgctgc gcgctgcctct 600
ctggcgcgac gcctgcctcc gcctgcctcc gcctgcctcc gcctgcctcc gcctgcctcc 660
ccggccgcgc gcgggaggtt gcgctgcctct gcgttgggag gcggtgctgc gcgctgcctct 720
gtcgacgcg ccggagagca ccgacacgtc ccggagacac gcgtgctcg gcgtgctcg 780
-continued

cggcaaggtgc gagcactacc gcatactgtg ccatacgcga cagctcaggca tgcaccagga  
   840
gaggtaccttt gagaacactc agcagcttgct ggagacactc acgctcagcy cgagatgyact  
   900
cgctacgcgg gctcattaca caacaagctat gagagggacca ctgctgggcc agaggatagt  
   960
ctacgcagcg ggcggcgcgc tgcacgcgaa ggcgtcagag ctgcgcagcg cacatccgaa  
  1020
ggggagttog gagaagcgtga tcggggagga ttacagggg aacaagctcg cgcctcaagt  
  1080
cattaagagcacgacggcactgc cctgctgagaa gctctcagca tgaacyaact  
  1140
ccgcacctc agctctggcgc agcgccttgcc ctcgatctct gaggagaggg ggggctcta  
  1200
catcgcacctctactatgacccaagggagactctctgcagtta ctaggctgcc  
  1260
gcctagcctg gcggagacagt gcctcctcaag ttcctgctcta acgagctgcgg aggccgtaga  
  1320
atagctgagagcgcacactag agatcagggct gcggccactag tgcaggtgcct  
  1380
tgacagcag ggctgcagactag tcagcagacta cagggcgccct ggctgggctc  
  1440
gacgacggcc agctcgcagac cagcgctggag gctggcagag aagaacactg  
  1500
tcctaatg actgtctctg gaaagttcgtg aactcttccc tgggaactct aactctttgg  
  1560
ggagagcttc atcacttcgaat ttctcctgga aagagcgtgct cctccgggtg aagagcgcct  
  1620
cagactgat gcggccgacg gttccgcagcn cgcagctcata agacatcagag aacaagctgc  
  1680
gcgcctgac gcggccgacg gcggccggtt ctaacgctctg cagagcgacg tggacactaat  
  1740
cagaacaccc gacgacgtgct cggctggcct tgcggatctag gctggcgggg  
  1800
actcgagactg gcacagctag gcctgctcct gcctgctcagag cgggaggagc ctggagactga  
  1860
gcggccggc gtcggcgcttc ctttttgctt gtcgcggcag cttcagcgccttc gcggccgcgtc  
  1920
tctctgacag ccctcgtcgg gcccctgagga ggcgaacactgc cctaccgggg gcggaggccg  
  1980
cgcggcgcgg ggcggcgcgc ccacccacgt ggcctggccct gcgtgctctc gcctgctct  
  2040
ttattgcgt ttttctctct ctcttttttt ggatttttttt tctgttggttt attttttatt  
  2100
atttttaag atgtacggaga aagaagtaac cggaaaagtc ggcttttttc gaagaagcag  
  2160
atgttatttt tattgtctct gcggtaggat ggggagacagg gcccaccctt ctagggaca  
  2220
cgtccggccag cccctccctc ccattcgggt cggagggagc gcggatacc acgggatcgg  
  2280
tcttcctcgtgc agctggctgctg actgcggccc aggccagctg cttcaggggg  
  2340
cgctgatgcc gctccggctg ccaccccccttc ctgcgatcgg gacaggtcag  
  2400
acccaccaacag ggggagag  
  2420

<210> SEQ ID NO 323
<211> LENGTH: 2253
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 323

gggaacactg gcgccttcggg ttcagaggtgg ggcgcagcgg gttggtagac ctggggaggt  
  60
aatccagttg gggagcaggga cgcggcggac aegggggcgg cggctcggagt ggggagcggt  
  120
cagactacgc gcggctggctg gctggctggct gctggctggctg gcggctggctg  
  180
ggcggcaacct gcggctggctg gcggctggctg gcggctggctg gcggctggctg  
  240
tacgctagct ggcggcggct gcggctggctg gcggctggctg gcggctggctg  
  300
tacgactcc gttgggaggct ccsacccctgg ttcggtctac cggacaactc ttcggtccag  
  360
tcctttacct gtaaatagtg gcccagctca gcgggtcttg tgtctcttca agatttcaca 420
gccgttctct tgtcaagcgc aaaaagttgt ctctgctcac aagcgcgttc agaatcagaa 480
gcaagaagca ttgcaagcag caagctgcttc ttctcttgctc tggattcaca taataacaac 540
cacacacggc aagcgcgcgc gcggctcgcgc actgtggatt catcctgagg ggagagtcgg 600
atctacacag aaaaatgaag atatataaaa gaggcaaggt gtttgggaag aaccttggaat 660
tgctggcctt cttggttaag gaaggtgttg taaagtttat tggcggaag aaaaagcaaac 720
cacagttttcct cttgcttctta aagttgattc taataacgca gttgaggaag cggagatggc 780
gctagctcga agagagagag tggaaatcata gttccccctct cgtcgatctct aatacctcg 840
aagctgtggt tattttactct atgtgatcag agtctacca taatctggaat atgcaatcctat 900
tggaacagct ttaagacacac ttctggattc ttaaagtttg aataagccaga gaaagtcaac 960
tratatatacga gaatagccac acggcctcgca tcctctctct tggagaaag gatattcagat 1020
agacattaag cccagacatc tacttcttgg atagcgatg aaggctaaagg tttgcagatat 1080
ttggtgctcga tgcctgtgcta ccccttccag gaggacacgt cctctgcgca cctgcgacga 1140
ccctggccctt cagaggttgg aagctggctag gcctgatcag aagtgggact ttcgagcagt 1200
ctgattcctt ccggctcatg ttttttaggg ggcgcacctct tttgagccaa ccaacactaca 1260
agagacccact cagagagatt cccgggtgaa atcctacact cctctctctg taacagaggg 1320
agccggccac cccttcttccag gacgctgca gaaatcccct aagcgacgcg caatgctcag 1380
agagatcctt gacacacagct ggtcacagc aaacctcatc aaccaaacca aatatgcaaa 1440
cacagtcagc gttcagccag aataagccaga aataagctat caatgcagcg 1500
ggctacatc taccagcaca ggaacatcgct acgtgaagtt attttaaat tgaatgctgc 1560
cctcacttctc gacagatcaca caagaaatatt tttttactct cagccaggttt gctcttcag 1620
cctcttttac gaaatagccac atcaataata aatagacact agaagttgag agcagcagag 1680
aagcgtcttg cttctatttct ttcagttgctt cggccgctgc ggcgcggcctt 1740
cagctgtgcag aagcaatgctt gctctctcttc gaggacacgg aacagcctgct gtctggggga 1800
agtgcacctc cttcgccctcc cccgagctgt tgcagtcgct ttctttctgc 1860
tgctgagttct ggtcaatctt ttcctggctc gcggctgtgtg aagctgggtt gaaagctgctt 1920
tgagtttttt ttctgtgtga aatttagatc atatgcctac ccgctgtatc ttctcctgtga 1980
ggcttttcctt ggaattgtgg tctgtttgct cccgaacccc ggtggccctg attcggttcc 2040
tgcttgctcc taccacacta tcctctctcc gaaagtgtag aaatatggaa cagctgcttc 2100
acccatatttt aagggatttc ctgctgatac gcagagcccc tgcgtctcgag acgtgtaaag 2160
gggctatttttt tttatatatt tggagctcata gcagctgtgtg aacctttataa tattgaaata 2220
aatatgttct tgatctttaaaa aaaaaaaaaa aaaa 2253

<210> SEQ ID NO 324
<211> LENGTH: 1619
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 324
cgcgacagtt tgtagttcgg gacccgctggg cagagctgcgc gcgggcgccca tgggtgcgcc 60
gacggtgcgg ccctgctgtag acgccttttc caaggaacag gcgtctctt catccaaag 120
ctggcccttc tggagggct cgctgctgac ccgggacgag atggcgcag gcctgccctc 180
cacaatcctg aacgggacct acgcagacct gcggccagtt ttcttctgt ttcaaggaat 240
ggaaggtggt gacgcagatg acgcccaaat agaaggaat ctaaagcaatt ctgcagggctg 300
cgccttcct cctgcacagc agcagctttta aacatcaccct ccgggtgact ttttcgtacc 360
gggcgagaag acggccagcc ccacaaattac ggggacacac aagacatgac agaaggaat 420
tgagagattt gcgagagacg tgcgcgctgac ccatcgcagc atcggctgac aagctgaggg 480
cctcctgcag gcagcggctg gtcgacagct gcgctgcata gcttccaggt ttatcctctg 540
gtcgacccag cctccctggt gcgcctttta cagatctcga ggaaggaag atacaaatcttt 600
caaatgattt gtctcctgct tccgcctctt tgcggatgta agtgggacacc caggtgcctt 660
tgcggatgta gcgggctctc agtggatctc atcgggtcctc cccgctcctc cccccctctc 720
ggcggcctt gcgggtcttg ttcggattac gcgtcctcag cggactttag acggggacac 780
aggggatatt ccccttttctg ggcggcgcacc cggccgcttc ctggtgcttc ccccctttttt 840
ggcggcctt gcgggtcttg ttcggattac gcgtcctcag cggactttag acggggacac 900
ggcggcctt gcgggtcttg ttcggattac gcgtcctcag cggactttag acggggacac 960
acaagtattt tgcgtccttg ttctctctct ttttttttttt tgcgtccttg tgcgtccttg 1020
gtcgagtgct aagcgagcacc gcgttccttg ctctctactc gtttcacacc atgctccctct 1080
ttttttttct gtcagccttt ctcactcgac atccagccc gcaaactaaca atccagccc 1140
agcattttttaa tgcattcggct gaaacaggggt gagttctcag tcggagagaa ggcagaatag 1200
agcattttttaa tgcattcggct gaaacaggggt gagttctcag tcggagagaa ggcagaatag 1260
agaagtgtggtt gcgggtcctt gcgggtcttg ttcggattac gcgtcctcag cggactttag 1320
ggtgttcttttt gcgggtcctt gcgggtcttg ttcggattac gcgtcctcag cggactttag 1380
tgcgtccttgc ccacacaacc atcctgtacag ttccactac gccgggagag agcgcttcgg 1440
cgcgtccttgc ccacacaacc atcctgtacag ttccactac gccgggagag agcgcttcgg 1500
agtttatt tgcgagtttc ctcctctttg acgggtctag gcctgtcctct ggcggatcatt 1560
gtcgagtttc ctcctctttg acgggtctag gcctgtcctct ggcggatcatt 1619

<210> SEQ ID NO: 325
<211> LENGTH: 5010
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 325
ggcggctcgg gcagggaggc gcgttagtg gctgctgggcc ggctgcagct gcggccgact 60
tgcgagtttc ctcctctttg gcgggtcctt gcgggtcttg ttcggattac gcgtcctcag 120
ggcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 180
tgcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 240
ggcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 300
tgcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 360
cgcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 420
cgcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 480
tgcgttgggcc gcgggaggtc gcgggagggc gctgcagcct gcgggacgag gcgggtcctt 540
---continued---
aaccaaaaac tgcggctgac agacgaggcc gagacgcgct tcacagtgac ggaggccgag  600
gagaagactt cgccgggccg cgtgccctat atgggagaga ctggaggagag aaggtgctgg  660
agaaaaagctga gaccaagacag ttcagcacag ccataaagct tcggagaaag aatacaatag  720
tcagccttg cggctggtct cccacagttg aaacttcgct cggtagaggc gaaaactcaat  780	ttgctagaatt taaactgagc aaagctggtgg tgggtacaca tttggtaaa atccagttca  840
aacacagcgcg cacaaactcag gtagacatag ttaaagagaa cgcagcagct tttaccttg  900	tggaggatcc tggaggcttac tgcgcgctta gtaacgctgac aacaattact gcctaaaactg  960
tcctgctat tttgctctat aaaaagaatt tggagattat tcatacctct gtaagttgat  1020
tcatactgat tgcacagctg ggcggaaact cccttgcaga aagtgtagca aactgtgtaa  1080
gttaaatagc aatggctggt tttgatattc tgggccccagcg tcaccccaagc acaactccag  1140
cagacatcttc acctctggctc acctgctgct acctgctgct acctgtgtgct  1200
tcccctttgtc caattccacact gattttcccac catcctgtgt atccagtttgt ctctatatcg  1260
tgtgctagcc aacaactagct gtcgctgagc aacagctttgt tgggaatagt gagaagagact  1320
tgcctgtgat cgggaaaaaca gacttatcat tagatgttgt aaccccgagaa agcagagatg  1380
tgagctgcttc tggtagaagct gataaaaaat ttctactact tttgaggcttgc  1440	ttaagagctt tgcacagcgcg cgggagtttgc tttgctggtat tggggtgggg  1500
ggcctgagaag cccttccctct cgttacatt ctccagcccg ggcagcagct  1560	tccagattgc gttcttttac ggagttgggt agccagcagcag aacatttttt cttgcaagt  1620
ggagctgttg agagctttggct tcgggtgttg cccctctgat gtaggggtag cactcttctgg  1680
cctggatattta acaagtttctt actatatata aatcggtata agcggccttt cgcagccgc  1740
acttctcagtt ctctctgctg ccctctgttt atccagtttt tggacaacaac atcgacatagt  1800
tgagacatag ggcagctggg catttttatc acctgagcg caaagctgggc agaagatttc  1860
agaaaccttt cttgcaagccc ggtgggttc cttcctctct tttctctcagc ctccaccaag  1920	ttttctctttg cagcagcatt atccatatgt ggtgacaccc aggcaacactc  1980
ataaggactg gtagagagcc actttgctgt agctgatgct gtcgacaact gtcgtagggg  2040
tcggctgtgatcc ttggctgattt aaaaaccccg atgggcttga agtggacct gacatagaga  2100
gggactcaac ccacagttt ttcctttgtga gggacttgg gcacactgcag cgcagacaag  2160
agaaaaagttg ctgagcttga cagttgagct gttctacttt gcgcctcttgc gcgcctctttc  2220
ttcgcaacgt aacacagatt cctcctccag ctccgctac tccactctcgct  2280
aatgcaaaacg tgcgacatag aaggggagTT atcactccat tccctcac tccatcaca  2340
aatgcaaaacg tgcgagtcag aaggggagTT atcactccat tccctcac tccatcaca  2400
tgacagactg gatatacttc gcggatggga ttcgcatccag ccacagcagc  2460
aacagcctgt gcagattgct gttctacattt gtcgctgttc ctcgtggttgc  2520
tttctgactt gtcagcttgg ctttacagat gttccacatttt tgcgctgat ccgctgctg  2580
gagatttttt gatttccattt tcaactttct atcaacactc aacagcagc  2640
taatgcaaaacg tgcgacatag aaggggagTT atcactccat tccctcac tccatcaca  2700
aatgcaaaacg tgcgagtcag aaggggagTT atcactccat tccctcac tccatcaca  2760
tgaggaatata tccctctctct ttcgtgattt ttgcagggcc gctggttcttt  2820
-continued

cctgtgttc ttaatgaanaa tgcctgacac cagttatcctg aatgatcccc tctgaactca ta 2880
aggyctgyc tctgycyag gttgaagtyg gttcctaatc tttgagtctt ctctoaactt a 2940
catttagtc taatcgcggt atacaacggtt gaaagacctc tooaatcgag atctaaagctt 3000
tttccatccct agcttacgac gttctcccat tctctgacag acacgtaaac tttcgacaga 3060
gatgagcttg ttccttgcct atagagctct gaaagtggag tcctttctgt ggtatattgct 3120
agttaaacat gtgtgtccca ggaataagcc ccctaaatgt taactctcaac gttacttatg 3180
aagaaggggg aacgaaagccg aacagttgag tatattttac ttttcctcgt tccttcccccc 3240
ataacccc tattattcct ttgatttttt tggattttcc acacgcacgt gaaagacgca 3300
cagtttaacc ggtagctgag cagagcctgt ctctgactct ttaagaaaat atatgctgagc 3360
aattttgccc caaggtgtaa atcctagggg agaccccttc gttccttttgg cactgatata 3420
tttgcgtt atattacagc gaacaggtcc actataasag ggtgttmtttt atataatagct 3480
aattctggag aagagctgctt toccataata tgaacgattat acgtgctggt ttttttattt 3540
aagaaggaca ttccctataa aacccgacca gttactgagaa gttggtccatt tacggtcagc 3600
actataagtt ttgagaaaaa aagcgctcag ccacgtacta tgaataaaggt tgaagctcag 3660
atttagaat gaaatattat ttaaggtggc ttgcttaagaa gcccagccct accagtaaac 3720
aagtctggtg ttccctattc caggtgatt tgcaccaagct atatcactt gctacccatn 3780
agygttgtct gtccttcgctg aaamtttattt aacgccagtt atccagaaggt taaatcctaat 3840
tccattggc aaaaaatcga ctaaatttt ctaaaatct acacattttt cggagcagt 3900
ctttccctaa tgaagaaaaa aagggaggttttt ttttactac cccgctgctt gtagggga 3960
caggtatact cattgacttc actaaggggt taagtaatta togggaacacg tgtttcccat 4020
aattttcttt ataacaaccct ctcttcaaaag tctgaagttgct cttcagtttt tacatgtact 4080
ccacacctta taattccctt tttttatttt ttagatcagc aacacttttttg tggcctttaa 4140
goctgggttg gcctttaccc aaccagtctaa aaatgaatctta cttcacaattt tgaatatttg 4200
cttgcttttt tggtaacttt atgggctcct caggtccttt actaattagt tagaagctatt 4260
acattatat cttgctttt tgcacagcttcttt cttacaatttt ttacactatt ctagctgtac 4320
tctctacatt tttagcgagg acaagttttta atcaagaaat gcocccctttg caaataagct 4380
aggyctgtct cttggaaccgt tgggaagcag gcctccctgtt aaggtctgcc 4440
aggtctgtct aacccctcgt ggctggccat taaatcgcag ttagagctctg ttcgagcctc 4500
ctgaggttaa ccaatcaataa acagcttgatt gcctagggcg agtgggaacc 4560
gggtctctag gaaaggtctc ataggtgctt tttgagttct tagtggctcc ttcctagatt 4620
tacagccct caattgcggtct ctcctctgcag cggtccttcttct tagsctatttttca 4680
atctctcctt ttatttggag ctgacatctagc gctagctaaa gttgaagaaat gtctagtcag 4740
atctctccta ggctgggaac gccgctccat ttaatgcgac gcactgtggga agtgcctcggt 4800
staggcctt atagctgctt tcgagcagt gtcctcctt gctgagggcg agaaccacct 4860
ctgtgcctaa attcgctcct ccggtctcatt tggcgacagt cggctcctcc cggaaaccctt 4920
ctctgactt aggtcgagc tttgtctcct gttgtactcct ccctcggcta ctgygtacga 4980
gagtagggc cttttgagg gaatgatgtta 5010
<210> SEQ ID NO: 326
<211> LENGTH: 2574
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 326
cctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 60
cacctggttc gctcagctt ctgcagctt ctaagccag tgcagctg ccagttcag 120
cacccctttg ccacagctcca gctctttgctg tggctagctg ccctgtgctag 180
cctttctgcac ttctttctgc cccttgtttgc ctcctgtttgc ctctttctggct 240
tccagctgta cacctgacca tcgacagctg acatcttgctg tgggtcttgc acagatc 300
tttggacac gctgtgctgt gctggctgctg tggagctgctg tggagctgctg gctgctgctg 360
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 420
cctcctgcac gctgctgctg tggctagctg ccctgtgctag 660
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 660
gcttctgcac ttctttctgc cccttgtttgc ctcctgtttgc ctctttctggct 720
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 780
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 840
tgggtcttgc acatcttgctg tgggtcttgc acagatc 900
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 960
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 1020
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 1080
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 1140
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 1200
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 1260
tgggtcttgc acatcttgctg tgggtcttgc acagatc 1320
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 1380
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 1440
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 1500
tgggtcttgc acatcttgctg tgggtcttgc acagatc 1560
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 1620
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 1680
gagacggttc ggctgggttc ttgagacggttc gctgctgctg acatcttgctg tgggtcttgc 1740
tgggtcttgc acatcttgctg tgggtcttgc acagatc 1800
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 1860
gcctggttac cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 1920
tgggtcttgc acatcttgctg tgggtcttgc acagatc 1980
tttggacac gctgtgctgt gctggctgctg tggagctgctg gctgctgctg 2040
cctttgaca cacctgacca cacatccag ccgtacacag ccagacctcc gcttctcggt 2100
ggctcatgct ttcttggtcc atcaactgta ttgggcttt ttgtagatcct cagoagcngen 2160
aeesgggtg aatcaacacc ttctctcgtc gctctctcttg cctctctctc caactctgcc 2220
tcgtctcatc ttctcttgcg acaacttgtct agagcccttc caggctagga tgcacatctgg 2280
tgctttgtcc tttcctgctc ccttgcccgct ctggcctctt ccccccacac 2340
acccagagct cttagtgtgt cctgtaggtg ccaggtccaa caacttgtgg caggcagagc 2400
agctgcatac gttgccacca aagcagtctg gctagcactc cagaagcaca tagaagcagc tataaagac 2460
agctgagata ctctctctac aatctgtttg taataataa tttttgagcc atctctggat 2520
attctctcttgctgaaatc ccagtagagc cttactccaa ggtt 2574

<210> SEQ ID NO: 327
<211> LENGTH: 1421
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 327
actactagc ggacagcctc ggagactact cagggctgtg aacctcccgct cggggccagt 60
agctgcttata acctgcccaag tcggctgctc ccgctccaca ccgagacccg cggggaagttc 120
cagatgtgct ggccgctgtc ctggccagga aacacccccg aggttgccag aagctcctcgct 180
cggtcctcacag tgcgtcagta cagctgcctgt gctgcaagtc atgcccacag cactgctcac 240
agcgacgact ctgtgctcagc tggacgataag cagctgaggg ggctgccccg cttcctgcc 300
cgagaattgg ccacagaggcc cttgggctgt ggtgtcctag gcatacgaca aggggcttgt 360
ctcctgctace tctggtggct cctgctgctc ggcccaacgc cggggcagcc ctgtaattgct 420
gctgcaacttg atgggcacct ccagaggaag cctttggggcc oacatctgta aaccaggggctc 480
cctcgccagc gcttggtgact gctgacgctgc ggtgtcctcg gagaagccgac 540
tataacagca gcgtgctccca aagagggagt gtggggtagt gttgggagga aacacagttcg 600
caactccgtgt gtcgccgctg tattctaacc gacgctccag gcagacgtcgc cggggcagcc 660
aacacagaga cactgcaccag tcggccccag ccagagacgg agctgctgctc cgcaggtacc 720
tttgcccccc aggccgagtt cgaatgttac cctgccacag cgggacgctc caagggccgc 780
cagctgctctg ctcgctctct cgggagctact gcgtgctgctc cgccagacct 840
caggaggaag tggagggcgc tggaggtcga cccacacttg gcctctttcg ggacactctgct 900
tctgtgctgc cccacactgc cgcctgctgc cttctctctc aaccactgtc cgcctaagag 960
tctgcctctc agcttgtagg aagaggcccc aggggcaggg tggggtcttg ggtgtgccttc 1020
ggatctgcca cctccctctc cttgggctgt aagggagactg ccagagccag tggagctgcag 1080
cctgtctctg ctcgctctcg ggtgcttcgt ctcagtctgctt cttccctctg gaagaagctg 1140
ggcacactt ctgagctttg gcacactgtc aggctgaggc ctctctctgc ggatggggtc 1200
cgcacactcc cctgcgagct gctctgatat agcgcctcct ctgttctctcc ggctccaaag 1260
cctctctccct ctcctgtgat gctttccca gcacgacagc cttctctccag caagagcttgctc 1320
ggcacactcc ctcctggggc ttcgctccca ggggtgatct tttctaaagct ggyagaggcc 1380
agcggtccac cttgctgtct cgtcttatgaa attactcact t 1421
-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQEUENCE: 328

ggacacagcg tgtcaccgcg cgccggccag aagacgtttg ggtccctgag tgtctgtgag 60
cacctcaga agacaccccccg ggctccac gcctcctgctt cctgtctgtgga 120
ggcgaatggcg ggcagagcctt ccctgctgctg cctgagagcg ttttagccgag atgctggcgg 180
ggttgctttg ggtgctgtgcc ggacagccac ctcctgtttg ccctgtttg 240
gacgttgttg atgcattttt aagggagacc ggttttgaggt gtttttttttt gttttttttttt 300
caccccagcg ccagcagcag ggggggggg ggggggggg ggggggggg ggggggggg gg 360
agctgttgag agcagtttcc caaagctgctg cccctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
aagcagctct ggcacactcg ttttcccctc ccactgcctt aactctctgc ccagggagagc 2220
cagccctccc ttttcccttc gcattcagcc gggggcgttc tgcctcctgc ccccacccc 2280
ttcacctgag ggggggctag gcgtctccct tttacccagcc ccagggagagc 2340
gatgagggg aagttcctag gcctttcact gcctggagggc 2400
tcgagggg aagttcctag gcctttcact gcctggagggc 2460
gttgggagag ggcacactcg ttttcccctc ccactgcctt aactctctgc ccagggagagc 2520
gttgagggg aagttcctag gcctttcact gcctggagggc 2580
tcgagggg aagttcctag gcctttcact gcctggagggc 2640
coccaacttgc ggctacttgc tttttttttt gtatccagct cccatgttgc agatgtctgc 2700
gttgagggg aagttcctag gcctttcact gcctggagggc 2760
gatgagggg aagttcctag gcctttcact gcctggagggc 2820
gttgagggg aagttcctag gcctttcact gcctggagggc 2880
coccaacttgc ggctacttgc tttttttttt gtatccagct cccatgttgc agatgtctgc 2940
gttgagggg aagttcctag gcctttcact gcctggagggc 3000
gatgagggg aagttcctag gcctttcact gcctggagggc 3060
gttgagggg aagttcctag gcctttcact gcctggagggc 3120
gttgagggg aagttcctag gcctttcact gcctggagggc 3180
gttgagggg aagttcctag gcctttcact gcctggagggc 3240
gttgagggg aagttcctag gcctttcact gcctggagggc 3300
gttgagggg aagttcctag gcctttcact gcctggagggc 3360
gttgagggg aagttcctag gcctttcact gcctggagggc 3420
gttgagggg aagttcctag gcctttcact gcctggagggc 3480
gttgagggg aagttcctag gcctttcact gcctggagggc 3540
gttgagggg aagttcctag gcctttcact gcctggagggc 3600
gttgagggg aagttcctag gcctttcact gcctggagggc 3660
gttgagggg aagttcctag gcctttcact gcctggagggc 3720
gttgagggg aagttcctag gcctttcact gcctggagggc 3780
gttgagggg aagttcctag gcctttcact gcctggagggc 3840
gttgagggg aagttcctag gcctttcact gcctggagggc 3900
gttgagggg aagttcctag gcctttcact gcctggagggc 3960
gttgagggg aagttcctag gcctttcact gcctggagggc 4020
gttgagggg aagttcctag gcctttcact gcctggagggc 4080
gttgagggg aagttcctag gcctttcact gcctggagggc 4140
gttgagggg aagttcctag gcctttcact gcctggagggc 4200
gttgagggg aagttcctag gcctttcact gcctggagggc 4260
gttgagggg aagttcctag gcctttcact gcctggagggc 4320
gttgagggg aagttcctag gcctttcact gcctggagggc 4380
gttgagggg aagttcctag gcctttcact gcctggagggc 4440
aacagtgcct gctgaatgat ttccaaagag aaaaaaagtt tgcagaatg ttgctcaagtc
aacacaagta gccagctttg cttatgtaaa taaaaaaggg gcctactttat atagacat
ctttcttgcc aagtactgct gtaaaataat gctttatgca aacc

<210> SEQ ID NO: 329
<211> LENGTH: 2076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 329

cgggagaagg gaggagagag ggggagagag gagtggaggg gttgaggg gctgaacatc  60
gcgggggttt ctggtgtccc cccgccccgc tcccacaata gttacaacgca ccggagacgc  120
ggaggcgtcc tccctgctgt ttgctttaccc tcggggcttc cgaatccgggg ggagccggtt
200
gctgggcttt cttacatctg aaccggttcc tttcgggggt cttacggtgg actgggcttgg  240
aggggcgtcc gcaagaatgg gagcggccag cccgggcacc cctccgcccgc cctcgggttc  300
gccgcggggg cttcggccgg aggaggtggg cgggagaagg gcattagggg cccggggcct  360
cgggaggcgg gcgggagcgc cagcctgtgc cccgacgagg gacccggctcc ccaccaccgg  420
tcctcccaccc atgcctcctgt gcgggtctct ctgtctggct tcgctctgcc gttgctgcgc  480
gctggccgg ggctggcgag cgggacgggc cgggacgggc gctgtcaggt gggcggtcgg  540
ccttcgggc gcgggcgtgg ccgggcgtgg gcgggcgtgg gcccggtcgg ctgctctgcg  600
gggaaggttt cccgggagtt ccaatcgtag tggactctgct atctgtctct aaccaagaata  660

500

tgggaaaaat tcaacaggtgc acgtaagggc caagagcttg ccacatattc gacaacgac  720
cacccgaccg toccataagg aagagactat aaatgttgt ggcgacaatc attaactcaca  780
gatcctggag aattgagtatt atgaaggtgag aasaactcag tgcgtgcccc gggaggtgtg  840
ttaagagtgg gggagaggtg tggagtgcag gcacacacac ttttttaaac tccgattgtg  900
gtctcgcttg agataggggg gttgctccag taggtgggg ctgctccgca gcacacaccag  960
cagagacag aagccagag aagttatttg aatataaatgc cctcttctt caccggccaa 1020
aaccggttgg atcagagcc gttcatttg ttttgcgcgt gctacccgg gaagccagac 1080

ttcagacaa gttcattgg cttttgaaag gttcattggc gcacacaccg caaagtgccaa 1140

ggcacacag gaaagctggcc ccacacattt cctgggacatt aatccacatct gcagatgtgct 1200
ggcgtggg aagtttaagtt cttttcatgg ttgaagcgct gaacacagat cttggttccaa 1260
tggagcatttg gcagggcttg ggacaggtgg tggagccttg tctggacagcg 1320

ggggcgtttcc gcggggccttc gcgcaccagt gcagcagacta gcagacagact ctcggaggag 1380

ttgctcattaa aaccacacct cttccagcagc atgctgggccc aacaagctat tttatgaaag 1440

cacaacactg tggtaatgta aaagacactg ccggagacaa atacccatct ctoctcgaag 1500

ggcgcggcct gcttctcgtc gacagcctgg ttcctcgtc cattcgggac cagctggtttc 1560
cacccacccgc ctcagcgcct cagcgccttg gtcgctcgt gaactggcct gtttcgggctc 1620

gggtcgttg catctggcttc ggtgcggcgc ggggtttcag cgggttgttg aagctctggc 1680

tgactgtaaa tttatgtaaa atctcagactc tttcagttgg ttttctttgc aagatgctgct 1740

gcccagctac gacggtcgcttg ggcggagaga accgctttcg gcctctgttac caaccacag 1800

aagctcttgc ttcgctgcat atgctgcatc ctccatcctaa aaragacgag aacgtctgg 1860
<210> SEQ ID NO: 330
<211> LENGTH: 2619
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 330

tgggccccag ccccccagct gccgggtcag actgtgctgg gccggtcgct gcgccgggcc gctggtggtg 60
tgggagcaag cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 120
agccagcagc cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 180
agccagcagc cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 240
tgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 300
tgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 360
tgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 420
tgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 480
tgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 540
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 600
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 660
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 720
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 780
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 840
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 900
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 960
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1020
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1080
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1140
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1200
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1260
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1320
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1380
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1440
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1500
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1560
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1620
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1680
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1740
ntgggagccca cccggctgct gctggcagtc ggccggtctg ccccgggcag gccggtcgcag gctggtggtg 1800
tccacccctac actgtgaagg gtacagatta ggtttgccc acgcaaaaat aaaaattgtat
310
aacaatctct ttttgtgga aacaacccca gtaataactc cagagacagg gaaaagctcag
1920
ccggttccag aagaccaaac tgaactctca actgaacag cagctgactg gcaggccttt
1980
ggacggttgg ccacgtccct cctgaatctt tcttcttctt gcttcttggg gtttctagga
2040
atgytgaaag cctctggtgct gcctggtcgt gcacgctgaga tgggtgccc ttttctagag
2100
ttaacottta cagagcagtgg agagacccag cagccagaacc tyccagaacc actgaactga
2160
ccaaagcagg ttccagtggag ggttgtggtc acaccaacgaa atgytgtgcc ttttttctag
2220
cctggggtga cctctagacg tggaggtgtg ggacacccag gggcccccgt gtggagacaa
2280
catcttagc agacaccaaac ttaacagcct ttggcttctg tggacaaattt ggcagtttctct
2340
tggtatggag atggagttt taccagttt tgaacccaga atataagtttta ataaagttttt
2400
aaagotgacag agtttgtgag taaaagaata tataattcctg ttgtatattt cttttttatttt gggcctttttc
2460
tggtgacagc atctggctctg tggcttccgt ctttattagtt cactgtacct tcaacgcagac
2520
gtggaatgt toccattatt attttgttct tcaacagttgt gaaactggaag cctcatagagg
2580
tggaacact caccagacag ccccggtttg gtcggaggg aaactttagg ttccattagaa
2640
atgtgctcgt ctatttattcc cattttttccc cccggtttttt gagacccaca atgtgctcag
2700
aatagccaa cataaacata atgygtctgg agttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ctgtagaac agaaggacga agaatataa caaatatttc aatgtagctc ttcaccttggag 1080
gagtcttgg taagactac tgagaacaca caatcaacct caaggaatg tggtagtaacc 1140
aatcttcaca atctatgtgt caaagaagct atacgaaagtacctt tctttagcatt 1200
agaagaataa tggagaaaaa aactcgtaga tctgggagc atctacaaac acttagggg 1260
cctggtcgcc atctagctaa tgcctgaaaa gacagatagc aagatgagtc aagtacagct 1320
tcatacaga aagagattag taactaagag aagtaagag gctctggcag aaggaagctc 1380
tgcaaaatct gtatggaag aataattagc atcttctttgg tcctttgagc acatcagtc 1440
actgtaaac aatgtagctg aacagttacag aagttcttccaa tggctgacatg agtcttacta 1500
tgcaaaaaga aatattttgc gttcttactc aoctcctag taggcatatg tagtttgtct 1560	tatttacctg atgtagttctg tgtattcaga tgcatttaag taactgagat tgaattcctt 1620
tagcatgctg taacatagta gaaaaaaga gtcattggaca gtgtttagtg tgggaatat 1680
atottttaga tttcggatttt ttcgggtatt taggtgttat attcatttatt ttattccttga 1740
tttaattgga accaaagacg aagataaagga cgaatactac tataactgaa cacaaatgtg 1800
attcattagga tattgatttt atttcaagtta taatgctag aatacactctg gattttttat 1860
ttcattaag taacatagctc aattgagcct ttcctgttat aatgtagtg tattagttga 1920
atcctccaa taacatattt tgtttttgtt gaanaagga taataagttct cttgctggtg 1980
ggaagagata gattggtttgtt aaggggggttt tgtggtttttt tagggctttct tagctttttc 2040
tgtaaagggg taacacgagga cgtgtagcagga ataattggta aagatagtact cagatgtg 2100
aaaaggtttaattgattcagtaa aacaacctgt aatgctagag aagaaggttca 2160
gagagctgt aagttgatagagagtaagtggtc ggagacactat gtagttctg agcagagaca 2220
gaatagttcattagcttctatta cacaaaccct cacaaaccct cacaaaccct cacaaaccct 2280
tttaattgtg ttctctctcgg ggaggaggggg gatttggggg ggggccccag agggtttaata 2340
gagggcccctttt tccctttgg ggaggtttaac atttttttcttt gaggatattttt ttataatttct 2400
gtagagcagcg agggttttaa tgggaactaa cattagtaac ataaccctgg tgtctatctt 2460
tgctctttgc tagggagcct tgtggctttcc cacaaacaccct ccccccccctg gaaacacgctg 2520
cgtttagctg gggcaatttgg 2540

<210> SEQ ID NO: 332
<211> LENGTH: 1474
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 332

aasagaat caagaagctca attttttta catagtcac gocggaata acatgaaaata 60
aatttacg aaggttgaag agaatctcac aagtagaggtt caatcgtctg aagggagcgc 120
gagctgaccg cacagcgagc gggcgcagcc caaccgcggc gcggccctag tagctcacc 180
aaccggaga gocggcagctg gocgcagccg gccaccgctga caaccagcgc ccaccagctg 240
agggagggc cagcagcagca gocgcagccg gcccacaaggc cgcaccaggc ctcagaggcc 300
gccacaccg agccggcca tacgcagagc gcagcagggag cgttggcggcg gcggctcacc 360
atcgtggcgc ctgagctggcg cagcacaagc gctctggcag cgtatgcatg tgggaacgta 420
aataggtgta acgttaagga cgatagctgt tcctcaccaca gtaatgacac aagagaagat 480
gtatgtgcc acaagacttg ctaaaagaa aataaactca ggaagtaact tgcgaaggtta 540
ggagatggaag agoctgygga gttgaggtg gttgaggg gaaaggtgc gggaggaaga 600
aatgttacag gtctggtgag gttgtcagtt ctaacggata aatagtcaag agacgctaac 660
cattactgc gtcggactg ctgcgaagtg ctcacaagca aataacgcct aaatattccag 720
aatgtgaga gtggggaaga gaaogagggga tcggagcggt ctcggaagc cagggcagaa 780
aacgagcgcg ctacggaggc gaaggttccc aacactatac atggagacac ctatggyggt 840
cggacacagt atccacaccc tcctggtcag gggagagttg gggaggtgag ggacacagag 900
ggtgcaggagt acaacagtg agatgtgagg cagtatgtgta tcggagccat agacagccag 960
tcggaggg ggctctctcg caaaaaagc acctgaggg cagccatagga aagacaaaaa 1020
gaaatcagaa gggagtgcag cccaggtcag cagcctgcc agatcggcta cggcgaacac 1080
ttcattacag gaacgagaca cccagaaaaa cttaaacca acatggcagaa aagacaaaaa 1140
goagcgacaa cccactgctg ggacggagag gcgtcgggag ggggttctacag 1200
taactgcgg gtttcacttc ctacccatct cccgttttatt ctcacaacac ggaagataat 1260
gaaatccag ctaaaagaaa tgaaccaaaac atgggagctg aagacccaaa gctgtttgct 1320
tttgctgctt gaacacatca atagacatc ctcagacatc ttgagcagcct gggggtttta 1380
ttttgtttag ctaaaagacg tctctttttgt gttaattcag acctggttttt ttaaannac 1440
cgcttttttc ttcataacgc ttaaaggaat ttcc 1474

<210> SEQ ID NO 333
<211> LENGTH: 4079
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 333

4079 bases of DNA sequence.
ctgggctgtc gtgtgtacgt cttgctggc ctgctccttg ggaagatggg ttctttcaaac 3360
ggcccagcgg cccgagggc aacggctccc agtacatgc ggtgaagatt cctctggaag 3420
acgacagac gccatcaggg gagaagaggg gacoaactct gacacatcag 3480
cggcggggaa gggcggaggt gacccacgca tctctgtgct cgcggcgttgg cccggcagctg 3540
gagcggatgg ggaagaagcg caagggcagc ctagggttcc caggttccag cctggcgttg 3600
gggcggcctc attttgtctt ccagatagg cctcttggga tgaagaggtt aagttgctgg 3660
ctggtggtct ctgtaaggat tggcgagctt ggttttcctta ggacacaga cctcttccac 3720
ccacgagac accttttcttc ctgatattg cttggagatt ggtgagttta caagtttagg 3780
tocggggag ggccagtggg ttgggatgag agggagttgg tctcttgatc caaactgtagg 3840
gagactgatt gtagaacttt toctctctca ttactctcct gtcacaggt gcacagggacc 3900
tgggctccag gaaacgctct aacottagagg ggggggggct tggggcagtc acgtctacag 3960
gctgtgtcct cctttagtctt ttctcttca ttttgggaa cccaaaacaa aaaaaaaa 4020
ctttttttct tctctcttatt tttttttt tttttggaac cccaaaacaa aaaaaaaa 4079

<210> SEQ ID NO: 334
<211> LENGTH: 3373
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 334

gggcggcagact ccctccggcct cttgctcacttt ccagcccttcctt cggcggttaatt cttcttctccct 60
agtttgcctt gacagcactc cgcacacatc ctgctggagc gaagaagaggg 120
acagacagac gccagcgtgg acagcagcagc ccacagccagcg cagcagcagc gggagagatcc 180
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 240
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 300
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 360
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 420
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 480
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 540
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 600
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 660
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 720
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 780
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 840
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 900
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 960
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 1020
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 1080
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 1140
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 1200
gcaggtctgc gcacagcagc cagcagcagc cagcagcagc gcagcagcagc gcagcagcagc 1260
--continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 335

gtccocgac ggcgctgcg ccocctctggc gcaggccacc gaggccgocg ccgtotagc
60
cccgcgtcgc ggcocccgtg aggcogctcg ccttcctctgc ccgctgtgcg
120
ggcgacgct aagggcagca atagaatccca tcaagcttca tcgaactggt actgcttaa
180
tgaggacaac tglgtgctca acagtaacct tcacacactt cactctgctca acctgcaaaa
240
gaaatttga cgccggcagc gcataaattt caaagtccaa acttcgttca
300
tctctttct ctgagagagc ccagctagct caccctaggg gccggcggtgc tggctgtaag
360
tcgcacatc gtcctccagt gaaagtaagtc acgtgaagtc ttcagctgg
420
cctggggagg caaattacagt gcagggagcc agaagccgctc agggagcgcg cgggtgcgtat
480
gcaggtgcggct ctaagccggc ccggagctgg cctagctgtgg ctagctgcgc cagaggtgg
540
agccggcctt ttccccagc agaagttttg ccagcgttca gttccagcag ctcctgcgag
600
catttttaag attaggcttg gcagtaactggt gcgggtgggt gcggcggtggc gcggcggttc
660
cacctcagc agccaggggg cggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
720
cctctgttg gtgctctgag cccacactgc tccatgtag ttcacaaaag cggagagctc
780
cacctcggtat cttgggtgct ttgagcttac cctagtctgac cctggcgtgc cctagctgtgg cctagctgcc
840
gggctcagact cccgctgact cccgctgact cccgctgact cccgctgact cccgctgact
900
cattggcttg ttgagcttac gttcgcctgc cggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
960
acagccaccc cttgctgcct ccgggcttac cttggcggcgg cttggcggcgg cttggcggcgg cttggcggcgg cttggcggcgg
1020
cctgtggatat gcgagcagtc ggtggcgtgc tggcggcgg cttggcggcgg cttggcggcgg cttggcggcgg cttggcggcgg
1080
tgctgtaaag tggccagctt ccggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1140
cggcagcagg gcggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1200
cctgccctcg ccctcgcttc ccgcctccct cccgccggtc cccgccggtc cccgccggtc cccgccggtc cccgccggtc
1260
cctggggcgct ggagctggcc ttggagccag gcagctggcc ttaacggctt cctccgtgcc cctccgtgcc cctccgtgcc
1320
cctccgctcag ccagccccag gcagcagctc cctccgtgcc cctccgtgcc cctccgtgcc cctccgtgcc cctccgtgcc
1380
cggcagcagg gcggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1440
tccacctgcgt cggatgcgagg gtcctccggt ttcgctgcgt gcgatgcgagg gtcctccggt ttcgctgcgt gcgatgcgagg
1500
cacccgccag cgtttcttcc cttggctgcct gttggcgtggt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1560
cagcaggcctt gcggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1620
tgtgctttccc tgtggctgcgt cctccgtgcc cctccgtgcc cctccgtgcc cctccgtgcc cctccgtgcc cctccgtgcc
1680
ggcgctggag cggcgcggcc cttggccagc gcgtggtgcc ccggccaggt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1740
aatggatatt tccacatttt ggaagtagct gcgcgctggt gcctgctgtc ggctgctgtc ggctgctgtc ggctgctgtc
1800
aatgctggag ccagccggtt gcggccacgc cttgcggccg ccggccaggt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1860
attgcgctctt cctgccgcgg gattgtgttt gggggtggtg ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
1920
cctggtggagt tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc
1980
aatctggtg gcagctgcgc ccagccccag gcggcgtggttt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
2040
cctggtggagt tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc tttgctgcgtgc
2100
acagtgcggag ccagccggtt gcggccacgc cttgcggccg ccggccaggt ggtggggtgt gcctcgcgtc gcctcgcgtc ggcctctctc
2160
acacacatctttataatcaacttcgtctgtaagcagcagcacgttctttttcttcttctttttcctttctttttttctttctttttttttctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<210> SEQ ID NO: 337
<211> LENGTH: 6633
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 337

ttctccccgc ccocccggtt ttgtcgaagt ctggygggttg ggaocygacc ccctgatygcc 60
gtgaagacca aagccgaagg cggtccctgg acactggggag attcgagacc cagggagttt 120
gagagaacct ttactttgta agagaccaag gtggaggggg gttgcttttc ctgacacgta 180
tttaacctga gcacaacctg aggtttgaa ggtggacta tacacattgaa ccacattccaa 240
aacggtgttt ttcggcccct tctctttgta gttacagggga gagaacaggg aagagacgctg 300
aagagacat ttcggagacc cgggagggac ccttctctct cttggtggctg aacctctgag 360
cgggaataca tccgagagca aacttccagcg acotatgccc accttcoccc ctgcttccct 420
gccccttggc tcgctcttca ccagcttggag cccatccctc cgccgcctttt cattaccctc 480
tgatcttcca ctcgcagaaa tcaacagctgg ggctcggacat ttcaacccct ccggccttgag 540
cttttgaggg aagccagcgg cccggcgaga cccaatccctgg gcagacagag ctggagaggg 600
ggaaaccacga atagagacga aacaagccgg cccccttttc gccgcttggga aagagagcag 660
tgccttcgggcc ccccccagtg ctggctggcc ttcctttacc aacccgccct acagagagag 720
gacagagcgg acatctgagcata actggtccag taacccggga cggatcggccg acatctgagc 780
accctcagga atggccctgt atttgctcct ggtggcaggg ggcttcccct catttttcttc 840
tccggtccac tgcagcttgg aagcctgccg acctgcctca cgggctgcttg ccctttttctc 900
tggctccac cgggcgcccg aactgctgcttg aacgggtccg cccggcctgg gcagacagag 960
tgcggccaac gcgaacgaag aacacgaacag aagcagagct gacttcagtc aagtaagcttt 1020
agcgccagca gcagcagcttt gcagctggctt gctctccacc aacagctgaa atgtgcttgg 1080
aacagattgt tgcacgctgg ctggtttttta cattgaggtg gttgaccttc aatcgccattct 1140
cctgcagggc gtggccagcc aagccatccgg tgcgctgaggg gaaccttcaag ccaccttctg 1200
cacacgtgtg taacatggtc cggggtctcc gcggccgcct caagttcctg ggtgctgctg 1260
atggctgcg cccgacggtg cccggcttgg aagacatggt gggctccctt gcgggctggc 1320
catggagaaa gggtaggtgg caacccaccc actctccacc cagctggcta ctggttgatc 1380
gcaagagtc cccacttggc cttgtagaggt ggccgctgac ccacccggctc ggtattcagc 1440
gtgaaaaacc aacgattgct tgctgtagaaa tctactctgg atacacatctg atggtgaaa 1500
gatgagggat ataggtatcc gacggcactg aaggtgctgc cctggtcaag aagagagcag 1560
tgggcaattc acctttagc tctcccaatg aagttgtgg aaggtgtata ctcctaaaat 1620
gtctacccac gttccctccat cccctgctgg cttggtccag gatcaccactg gttcacaagc 1680
gggccagcgcttgagctggca cagttgtggg ctcgccccct ctgttataggt aagttggtat 1740
atccagatt gcacatagct gtaataagta aacttctctgg actcttttgg ccaaccatgt 1800
cctccacac atacagagca gcctctccct cagagagggt acctgtggag ggcttgctgac 1860
ctccggaaca gtcgagagga ccctcgcctg ggtgctgccg gtcgaagctgc tcttgggacc 1920
tggaacagca gacggtcagc tcggtgctcc ccctccgcttc tgtgaacacta cgtggtgctg 1980
tgcagtctctg tgctgttgg tgattgtgat cattctctct attgtctctg tttcctttg 2040
gaacagaaa cggaggtatg aatattgctg gaggtcttatt ggtataattcc gagcggatgg 2100
aacatgaatat atattagttgg aacccatgaa gactgcttatt ggtactaatg ggtatatttc 2160
aagagatgag tctgggttcc ggcgctgctt ggcgttggag gggtttgag ggtttgttaa 2220
aggaacagoc tacatgattaa ggcgtccctaa acgtgcctag aagagttcag tgaaggtgtt 2280
aaaaccccaac ggcacagtcct gtaaaatcata acgctctctg tctgaagctga agataatgac 2340
tccacgtgg gcaaccttgaa aatgttata cttggcgtggga gcttcagcaca agtcaagccc 2400
cattaacatc atccagcgctt atgtctctta tgagattgtc gtcaactatt ttgaataagaa 2460
tagggatatg ttcaactgca aacacccaga gagcgcaccag aagagcttgg atatotttgg 2520
atggacccct tgggtggggaa gcacagggag cttgctttttt ttcattttttt aaacacattg 2580
tgacactagc gcacaggtag aggctgatac tacacagttg tcgcacagac tcggaaagga 2640
agaggtttct aataacccgc acaactcaag atacaactct gatcgcctca cctctataaa 2700
gaacacatc atgtgagctg ccaagatccaa aacccctcttt tcaagtgcata actcagaaa 2760
ccttaactta tgggtatttt tgagttccct tctcaagtgg gcgcgggagga tggagtttttt 2820
ggctcaaan aatgtgtgct aagctgtttc cggctgcttc aacgcttcgg tggcacggc 2880
aaanatttgac agagctgttg acctttgctt ggccacagcc atccagcattt gtcagtaact 2940
tgctgagaga ggccagtgcct cttcgccttgtagggagttgg gttcctggaga gatcttttta 3000
caaacactac aacagnctggaa gtagagtctg ttctactttt gggactttctt 3060
tcctttgttg ggctccctttt aaccccttctg gctggaggtat gttactttttt aaccaagttt 3120
cagagttgty caccgcattg ccaacgctga acaagctaca agtgaagctt aagatacttt 3180
ggtaaagtct gccaacctgg aagcctggga gagacccccc tttaaccaccc tttgtgccag 3240
tgctagactg ctcgctggct gccagatataa aagagtttac gaaaaatttga aacccacattc 3300
cotctgagtt ccactctggct cttgcctgcag catctgcgttg ggctacagcc aacggctaat 3360
tgctgtcacc tccaataacc cggagagaca ggtgaaggac tggagggggt tgggttggtg 3420
gcagacattg aagcgtgacgg gctgcctccat cttcactctgg cttctacatt aacccctctc 3480
tgagagggac gacccggaac aagagacac acaagctcatg cacccattga aagagttgct 3540
cattacagcg ggtcacttct gcctccctttt ctcctagagc gcggccgagcc caacgttga 3600
cagctgctat gttgacttac ctcggattga cttgccagac gcggccctct cttgctactt 3660
gtcagctgcag gatccaggcg gtcctccca cttgctggggc caaacctgga tccctgctcg 3720
aaaaccaact tatagcatgg cgaggttga gggagaggc tgggttattg gattaagagaa 3780
gttccacgcag cacggctctgg gcgggttctct ttaaatggaat atctctttgatt 3840
gaacccacct acgtgcctggct ctcgactagc ctcagagaca cttcaagtct ggtggaggtt 3900
tggtggagtc ggcacagggcta caaaatattc acggacagca acaacctggaag ctcagcttgtg 3960
atggagttgg cgcacagcga caaaatatct aacgacagca acaacctggaag ctcagcttgtg 4020
gtaaaaatgc tcaacagcgg ctggctttcggt ggtgtgcattct ggcacgagac 4080
aagagaccccg tcaacagcgg caaaaattct acgcgtgct gagcggaggt cttgctctgctg 4140
aatcctttatat gcaagttgtgc gaaaaacctt ttttgcacct ttaaactgctc tggctctata 4200
goatttgcct atttttttta tggtaagac ataaagata ataattacaac aacottttt 4260
-continued

aataagattg ggtcatttag aagcctgaca actatttttc atattgtaat ctatgttttat 4320
aatcactca ctgtttcag taaattttaaa ttagtaataa tytaacattga tttoocctcca 4380
gagaaagcag aatattaaaac aatcctcaat aatgtagttga tgaattgttac agttttttgc 4440
attttatata atatctacttc acatccaaacact ttcgttttat tgtcctttgtt gatattaatat 4500
tagttgaca taggaacaa atgaagagta gtgtgtccca ggaagtcgag aatttttaact 4560
gctactgaact gttgcccaaa tttcatgtagttataaaaacta ttaattgccc tctgaataaa 4620
tgagattaga aacaaacaaa actcattaagc ctaaaaggct ttaaaatagt aagccataacc 4680
ttcgttagc aataattttttttt cttcattgc tttgagaaac ctttggacaa 4740
agctctgca tggacggaag tttttttttttttttgctggaaaatgctgagcata aagcactcaaa 4800
cccaaaaaatttttttcagctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued-

```
aatgtatca cgaatgcgcc tgcctactggt ttggttttaaa cagctgtaaa tgaagctcctt
6600
tataatctca taatgtatat ataatattaaa gtt
6633

<210> SEQ ID NO: 339
<211> LENGTH: 994
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 339

tgctgaggct caaccttcgg gaagaatgcc ggaagagctta aagctgcaaa cgggtgcgag 60
aagcgctaga agggcagctc agggcaggtg tctacttca acacacacat taaagcaccg 120
cagtgggagc gggccagcgg caaagcagcc agtaccttc gaagaagcctaatgg 180
gggagtctcc gtctgcgttg aagacacgcc agtcaagcgg ggctcgctcc 240
tggcgcgaag agaagcgtacc ggcgccagag ggagagcggc ttgagctgat caaagcctac 300
atccgagaag tcaagcctgg agagagccac tttggacttc tgcggctcaaa gttcgcggac 360
tgacgcgcgact caagagccctt gggacccctt ggtgccttca gcagagtgca gatgcgcaag 420
cattgtagc aagcctcgttg ctgcgctggc aggagggaga taagcgggcc ctggttcctg 480
gattgcgcct ctcagctatt ctctccacag tggagaggtt ggagagcctct ccagcctggc 540
cgggagcagg cagggggtgt agggcgcgc aggcgcgtt ccgccctctt gcagggcagc cagggccag 600
accccccact ccctccgccc gcacacttag attatatttt cccccaacagg ctggagaggg 660

gcctccacg atgtgggccc cttggggccc ccaccccttt ccaccccttt tgggggtgcc 720
gacccgcaga tttccctttta aagcccttgc ttccacgaggg ctggagaggct ccagcctggc 780
gggcagtgct tgggaggggg ttgctcctag aagacgtcct gcagcaccag cagccccctg 840
tcccccagcgtcggggcag gccacacttt tcctgcttag gcacagcctcc 900

tctgttcacgt ggaagaagggg aacaactcatt cggagacgct cgggccccat agaaccagt 960

cagacoctttt cccaccccaat taaaagccaga acca
994
```

<210> SEQ ID NO: 339
<211> LENGTH: 772
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 339

```
agctgctgcc gaatttgcga cgaagcgggt cggagcccctag ggcgggcgaa attoaagtcc
60
tgtaaacccc gctggtcttct tttggtctcag tattggtgct cacgagttt ggcgcgattaa
120
gatccgatct ttgctgcacag ttggctctaa gggggggagcac ttttttaggg agttggcacc
180
cggagatct cggggcgagt ggggtcccaat agggtaca gagaaaaag tccaccgggt 240
cataagggat tctactggttc cgggagtaag tttgtgttaa ggagagctgta ctggaggtgc
300
cattgactt cggggccggt cttcgactgta aatattttcat ttctggttat gagcctctcgggg 360
cctgtcttttc atggtggccag cttgctgacs tacaaagcctgc tggctgctttg ctcatctctg
420
ccttaagagc gttgggtggc atgggagaga cttggtttttt ggagggattc tggatggctttt
480
ttctgtcagag aagacgtgctt cacaaggccc aacataagc ccaagctacc 540
tctggtgctc tcggagtgtgc cggagatgtt gtcagccagca agctgcaactg cggccttcgag 600
ttttctcttg ggttgctccct gagtaagata atctggacgtg cccgccgttc tttgtctccot
660
```
---continued---

gcttgctgt gccccatcctt atcaagagac cattggaagt ggtagagattc aagatccaag 720
attgctttta aatgtagaac tcgaacatgaa gttcctttgt agtgcctaaa aa 772

<210> SEQ ID NO 340
<211> LENGTH: 7365
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 341

gagcagtttgt aagtcgctag ggaacggctag aaggtcatgga ggggacact aaggtcaagt 60
ggggaacct tcgctagactt aactggaactt tgggctgaggg gaaactatg tatctttataa 120
aacgctgtctt gctatgactag aacgactctg gaccaagctgt gtaaatcacttc 180
gttctacag aatagccaggt cggcgctgctc ggctgtcagct tcgactacacg cagagagttc 240
tggaaccg aacctgactc gccagtaactc tgcgtcagctc gggggtactg aagctgctg 300
tgtaactcc cttgagtttc tgcgagttttc ttcagagtc gcaagctgctc cagagagttc 360
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 420
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 480
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 540
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 600
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 660
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 720
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 780
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 840
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 900
taataacctg caagctgctc gggggtact ggcgtcagctc cagagagttc cagagagttc 960

<210> SEQ ID NO 341
<211> LENGTH: 7365
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 342

tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 60
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 120
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 180
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 240
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 300
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 360
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 420
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 480
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 540
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 600
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 660
tggagagtgtg tggagtctgg cggcactgct tggagtctgg cggcactgct tggagtctgg 720
-continued

ggaaagaagt gttgacaggg aagacgccag attgaccgct tcacacagaa gggtccacaa 5340
ggaaagccg atggagtggct ctggcgtagc ccacagacaa tttatgtcty tgytaaaagt 5400
tgcacagaaa cacacatata cttatactaa tctaataact gaagaagata ctgatgttgt 5460
tttgaaacga gctggtcggt ttgggtgtaga ccggacacctg aataattttc tggatttgtc 5520
ggagagagaag tgggttagtta gctatattcg tgtgaaccag tcctattaaag aagaaasat 5580
gctgaaatag cattgattttg agatcagagg agatggtgct aatggaagaa aaccaccaag 5640
tccacagccga cccagacgaa ccagacagaa aagatattcgc agggggtcag aanatgttgt 5700
catagtggccct tcccccaacaat tcgacacaga tcacagtggc tggaggtgac agctgtg tgt 5760
tggtttcggtg gtagagggcgt tttcatacatt ccccttggcg gcaagagttc accaattgt 5820
ggtttgtagc ccaagttgct ggacagagag caattggtcct catgcaatgtg ggcaagatgtg 5880
tgacagcctg ctggtggacc gastrgggtggttggtaccagt tagagacctg accagtagcc 5940
ggacgtggac acctacatcg tacccagat tcccaacacag cactactgcg tcgagaccgc 6000
cacaatgata gacacacgg accacagaaag ctagcttacta aagttggcttc tcacggcctt 6060
ggacagcctact cctccttccct agttcttctt cttccttgcc tgtgctggcc 6120
tcaacgctga aaggtctctct ggtcattgca gaactcttcct aagatcttctt cattctagc 6180
tctctttgga tcgactccagct accacaaaaat tggtaggtt ttttattgctg gcaagattta 6240
aaccacttaa cccagcccccga attgacaggag atcgaccctt atatattactg aagcttattggc 6300
ttctatttcc ggtgtttgtt ggtttagggc tggaagcaca gatgtgcttt gcttcaagag 6360
aattgtgttg tcctcattact ttttctcttc aaccacttcg gttgacacat gcagacagtc 6420
agacccattc atgggaaggag agtctttgct atcagatttg tgccttaagag tcgaattgtg 6480
cctgctgccg tcctcattact gttgagttga aacctgggca gaaatattcct aggcaagatc 6540
tagaaatgg aagagacactc gcaagcccttg ctggtgtgct caagocgtta atccagcagc 6600
tttgggaggg caaggggrgg acgtacatgg agtctggcac tgcgaacgca gctcgtcaggc 6660
caagtgtagaa cccacatctc actaaaata gcaaaattag cccgcctatttg gttggaacc 6720
cctagacccc agtctattctct attggctagggcc gggaggttggag 6780
ggtgccagtg agccagacgt atacacacggct accacgctct gggacacagtt gagaocgtgg 6840
cctcagaaaaa aaaaaaagaa aagaaatagaa angctttttcc aagttttttc atgattgatat 6900
attgattca gattttctaa gatgtgcttc tccaaagcga gaagattttg aacaaccctgt 6960
tccacacgag acgtgctctaa ttcgggtttt gtaaatgacta aatattttt gtttgtgtc 7020
caggtatata atccacactc ctttacaatt aagacocctg gcatgaattc ttctcatctt 7080
ctaattacg gacccagacg aagacagctg tttgtgacct ttttttctttt gttgcttttt 7140
ttgctctctct tggctgaaacctcatacagctc ataaataattt ttttgctttct gaagagaaaa 7200
aagattttctt taaaacccagc aatcttattc cttgcctgggg ggaactatgcc 7260
ttccccaccc cccccccgct gcaagggcag tgaagaacattttaggcatcttgg 7320
taaatcttggt gcttgattc ataagacataattgcgtgca 7365
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: (0)...(0)
<223> OTHER INFORMATION: n = a, t, c or g

<400> SEQUENCE: 342

atgagact ccgaaatgcg gcccaggtg agccaaagat gcctctgagct tcatatgac 60
agttgtaaa gcaagttgag gcgaatgaga tggagaactc aaatctttga caagagctag 120
aagataatt caactaattt acacaaactg gcacacgagcc acatataag tggagcta 180
ttaaacat ccaaggaagt atggagatag aagctatagc ttctcttgga cagatgtatt 240
tattagacg cccttaaaggg cttcatctag atacgtaaat tttctctgga gtaaaacctc 300
ggctaaaaa gtcctcttca tgttataagg ggcgggagaa agcttgatca agcagtctg 360
gagagtacg ctctgtctct ctgtggtcat tttcgcagag gggatttgta aatgagacag 420
ggaagagao tggatattta ggaagaacct gcagagagac gtcattgctt cttggtgatc 480
tggaacagag agagaaaccg aaagacgctg atacaatcaca actcgaagct tctacaat 540
gaatagagc tttcttttta actgaaactt tttctcttaca aacagatatg accagaaagc 600
aataggaata tggagacgag ccaatagcag tgtgtgtgga acagagacta ggtacccgcc 660
aggtatagca aacagacaga gccaagaaag tccgcagatt cccagaaact caaaggaagg 720
taatctggtg aagacgtgct tttcagttcc aacgaacagct aagctacgacg ttcctcagaa 780
dccagagta aaccagctca catgatctgt gcaggccagaa taaaaagcgt gcaggtggag 840
aatcacaact ggcacatcct ggttaagctgc aggcttacac taoaagaatg gaccaatagaa 900
cagcagctg tttcgatcac agtacacacc actctgcccc tggaggtcgc acaacacctc 960
tgggaacacca ggtggaaagt tttgattctaat tgtttgcta gttggtcttt catgtgataag 1020
agttatagct ggcataattg ttcaggttctt ctctctctct attatggtacg acctgataagg 1080
gacagctttg atgtcttctt ctctcatcttc agtcttttacag ggtgaatcgac aagactcttg 1140
tatttgtggg aatctctcgag gcggatagag ggtctgagggc cagggcagct gcgcgagtcc 1200
acaaaacatc taocatcagc ccctgagcagc agagagcgcg agctgaaatc cgctcctttc 1260
atttttggg aacgatactg gcctagttgc aaccctcgttt ggaattggcg aggctgctag 1320
aacacagggt gacccagcag aaaaatctca tgggacgtcc tgtttgagct cagatctgctc 1380
cgccgctgt gctctctatt aaaaactctttc agtcttctgag cgtctctctc gctctctctc 1440
aataggggg ccctacagcc attcgacatt tattctagat ggctgagttt tggctcgggc 1500
atcattaat gacatactgt caaactatac ggcgatagac tggcaattgc tgtgacaact 1560
tgacctttgg gatgttgcag aaccagctca cgcctagctc tgtaaagagc tgcattgag 1620
cactttggg ccctctatat ttggaagagt aacaggttaca gcgatttact gcaagtcttt 1680
tggaagatt gccctctgcg gcaagtattg atagtaaaaa gcgyttgcga gaaggtggaa 1740
gttgagacag atctgagtaa tgtgttgactt aagtttaaaa ggcacaaacc tctcaaacgr 1800
tattgagtc cttctctcact ttttgtctcag atgctcagt gatattattg gatattattg 1860
cgttatagct gctctctgcag ttttttctgc gcaacttact ccagcagata gctctctcct 1920
cattagcttt ttttggagat ggaagttgga gccatccgag ttttagctct tgtgataag 1980
caatagag cccagccca acctcaagag gagcagcttg cttctacatt cttctacatt 2040
actaaataatt ctatagtctg ctaaatgctag tggaaacctag tggaaacctag 2100
cagcaagaagaa tcaaatagac caggaagcat tgtggccat ggggcaagt ttgacagctca
agacacattt cttccttcag cacaacagta tttgcattgg aaagacgcca accttttaagaga
ttctcagcgg acaatgcagt gcaagtcagaa cccttcagta ctctacagttt cggacagtta
agcaagacag caagcaagact cttatagttg attaaagtttt tgcagcaaat cgaagctagt
agtaagtcgc acataagtttt actaatgcgta aacataagttct cttttcccaact tatttgast
ctcaagctgt acccagctcc ttctccctcag ggcagacgctt gtagattcct cggctgtga
aagataagaa tttgaggaga gaaogccaa tttgctggtg ccaactacat ccaagcaacag
aacaattttt aacatgctcc acaagctccc acaagctccc acaagctccc acaagctccc
ccctacacca ttcacactt ctcaacacct ttctcctcag gcaagcaggct aacatgcttt
ctagcotta tgcacattta gaaatacaac gctttccaaa ttttagatca aataagttca
gtataggctga tgttaagtga aaaagagctc ataaaacacc tctagctggg tttcctcccgt
aagatgtaga tctcaggttt tgcggatttt tgcctacatc accagcagct tgcacataaa
atcataaggctca atataagatg ctggagacta acagcaaatg acaaaactaataa
gttttaata ttcagagag cagctgcaag cttggaggca aagcttcaac gaaaaagaa
agtgcagaag acaaaaacc aactaaagag acatgagaaca aacatgcttt tcaagcaggct
ctggaagag acaagcagtg tcaagagtgg atatacaat ctcataagtt cagacgactt
cttacataa aacaagctcag acaagctcag acaagctcag acaagctcag acaagctcag
ccgcattag tgcagaggta ctgctcaact ctctatcact gcaagcaggct
agataaggctc ctctccagt cttgctcagag tgaagcttgtg cttgctcagag tgaagctgttt
tttctctact tttctctacag aacagagatt gtcgtagaat gcagcagacc
caacatgttt ttcagagag acaagctcag acaagctcag acaagctcag acaagctcag
gataaggtc ccgcaagctcag aacaagatg gttcgagact cttgctcagag tgaagctgttt
eaagaaagct tgtgaagtgg atatacaat ctcataagtt cagacgactt
agataaggctc ctctccagt cttgctcagag tgaagcttgtg cttgctcagag tgaagctgttt
tttctctact tttctctacag aacagagatt gtcgtagaat gcagcagacc
agataaggctc ctctccagt cttgctcagag tgaagcttgtg cttgctcagag tgaagctgttt
tttctctact tttctctacag aacagagatt gtcgtagaat gcagcagacc
-continued

cacagcaga agtcctaaa astaaagcc ctactgtgta aataagagag agtggtccct 4440
agacaagtga agtaaatgt gcagtccaga gggttccggt tttcctcag atgtgatact 4500
atgtatactt tcgacggaag aagtccacag atggatatta ttgcttcacc agcgtgagtg 4560
cttgagcct cgagagagca tttaatcaca aagaaagga attataagca atggctccag 4620
ttgaaggaag tgcacatgag aagaaagcac atgcagagca gcaataagaa taaaatagaa 4680
aatcaaggaag agagagcagaaa aaaaactatgg attcagaaaa ggaacatatta gatggattcag 4740
atgtgagtaa tattgagaa atagaggttc gtagatgtatc ttcacaagcc aacaagttcaat 4800
ccacagtaac caaaaaagccc gccccaggtct ctcataaaaactctcct gttgcggaga 4860
aacaaagtaa agaagtctcgt ccacgttctg cccaatatc taccatacaca aacaggtttg caaaccctaaa 4920
agatgagtgaa aagagagagc agatgttgat gggtagatgta tggagcgctg gggagccac 4980
tataacatc tccatacagct acacttccaatt tgatgtaaac atogacgctct ctcataatgg 5040
agatgagtgca tggagagacg gttagaggag gacaaagctc atggatattt gaaaaacag 5100
atccacactt tagaagagagc aagaaagctag atggatctga aggaggaagaacctctcgtg 5160
taacatacaag taatatgtgac gccataaaaa cagaggagcc tggatattt ttgatagttct 5220
ttaaatctct tatttccgagc ggaaggggtc cacagctctt cgggtgatgaa aagatagttg 5280
aacaggtctg caagacatct gcgtctcttc ctagcccccag cunaaatcc ttagatgtaa 5340
agaaacagca acaaccttcccc cgacgttata ctacccaaac aatactagca tattagcaca 5400
gtgaagaaga aatgtagagc ttaaaaata atttaaatgc tggagagcttt tttcatcaaca 5460
acacagttta agaagccgctgg tttaatgttta gaaggtctct aacattttaaa cagacctcag 5520
caaaatgcag taatcagacg agagagagtc agaatgttcttt cgactctcatt ctcataaacc 5580
tatgtaaggt gactcactttct tgcgttcatt cagcgtcttct ccataaattg 5640
atgtgagtaa tttggtgttc tccgagggag gggtgatttt gaaaaagcag cagctgtattaa 5700
agataatca agaagaggtc tgggttggtc ttcacatcatt ctcataaccc 5760
ataagacaga agatagcagcag aagacagcag aataagtcg tcagcttcaaa ccataacccc 5820
agacacacgt ctccttctct cagcttcact ccagatacacc aacagaggg gccgaacagt 5880
atgaaatgt cacagaaattt gcagatttta aatactcagtt ttggttttct tataatttctt 5940
ttgctgtcct ctcagagagg aaaaaccacc aataaaatct gaaacctatac 6000
aagagactca agacccccgcc tctcagaggg aaccaatgtaa ctcaccaatac tcacagtctg 6060
tacataaacc cttccagagttt gaaatcccc cagtttcttt ctaagacaaa agtttcctca 6120
gttctctatt tttgtgcctt gaaataggtca tcggtgacgta attgtataagc tcggaatagc 6180
ccaaagaaac aagcccttcaag agacttgagct aagacagctt aagacagctt ccaaaataaa 6240
tgggtgcct attgagctt ccagttttgg aagataagac agacagttg 6300
cacacactcg tttctcctct cagtaatctgc gagaaggtttt cggagagatc 6360
ctcaactacc cgggaaacctgt ttaatatcctc ctgcgggttg aaccttctag acaggatgg 6420
cttcaggttc ccagtttacg ctcctctctc aacactcagtc aacatgcttc gaaagagcg 6480
ttcctctaa ttcagttttc ccagttttaa tcacgaaacc cccgatttc tggagagcag 6540
taaacactgg ggcagagggcc cagttggtc ctaaagactg aatgcgcaggg 6600
ttcagaggag aaaaaaattt tataaaagtgt ttgggtactag cagagttgctca ctaatttctcag 6660
ctsnactcct actcagggaa aatggtatt tgtgcacaaa aaaaattttt tgcctctttg 9000
tagccacacta acatcataat taactcttgct gttcgaat attcagtaat atgatttccg 9060
atgaaacagc tttaaccagc ctgtttgttt tactgctagc atgaaactga tgggtcaatt 9120
tgcaagtaaa ttggatcgc acatgtgaag catagagata gcatacagtg 9180
ataatctcga actatattgt gcotcaacac aacaacaaaattcggttaactgcctt 9240
gatttaaat atttacactg acataagtttt tctctgaagaagcgtctgattg 9300
ctgttaattt gcgtatatgcc tgtatttctgg ggtgccatgtc tcgtcttttt attaatgaa 9360
cattgaatat gcttcaacag aacactaastg acacacctcag atcataattt tggcttatg 9420
aatcgttacg tcgaatggcg aggtcttgac ttcacattttg ttattaatatt 9480
tgttttaaaa gctctcttta aagctttaaatttttatttttctctaggtcatctga 9540
agatgatagt tcttttcatg taaactaaaaaa cccgctgagc agatgctgttc ccatctaca 9600
tcctagcagc tggcaacctt cattcctcca tattttttatagtaccatg ctaaactctg 9660
ttctacacct ccacccctttt ccaccccttc ctcagcctctgc gtaaagttcat 9720
attcattgaa tagcactgca agagcagcag cacagactaa gcaattgaga ataatggccc 9780
acataatta ccacctccacct atattattatatctggcc ttcacatttc ctccttcagag 9840
tgctccgctc tcctggcttt tactgatttaa actgcttggc cccctctctctttgtgctca 9900
actggcctg acaattgcaac ctttttcccc cctttatgtt aggccagact ctcagcagctg 9960
aagatcataaa acacatcctttg ccactctcaag gattttaaatctcctgc ccttctggaat 10020
gattaacactctgga tctgcagtaa cttttttcag gcacccgctgtt aagttgaccc 10080
tttatttttt gtcattctcag ttcctggtta caagttcagt gtccctctctc aacagttggag 10140
gagtcctctc ccctctgcgg aaattaaccct cccatttatc aactcataa cttctctctca 10200
tatctctctc ccattcttttt tttaagatctc gaggggtgtc tgaagatacct agtccttttg 10260
cttgctgaat aactaataat ttttttaaggt tctgtaattttt cacaactctctc gtgcttgggg 10320
agagaaaaac tttttaacca tgggtgggca otacagatagg agtaatataa cctacgcttgt 10380
gctgcat 10386

<210> SEQ ID NO: 343
<211> LENGTH: 2191
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<240> SEQUENCE: 343
ggtggcgcga ccggggagcc ggagcaagac cccggggttc gctgccgtgc tggggcggcc 60
tggtggcgtt ctgttgccgg gcctgccttt ggcctgccc ccgggtccggg 120
agagccggtc gcggccggct accggccctg tagggccggt tcgggctgcggc 180
cctgcacgagg ccggccggtc ttcggccggt ccgggtggcc gcgggtggcc 240
gccccgggt tcgacaacgc gttccccctgc ggagccggtc gcgggtgctgc 300
tggagtttt tggccgctgg g tgtggagtgc aaagggcgg gatcctgcgg ggcggggtgg 360
cccagagcatg gcggcggcag ctgggtggga caagttgcgc tggcccctcc gcggggtgg 420
tggggcttt gccgtggctc agctgggggc ataggggggg gccctcctcc gcgccggtgg 480
acacacacag cctggccgca ccgggggcgc cggcccctgc gcggcgagtc acaacacagc 540
cagcagtcccc cacaaaaagc acacgttttccc tctggacccag gccagcag cagcttcaac 420
gtggcctcctg gccggygcoa cggcotaaccc atgaccttgt acatctgtg cattgacotcctt 480
tacctgagtctc atggagactg aagacgtgat gttggcagct gcctcgggcoc 540
cctacgccagct cctcggcttc gccgctcctt gcctcgggtgc ccttccctgg caagacccctg 600
cctccgctgag tgttacagctt tacctgataag tctggaascc cacgccccca caagagaaa 660
gatgtggaag ccccttggtcc ctctgacggcgtc gttctgaagc tgcaccaacc acctcagacc 720
tttcagagcc agtctgggaa gcagctgatc tccggaaccc tgttgatgcaac cagaggttgcc 780
cctgagcctag cgtacggctt gcctgagctgc cggagctggaa tctggctgac gcagctgacag 840
cgcggctgctg tcgttgcacct gcacgtgctgt gcctctgcct gcgggagccg aaaggtggtg 900
gcctctgctg ccccaagcag ccgctggctg caccctggag cacaactgtca caagagggac 960
atagcctagcc ataacaacac gcttgctgcaag cttggccccc cagctgtgctg ccctctctct 1020
cagcagcctg tccggtgacac cagctgattc gtagaaccg atccggagac cagcaaggtc 1080
atcccaaggt cagccctgccc gcagctgtctc ggagactccca gcaatgtgtct ccatccctcc 1140
atagagtgtct ccataacactt tctctccaggt gttttctctgt gcatacaagc cccctcccag 1200
acccctcaag tccatctcaag tccctctctgct acaaaatagg gcagcagcag gaacgcagcag 1260
agaggtgatc gcgtgcgctg ctgcatcaat gctctgacat cctcctctgt gcagctgacag 1320
gccacaggt gcacagtgtt gcacgtggctc cctggtgctc cactggacaata 1380
gtgcacgttg cacgtctctcc cctggtctgg gcctggctgac gggagcagag cagagagcagc 1440
agctctgtgg atggcagcctg ctttctctgg cgctgcccac tccaggtgctg cagctgtgctg 1500
atggcagaa atctgtagct gcagacacag gcccggcagc gcccggatgc gagaagsgac 1560
tgccggagg caccaacccc actcactcggct gcgggagctcc gcagctgtgc 1620
tgtgctgtgct ccacccagag cggcctggctg cagctgtgatc aacgctggagg aagagctgct 1680
gacacacacag atctgttgagc ctaaagcccc cctgtcttcgc gcctggcgggag caggggctct 1740
tgtgctctct gcagtgctgc ctgctccgag ggtctctctg gcagctgtgtc gcctgggtgc 1800
agcagcagct gaggccgctac gctggcctgc cagctgctgc ggcctgcgtc 1860
cgctggtgac tacgctgagct ccctgctctg ccctgctgctc gcctgctgctc gcctgctgctc 1920
gggctggctac cccctgctcc gcaactctcat tcctctggcg ctgctctgcgc gttggatggctg 1980
ggcctcttg ccgagcagct aggccgcttc tgtgggctgc gcctgctgctc gcctgctgctc 2040
gtggacggc caggctagctct cctctgtgctc gcctgctgctc gcctgctgctc gcctgctgctc 2100
gagcagcagcg ccggcagctg cagccgcttc gcctgctgctc gcctgctgctc gcctgctgctc 2160
gcagcgccac ccagccgcaat cctctgctgc gtcctgctgc gtctgctgcgc ggtggccgcc 2220
atttctctgggt cttcagcgtgc cccctgctcc gctctgctgctc gcctgctgctc gcctgctgctc 2280
ctctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc 2340
cccctgctcc gcctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc 2400
ggacgcagct gcagctgtgtc gcctgctgctc gcctgctgctc gcctgctgctc gcctgctgctc 2460
cctctggaga ctctgtgctc cccctgctcc gctctgctgctc gcctgctgctc gcctgctgctc 2520
ccccctggag cccctgctcc cccctgctcc gctctgctgctc gcctgctgctc gcctgctgctc 2580
agcttctgctc gcggctgctgc cctctgctcc gcctgctgctc gcctgctgctc gcctgctgctc 2640
gtcgaaggca gccatacagt cttgatamag ttgggtccaa tttatattca tttaaacttg
2700
tcaaggtata aatggcatac ccotattaatt tattttaaat cccatcacgtg tatagaasa
2760
aaataaaacc ttcacat
2776
<210> SEQ ID NO: 345
<211> LENGTH: 3160
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 345

cctcctcctcg ccggggtccg ttgggtccac cctcctcctcg cctcctcctcg ccctctcctcg 60
tccctcctcg ccggggtccg ttgggtccac cctcctcctcg cctcctcctcg ccctctcctcg 120
gatggggctg gcctcttctat cgggggtccg ccctctcctcg cctcctcctcg ccctctcctcg 180
ggctctctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg ccctctcctcg 240
tgggggctgcc cgggggtccg ccctctcctcg cctcctcctcg cctcctcctcg ccctctcctcg 300
gcctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg ccctctcctcg 360
gcgggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 420
cctcctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg ccctctcctcg 480
gcgggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 540
ggctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg ccctctcctcg 600
cggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 660
gcggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 720
cggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 780
cggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 840
gcggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 900
gcggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 960
tggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 1020
gcggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 1080
gcggggggtccg ccctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg 1140
ggctctcctcg cccctctcctcg cgggggtccg ccctctcctcg cctcctcctcg cctcctcctcg 1200
tttttgattc aacacactaa aaccatttcc agatctaaact ttttttgtct aacacactaa aaccatttcc 1260
agacacagcc caacattatt agatctaaact ttttttgtct aacacactaa aaccatttcc 1320
cagctctgg cagctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1380
acatgctctg tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1440
ttggacattg gacactatac tataactatac ttttttttttt tgtggtgggt gacactatac tataactatac 1500
acatgctctg tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1560
attataagc tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1620
acatgctctg tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1680
tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1740
acatgctctg tgtgctctgtc cagctctgtc cagctctgtc cagctctgtc cagctctgtc 1800
agagtttctc aaaccatattg aacagctatac tttttttttt tgtggtgggt aacagctatac tttttttttt 1860
atacattctt ctaccaggga ccagaggaaa cctcagaaaa agtgaanatt ggaagttcatt
tggtcaaga aatcgatacg atttgcaagta tagagctgc agataatgc aaggaatato
tagttacttaaac aatgacctgg caacaagccaa aaagacaaaa ggcacaacgat
acttttctcc aasatctgcg tggacaacct actcaccacaa macagtagag gagctgccaa
atogagaggctcaggtctca acttcgtgaa cacacagatt taggacaat gaaacgtgac
atatagaga ttcgacaccc actgaccttg atccagaga taacccctttt gatgaaacgc
agctatacace aatccaaaaa gttctgaattt ttcttttcccc agaggtatataa aacacactga
asaataacca gastaacactg aasaattggccc ttttttttttt tattcgaatc agaacccttg
agtggattac gatgttagaag aacacatcct tttctccagaa cactcttggtt ttatacctata
catccacgg attttgccac tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatacttct tttgcttcctg gttaccaactt gaaacatttc taggatcataa aagagttgca
ccttcctgt tagttcagtt ttataaaaaa gttggacacag actgtgatgt ataccyagga
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
ggtggacacag actgtgatgt ataccyagga aacagatagat tttccagtct tgtgacagag
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
<210> SEQ ID NO 346
<211> LNGTH: 2629
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 346
acttttctcc aasatctgcg tggacaacct actcaccacaa macagtagag gagctgccaa
atogagaggctcaggtctca acttcgtgaa cacacagatt taggacaat gaaacgtgac
atatagaga ttcgacaccc actgaccttg atccagaga taacccctttt gatgaaacgc
agctatacace aatccaaaaa gttctgaattt ttcttttcccc agaggtatataa aacacactga
asaataacca gastaacactg aasaattggccc ttttttttttt tattcgaatc agaacccttg
agtggattac gatgttagaag aacacatcct tttctccagaa cactcttggtt ttatacctata
catccacgg attttgccac tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatacttct tttgcttcctg gttaccaactt gaaacatttc taggatcataa aagagttgca
ccttcctgt tagttcagtt ttataaaaaa gttggacacag actgtgatgt ataccyagga
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
ggtggacacag actgtgatgt ataccyagga aacagatagat tttccagtct tgtgacagag
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
gttccgagaa gttgctcgtta cccacctaat gctgcttctgc ggtgtaaggc aagtgcgggg
agtgcaagta aagggacaata tttattgctg gactctggac cattaccacct cttcagacat
atccacacac cttctcttag cagctcgtcac ttttasact ccacatgctga ggaattggcc
agctcagaga gttcggttgt ttctatgtcg aacagagtgct gtcgacgagct gttatagaga
acccgacacg tgttctgccc gttgaaacagc gttgcttctg gttgctaatg
atatatatattt tttgctctctg gttacccacact ggagtgcgtta aagagctcttg tgtatatactg
gttcatcctc tccctcttctg ctattggtggc aacagatagat tttccagtct tgtgacagag
-continued

tccggccagt ggcctatcac aagcagt cacacacgcc gggatgtgtg aggctggtgc 780
tcctccata ggcgcgtgct gactgtgg attctctctct ctcctctctt gctt 840
tggaaggat tttggtgttt gattttttttt gatgtttttttttttttttt ggtttgtc 900
tgtactctctc tgttccgct gcttctc tcttcctc ttcctctctc 960
ttcacctg ttcttggggct ggcgttctgac cactttttt gttttgtttttc 1020
tgcgaact ctctctctcttgtaat gtctggttaag gctttttttttaa 1080
gctttcttct gtctttcttct gctctctc ccttccttctcttctct 1140
agagagatgt ggtgggtggatt cagctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
catcgaca ccggaggtca agttaccaaa gatatttact acaataagact ggcaagatc 180
cagagtaagg aaaaaagtgg ctaaatgtaa cccctcaact cgggacacaa tatttgaagt 240
gggagaagga gataaagcgc aagctggcaca ggctgtgagg gttccacagg ttgccottca 300
gagggagtgc cccggcgcg ggctgtgagc cttgaagctg ggggcgcgt gcaccacagt 360
ggtgacagt gttgagaggg aacgcgcctac cttggcgccc cttggacnga tactttcagg 420
gaacgcttt ctctcagcct ttctctcgcg ccctcaggg gttattagaa caaactcaga 480
cctgtcaggg tggccagaca acaactcaggg caagacactc cccacagagc aacaactgct 540
agtctcaccg aagctggacg ccaattttctg cttgctggcc atctacccat gaaactccccc 600
cctgtcagtg cctgtgctgag aagctgagcc cacoctcctgc cttggaacac ccagcctccc 660
gacgcctcgag cgccagacac ctctcagccgc cctttatctc ggtctcttga tacaagagcc 720
cggtttctct ccaagcgagt tgaactaatg gctagctgcc gggcocoaag cgggagcagc 780
aatctctct caacctccgca tcaacaaact gcctctcaca gccaatcacaag ggtgtaggaa 840
acgtgtaaaa gacgggtgctg cccggacacaa gccgactcgc cggcctggag tagtgggggg 900
gaaagccccct ggctctgtgt gttggactgc tcagttggac tgtcggcagtt agttggccca 960
tccggaggt tccctctcacc cgacgctagc tgtcagggcct gcctctcagg tgtggttgga 1020
ggggagtgtgc tacctctagt tgtctagggc gacgtgtaggg tagttaccaag aggggactgg 1080
gggagacccct tcctcagcgc aacagcaga ccgggacggc atctgactaa aagcacttccg 1140
caaaatotta gcggcgagcg agatgtggagaa gggagaggcc ggaacagctgg aagttggggg 1200
cctccagcygc gcggcgcagc gccgccctc acgccctcct gcctctcagg gacgtcagcagc 1260
caacagtggag tcggccaaaag agagaatatg ogggcagcagc caacoataac tgaagtccaa 1320
aacatagcgg gcacggtgctg gggcagcagc tagggcagttc cagcagagctgtt 1380
gtccacaaaa actcagcagcagc cagcagctagtg cggcggcttct ggctagcttgt 1440
cgtgataaacc gccttgatgc gcctgctagc tggctggccatttc agaataggtc 1500
agggatcgg acagcagactgt cggcgcctag ctcttggcctt gggcgcggcc 1560
caacacacacactgcaggtgactgc ttcggcctactct cctggcccagtg ctgagctctctct 1620
cggagcgagc cctctcaggg cttgagcgtt ctttttagct tcgtgggcttc 1680
tctcactcttaca accagactgg gcgtctcact agttgtctag tgaaatcgtca 1740
tctctctctc ctgttttatgc accagactgg ggtgccattc agtggttctctg gtaatattcgc 1800
gttgctgtc gcggaggggc ggttcgcttct cttggtggtt ccccattaaa cagactcttg 1860
agcaagtcgg ctaccagtgg ccggcgcacg aggggctggg accagagcttc 1920
gttggacaggt catccgagcgg cagcagcagctgc gatcctgtctt gggcgcggcc 1980
cctggccacgc gcgtctgcgc ggccgcagagc acacacacactgc gacagactgctgagcagtctt 2040
cacgcccata ctaaagcagcgt tcagagctgtc taactactcag caggttcgcttcgatctttg 2100
cccctccagg ccatctgtag tcgacgagctgc aagcggagct gcgtttgatttchtactttgggctcctggtgtcttgctaatcactgggtgctgcataactactagccagcagctgtc 2160
ttggctgtc gccctccgagc cactgtgcct cctgggaata gctagcagtc aataagccagcgctttctactgttctcttaactagcttttcagctttttg 2220
cgcttgtttctcagcgttctactggatgggctgctctctactgtctgtctcagcgtttccctgcctttctactgttctcttaactagcttttcagctttttg 2280
aataaagcttg gttcttctctg gtttttttttg 2340
aataaagcttg gttcttctctg gtttttttttg 2400
-continued

gcaaatctca ggotoaatcc ctctctctct tgtcaaatctca gacgtgtact tcaatcttatc 2460
aaattacgca ttctctacag tgtctctcag agcttctgact aagctctgtg actotctctt 2520
tttctctcttt tttgctatca atttagatga tgaagctga caaacatccga aagaactctcg 2580
ttttaacact gacgtgtatc tttttgctat tttatctctat atagttgtat 2640
aaattacact tgtatcttct gaggacctct ctctctctct ccaggtctgtc gatattttca 2700
tttgagcagtt gttattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<408> SEQUENCE: 349

atgcgcaagt ggggctgtcg ggcggcggga ggcggcggga ggccccctct gcggccccct 60
acccgggttc acaaatgcgtc aaaaaaaaaa gaatcagaaa ctggcaaaaa aatagattctt 120
tcaagagtt tcgctgtcag gcgtggttct ccgagactga gcaacatctga acacagttcc 180
cctgcctgtg atacatatt tggctcagtg gacgccatga gcagctttcc gtgggctgt 240
gatcagagt taggactgaa ttcagcggag gttaccttctg tggcaggttt atacaagacct 300
tttggataaa tcctggctttag tcgctgtcag aaatataaag cggataagaa cggagctgtt 360
aatatagtt tcattgatcc tgaacttac ccataccaag tttggaattag tgggaaaaagc 420
attcagtag tagaaaccca ggtagagaga gttatactgc ttcttttcaat tttggacag 480
cctttactact ccctactacta tcgtgctgtg gaaagaaaaag tgtcaggttt tgtaagagtt 540
tatggtgcaaa aacctctgcc aatctctctg caaagttttta ctcgacaggt tataagagtt 600
gaaacacac accggtcttt gcagacatgg atgaataata ttaaaaggaag ttcacgagcct 660
aatatatacc aatctctctg cgggacttac acctctctgt agctttcact gttcttcacc 720
aatatagtt tacagagttg tgaagacttac accatcagct tccccactact agctttctcc 780
agttggttc ggcgtgtttag aagctcttctc agtttactttt tttgagaaaaa aatctctctg 840
aatagacttc cgcagttcttt gaaagaaaaag gtcgaggttt gtcgttttttt cctctcaggt 900
gctgttacaag ttggactact tagagctttgc gtcgaggttt ttcagttttca cctctcaggt 960
gaaagagt tcagatcagat caggtttttc aatagatttg ccatacagaaa agggagacgg 1020
cgcaggttt agttggttc gtcgaggttt gtcgaggttt ttaaagagtt 1080
aacacagctg tgtatcagct gaaacacatt aaagatgaaa accatagatt ggttttttttt 1140
aataggtcttc tgtgaaaaac ccctctcttg ctcctactgt agaaacacat gctctctcag 1200
cocaaaggtgc tgggtcttaat atgcagcgtgc ttgacagctt ttttaaaagag agatcctaat 1260
ttggtgcaagt tagaaatatt ccctcagctgt gttgttttttt cctctcaggt ccctcaggt 1320
aagaagttgct atacatattt tttttaagag tggatcagttg tctttaagaag 1380
aatagacttc gtcctaggtg tctttaagag tttttttttttttt ccctcaggt 1440
ggacacttc cggctcggc tggatcagttg gtcgaggttt ttttaaagag 1500
cgtctgcccc gcaacactct ttcatcgccg ggcgtgcttc ttttaaagagc 1560
gctgaaatt ataaatgttt cttcttcact ttttaaagag tttttaagag 1620
ccagactctc ttcaccactc aagtttctgc gacgtagaat ttgatcagttg 1680
agtcttccg acataagaggg cttcagttct aacaccactc acatcagttg ggcgtgcttc 1740
ttgatcttc gttgtttttttt ctcctcaggt atccagctct ggggaaatag 1800
cagaaaaattc cttctcaag gatttaaat tttgacaag tggccttttc ggtttaag 1860
ccagagctc gggaaaataaa gcataaaaaa ggtttgtttg ggtttaagtc ttttaaagag 1920
aacatatttc ttgatcagttg ggcgtgcttc tttttaagag cttctcttcc 1980
gatctgcttc gcacaatttt gcgttttttc aagtttattc ggtttaagtc 2040
atatttttt gatcttccg cggctcgggt ggtttttttttt ccctcaggt 2100
itttttttttt cgcgtgtgct aacaccactc ttgatcagttg ggcgtgcttc 2160
cctttctcat acctcagga gcaaacatcc ttttaaagag ttttaaagagc 2220
ggccagcggga aagttttatt tacccttgttc aagaggaagt aataacgtga agtaaaggt 2280
gccagtttg tcgtgctcgt tgcgcagact tcgcttttct attcagggga acaagcattg 2340
atgatgacta ttgtaaatc ggctcagaca ttgctgggaa ttaacaaatc taacttgccit 2400
cagtctatgt gcctgcatgg acttacgtct cttttgagga gagaagcctgc aagctcctgt 2460
tatatttctca caagtgtaac caaattcgga aagcgtacct tcctgtgtgt ggattgaaca 2520
caccttaagt ttctttaaga tgaataacaa cgtctgagac ctgacttgtaa tattcoata 2580
attcctgagt ttattttcata ttggtgtcag ggcattgtga ctggatggct tgttataacta 2640
cccaacatag atgycttggga aatgtggaac aatgcacaac gactatcagcttg gccctggt 2700
tctcatcaca tggattggca cttaacaaac tttcaagggga gatttcgaag actttgctca 2760
aaccagtatg ccgatcagtt ttgaaatatt gtgtgagca gaacacacgc agaatttaca 2820
gagcttcttg ttgataacttt cagctgagca taaaagagaac cattttttaga aactttgctga 2880
aatggaacag ataaacaccc acgcatatt tctgatttt aagaattatac atacacacac 2940
actgtggatact ttggtgtcag aatgctggaag gacaactacg ccacagcaga agctgtgctga 3000
cgctataag ttatatattc tcacactact ccacctgtga atttactggt attttttttgt 3060
catagcttgt gctgtggagaa atatgaactt gcgcagcaga tctctaggaac atctttttgta 3120
attcagctgg gtttatacgg tttagctatg gatggtcttg ttgactgtgt ggagacgca 3180
tttataaaga tttataaccg aagctggttc atttgtaga agatacagaa gaaamattatc 3240
ataagagat aagtcagaaaa gcattttgatt caaattttag ttcagagaggg ttagttaatc 3300
gcccgagta aagctcggca aacgccacaa gaaaccacgc cagagagagg tgsaaacccac 3360
aaccagagac atgatagtttc atctcataac gcagactttt caacocctgg gtttaataaa 3420
attctattta gttctcttttg tctctcttact aatgaaaagag ttgatgaaa agttaaaccag 3480
agagatgcac caagggcagga gcgactatgt ttaaaaagaa aatctccttc agatcttttg 3540
aagaggatgt tgctgtctaat tttgaagagaa atttccccct gcagcagagc gcttataaatc 3600
agattgtctgg ctgaaagtgc ttgaaagca aaataagaga aatttggcag aacaaaggt 3660
aagacactcc atggtgaggg gcaaatgcccc tccactttgg gcggaagaaa aatctctctg 3720
attctcaga taagaaaagaag tcgcagaaaa aagttgctga aagaagaaag ggggagacct 3780
gatcctatgt cgctttttgt gaaatgttgat ggaatgatca tgtgctgacc gttgaaaggt 3840
gcagagagaa ggcctgtgac ttgctctatt cttgtaataa aattttccca acaattcaag 3900
gagaaggag agctttgtag ccgatgtaga aaaaaactct tttctctttg taaacactag 3960
gcagacagaa tgaagaaacac gcctccttgg agtatttgat agaacactgc agaaagttt 4020
ttggaagaga cggcactctgt gttcatcaag aggattagg agcagagacc 4080
gaaagacacta atacacaatt tcatattcga gaaggagag atggatgtgc tgtatgtgtg 4140
gatagagaca ataattttttg aagaggaatg aagatcagact ttctccttcg ctatctctctc 4200
gggagagat aagggtttcc tttgaggtttt tgataaaaaag agaatcattct attttcacaa 4260
gcagaaatca aagcacactgc aagaaatctt ttagactaca aaaaaaagctc ggttttttga 4320
aatctctctct ctcatcttcct atattcctag aagctcagag atgtttcagc taatttttac 4380
agatagagag aataggccct ttgttgttct ttcacattcct ggtgctgaga ccaagcaaac 4440
aagttcacta gtaaaacggg agcgtctcata aaaaagacac cgtcttcgaga tacagtccc 4500
-continued

aagcocaaga gacccocaag acagaagaga gtttgaaagg cttgataaca tgtctgatat 4560

tgagatgctg doatcaacg acagacaca gaccaaaag gcaagtcgct gcgaggcaag 4620

aaaggaagg cttcctggct gcaaaactaa gggcgattaa acccttggcg gaaaacatat 4680

aaacaaaac acgacacag gcagaagaga tctctggctg acagttcaca attgtgacat 4740

tccgcttctg acaccaacg tgtctggccac gacccgctcg ggcctgaaa 4800

ggaagaaac aatttgcggta ggtctgaaag gaaagagagt atatggagtt ttgcaatgttt 4860

aatctggc ccacagacag cccacatatt tcacacaact tccagtttgt ccctcttttgtc 4920

tctctgctt cacacattgt tacatctgctg tctttatcgt attgtgctt aatgctcgct 4980

atatataat ttgaggctag ctttttaacact tttgtctcta caccatacagt tttagctct 5040

attttactca tttgacgtc actgtacgct gtctgctctg gtaaatggt ttatagctgc 5100

cgctgacatg acagacattt aattctcgct gttcaacaact acacaacgct tttttggaat 5160

gaatttcaa ccttaaactc ggaacctgct 5189

<210> SEQ ID NO 350
<211> LENGTH: 1536
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 350

ggggaggggg ggacctgctg gctgctctcc ttgctgcgcccc gcacattggc ctgctcgcgtgt 60

coccgocggc caatccgctt ccggcgcccgc gcagcactccc ggcccgtcgc 120

tcggcggtcc gcgcgctccgg ctggtctctcc gcctggccgg gcggggcgcc gcgggtggtt 180

cggcgccggc ggagccggtg gtaactctgg gatagctcag cagatcaca acactctcctc gtcgaggtgc 240

gagaagggc cccgcattcg cccggccacc ccgctggact gcggcgcttca gggcgggttcctc 300

agcttcgag tagaactccct ttcgctgaca acacacgcttg tttctctggaag gggcgggttgt 360

gcgagggct ttttgcttat ccagggactgc aactactgct aaagctggc ttcaagggaga 420

gccggttaga tccgatccttt tcgctgcgaa cccaccggtg agagccactg agtctctgaa 480

agcagccggc gggcgggcttg cccgttttact gcgcgcctcc ggcggctttt tgggggagaa 540

tacaagatga tgcgactgta cggagagtctg aaccactgca aagaaggtct 600

gacacacacca acaccacccc tccgagagac acgactaca tggctggtctt ggaatccgag 660

gacttcctcc ggtgctggct gctgccatcc ccgctctctt gcctgtctct gctgtgctcag 720

agcgaggttg ccctggctac ctgagcaacgc cggccgctgg gttgcttggctc 780

aacatgcgca gtaactcgcct gcaaacagct gacactccgg ccgctggaca 840

ggtgacagca tcaactcttt ggagagctg ctcatttacc tcaatcagcg ccggacactg 900

aacaattcag gtcagcaggtt aaccgactt catctacat tcaagtatct tacaattagtt 960

aggagaacgt atcgactccca agctgagaca tttttgattg aaggttcacca tccgactccca 1020

acttatataa cccagagttcg tttttgggggt gttctcagag gacgattgag gatttgtctt 1080

gttctcagag ggtgggctgt gtcgctttctc aacctgctgcct gcagtttttt aagtttggat 1140

cccactgtc acaactgtcct cgccacacctcg cagtttttttttt ctttatggaa tgtgctctct 1200

gcggcagtc aacctctcttct cttttcccattttaaagggg cttggttcat ctcctgtgcag 1260

gttgctgcgctg atgctggtgcct gatatcagcc agtactgctgattttttt ctttttttttttt 1320
---continued---

cacaacactg tataagcatt tcaatccocac gtacttatata aagaggtgag tgaatttcac 1380
aagctatatt ttggatatatt ttgatatatt ttgaagtatt catcaagttgt tacctcaaat 1440
cagggggaccg tgaattccac catctgacgat gatgtaagaggt ttgagttataa acctttatgt 1500
tgtggtatat gttgccattg ttcacgtg 1536

<210> SEQ ID NO: 351
<211> LENGTH: 2386
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 351

ggagggagagc gcaagcggagc ggggtgcgtc tctagttcag caattctgtgt ttggtgcttc 60
ggctccagcg ccctcgggct gctagttcct ctcgacggcg aagctgttttt tttggcagcc 120
aacgcctgggc aggctggagtt aacgctgtcc cggagccgcc tctctctctc gaccttgcttc 180
acaaaagatgb tccgggggct tcagctgccg tcgctttgcc taaggcgtat tgaatctctc 240
tccctctct tccagctcttt atctgcgttct tgtgatcagt tagagcagttc ctaaccaagag 300
acagacatct cattcgtgcag agttgcttttg caaattgagc tgcggagttc gtcctcatt 360
atagagttct tccagggcag cggaggtctg gtttgtgtac tgttcttccca caagttctcc 420
cagaaaaagt aacgcctgggct tcagttcagt gatcgttgag aagacagcagc tgcgggggaa 480
aggagcgct tccgagcagt gagaacggcc atggccggtc agtagctcc tggccacgcc atcccacgcacc 540
cggagagac ccggggcagg tgaagcgaga actgtgcccag cggagagtttg gggagcggcg 600
aggagcgct tccgagcagg cgaagcggagc ggggtgcgtc ggcctggagg cgcagcttgac 660
cggagagac ccggggcagg tgaagcgaga actgtgcccag cggagagtttg gggagcggcg 720
gcacaacactt ccagccggagc gcattctttt gttgaacctct tgggaacagc ttgcagcgcag 780
gagagcgct tccgagcagg tgaagcgaga actgtgcccag cggagagtttg gggagcggcg 840
aggagcgct tccgagcagg tgaagcgaga actgtgcccag cggagagtttg gggagcggcg 900
taccccttccg tggccaggtct ccatccctg tggcagccag tggccagccag 960
gagacacact tccgagcagg cggagagtttg gggagcggcg 1020
tccgtgagc ctatgttcttc acaattcctgt tggcagccag cggagagtttg gggagcggcg 1080
cctttttctc cagacagctcc acacattcag ccaattctgt tccaaagggag cttggccccgt 1140
tgtgctgttt gcgggtttgc ggccagggccc ccctacatca ttggccagcc cctttttctc 1200
tagaacacc gccctgggttt ctcgccaggg gcattgcagaa gtttctccag ccccttttttt 1260
agagagagc ctagagttgc ttggtttttg tttttttcag ctagatgttg atcatctttct 1320
cctttttgtc cccacccatac ctcctgggcc ctctccctc acctccacac cctaaaggcag 1380
atagagggcc cctttttgtc ctagagttgc ttggtttttg tttttttcag ctagatgttg 1440
gctttttgtc cccacccatac ctcctgggcc ctctccctc acctccacac cctaaaggcag 1500
agagagagc ctagagttgc ttggtttttg tttttttcag ctagatgttg acagacgacc 1560
cctttttgtc cccacccatac ctcctgggcc ctctccctc acctccacac cctaaaggcag 1620
tctgcttcct tggcaactcct cagagaccc caagttccct gcagagcttc ctgcagcttc 1680
tttaggctcc agtaggctcc cttttttcag cagagaccc caagttccct gcagagcttc 1740
gctttttgtc cccacccatac ctcctgggcc ctctccctc acctccacac cctaaaggcag 1800
ctatggagac cccagggctct tccctacctc aggcaaggaag ggcaaggaag agaagctgtc 1860
gcaggtggtt gggagttggt gactagaagg gcagctctcg cttgcccaggg cagatctgctg 1920
coccatgtgct tcagcaagtgt gcagagcggc tggcagagag gcaggtgcag cggcagagag 1980
cctcttcgag ccctctccct cttctgggttcc accatgtgccc ggtcttcattcc ccaaggggttc 2040
cagccacccc cggcagctctg cttgacatcct ttgagactag ttgatttaccc ttgtgaagagat 2100
gatataactatt ttggtttacaag cgtgctgtgtg ttttaagttg aggactgtcgt ggttggcagtg 2160
gcctgtgtcc gcagctgtctgt gcgctcgccg gatgagccgg cgtgctgtcc gctcctcctcg 2220
cctgctctcc gcagagagag cggcaggttcct gcgtgctctgc ggctgctcgc ccctcctcct 2280
accccccaccc cagactgtccg ttcagcctttg gaactagcgg ttcgagacggt ggaaagggcag 2340
ttcagtaaact cagctgttttc actcagtgaa aaaaaaaaaa aaaaaa 2386

<210> SEQ ID NO 352
<211> LENGTH: 1270
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 352
agcgctgagcc aacacttggtgt gcagacgccc cagaggttccc cggcaagccc gagggcgccg 60
gcagagcggc ccagagctccgc cagctctcgc agaagcgggac acactcctcgct ttcctcctcg 120
cctctccacc aacacttggtgt tttggccgca gctcgtgctcc cagagacgccg gttgcccagc 180
agacagagac gcagagctccgc cggcaagccc atggcgctcc gcctcctcgct gcctcctcctg 240
ggggtgctcg ctgccttcag aacacttggtgc gcagagacgccg gttgcccagc gttgcccagc 300
gaccaactct ccgggaaagcg tggagacctg tccgagttcag aacacttggtgc tggagacctg 360
gtccttcaggg aaagagcgt ggtctgctgct cccctgcaggg tggagacctg 420
tctagatgtg aacacttggtgc cggcaagccc tggagacctg tggagacctg tggagacctg 480
coccaaacat ccggctatag ttcgctcagag gcggccaggg tggagacctg cgggactcag 540
coccaaacat ccggctatag ttcgctcagag gcggccaggg tggagacctg cgggactcag 600
ggggaaatgc aaagagcgcga aagacagccta cccctgcaggg tggagacctg 660
caaattctgt ccatttggagc aacacttggtgc tggagacctg tggagacctg tggagacctg 720
aaaagctcag aacacttggtgc cgggaagcgcc tggagacctg tggagacctg tggagacctg 780
atggagctgc ttcagtttccc aaaaaaacctcc ttagacgcgg cttagcgcgc ttcagtgctc 840
gtgcctctgc aacacttggtgc ttcagtttccc aaaaaaacctcc ttagacgcgg cttagcgcgc 900
tatgagcagc aacacttggtgc cgggaagcgcc tggagacctg tggagacctg tggagacctg 960
atatgacctgc ttcagctgaat atcaagatcag gcggccaggg tggagacctg 1020
cataatgctgc ttcagctgaat atcaagatcag gcggccaggg tggagacctg 1080
atgagcagc aacacttggtgc cgggaagcgcc tggagacctg tggagacctg tggagacctg 1140
agcggctgc gccttggcttc cagagagcggc tggagacctg tggagacctg tggagacctg 1200
cttatttacc gcgtcatcctc ccttctttttt tatttttttt tattttttttttt tattttttttttt 1260
aaaaaaaaaaaaa 1270

<210> SEQ ID NO 353
<211> LENGTH: 1600
<210> SEQ ID NO: 354
<211> LENGTH: 1842
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 354

cagaagagag acatacgatg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
60
cacagcagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
120
cacagcagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
180
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
240
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
300
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
360
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
420
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
480
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
540
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
600
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
660
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
720
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
780
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
840
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
900
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
960
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1020
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1080
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1140
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1200
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1260
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1320
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1380
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1440
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1500
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1560
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1620
tctcgatcga tgaacttctgcc agaacagctg gtaatatagg tctcgatcga tgaacttctgcc agaacagctg gtaatatagg
1680
-continued

```
aggtaggtta ttgagcctag tgtggagagat gagaagaaca aacggtgagg accgtggcct 360
attacattga aacacacacc gacgaagaat gccttgcctt gctttttgtc agagacactt 420
cctggctctta ttggaacat tttggtgacat atcgatgtgc atgtatggtc cgc tgtggtt 480
aattgtaactg ctaaacagct taagttgctct atatctgca tcatagcag gctaacagct 540
gctgttaaac atataggagg cgtatacag gagaggttag aacctctccaa aagattactg 600
attgtttaactc agtccttacgc agccgagcat aactagccgg gttccactc aactcttcct 660
ttttggcttg aagaagacac ctttctggtct tttctcattag aacgccggt caaacaatcga 720
gatggggcg ctgacaaaag gctctctcgt aagacaggt ggcctgtctt aaatttgatg 780
ttcctcctctta aatgtaaaag aagtaatctg gcccgttgct tttctctctc 840
gctgttccatt ttttttttt tttttttgta gactttccac tatcttacta aacgagatcc 900
agcactcact ccgcatcaact ataaatgtgt ctctggaacc cttgttatata gtcttgtgta 960
attacattga attacactaca gattttcct ttcacatca gttcggaaac aacacgccat 1020
ttttaactatt cttcgggga aacattcttt tgttctcattt tattgaggga actatatttcg 1080
tgyttgaaa gattatgtct catcagttt tctagcattt tcttctgttt tttttttaac 1140
cattgttctc gtctgtactt cgtgtaagcg tggcatagttt taatattttt gttggcctac 1200
tttttttact ttttctttatt gtcttgactt ttttttttt tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

tctactcatc caggccaa cctcgcggtg gtcgcggagg gccaggtcct cgatgggaag 540
tttgcactt tcycactggt gaactataac accaactcga gcacaagctc ggcacgctc 600
cggtgtacac ctgctacgga gattcttgct gggggttttt atattgagaa gaagtcgtaa 660
cctgctcaca tcgctcggag gcctcgtta gcgtggcgtt ggcacgctc gtcctgata 720
gttggactg acatagcag aagotgcccc cctcgtctag aggttggcag ggcggcagtgc 780
tggcctcctg gattcagaga ctcgcaagca ttcggcagca cacatcctgc tcctcagttg 840
aatggtactc gcttggtgcc ccaaccccaac caggtgctgc agttgagaag tgtgcgggggc 900
tgcagccccc tccgacaagc gactctcgc gctggtggcg actctagctg agtggagagcc 960
tgcagcctc ctggcacaagc gactctcgc gctggtggcg actctagctg agtggagagcc 1020
aatcccaaca ccagctatca gatttggag gttggtagcg gcactgtccc ccataactctt 1080
gttgggtca aacacttctc tcycactggt gtcctgactc acacagtagcg aagctgataaa 1140
aatcgcggtca gatgtggtgc gcttgggag gcatactgca ccaacgcttg tagggcaca 1200
gttgggggc cgcctcctcag gacgtgagac tgtcatctcg cagacactcg ggcagcaga 1260
accagggcct cgggacact cgaattttcg actacagggc ctcaatagaga ccggggtgct 1320
aagacatgct ccttacctcg ccagatcctcgg cctgcatcttc ctcggggtat gcacatagagc 1380
 ggctccctg ccagatcctcgg cctgcatcttc ctcggggtat gcacatagagc 1380

---continued---

---end of document---
gtgggtggcc tggcgcctcc tggagataag cagctgcctc acagtgaggc caagacccca
atgaaggtgg ggcctggctt gatctggtaa acacacacc aagttgatgtc
2820
tggagctctg tgtgcagttg gtggagtttg agcactctgg ggccacagcc ctatggcagg
2940
tctgatcgct cggagcgact agacgtgcta ggaaggggg aagcgtttgcg acacgcccaag
3000
atctgccaa tgtgtagcct caggtgatgt ggtagtagta tcagggatas tcaggggatg
3060
cggcccaacct ttaaagaaat agcccaagag ttctccacacct cggccgagac ccacacacgg
3120
tatctgcac ttcagagaga gatgtggtcc ggagttgtcc tggggccaga gcccctggtt
3180
ttgaaagaca agaagaatga ggaagagtga cctggaccgc aacgtgacag cagcgaagag
3240
tggagagcg agaagccagc ctaagcagag ccagacaccc gttcggcagc tgggtgacag
3300
tggaaacac tttaagcggc acagtgagga cagcgccttc taaagacaaca tctggtgata
3360
agtcacatga acagattgga ctgcttgatc tctgagttg agtctgtcag tgggagactt
3420
agatgcatgg gggccaggtg cagctctctg tgcagaccac gtcggtgct ctgctggagag
3480
agatgtgctc agggaggctg aacagagctc ggtctgtgag cttccagggg aggccctagc
3540
tgtggagcc gcagcaggag ccggccagcc cggccgcctc gggaggagtg gggccttcgctg
3600
ccggcaccac cggctgtgtgc ctggctctcc gctgctgtgc ggcgcggcgg ggagagcccg
3660
gatgtctcaag gttaggactg gccgatgaca cccagccgct gattcgctgg cgggctgctg
3720
ggaggacctt cttcctgttc gctcggtgta tcgaggaagc gtgataagag agatagactg
3780
aggggtgttg aacacatggc cggcaggagc aagacagatc cccctgctcc ccggggcccg
3840
atggccctgtt agggaggctt tgtagctcgg atgtgtgctt ggtggagcgt ctagctctcc
3900
tggggccaga cagagagggt cccatccttc cttggtcttc gcggcggcgt cgggtgctttc
3960
actgcaagtt ggaaggtgta ggtatgagtt ccagagcttc ggtgggtttt tcggagacgg
4020
aggtatcag ccctgtggggt gttcgccact gcctggaggc ggtatagagc gatggagctg
4080
tttcggctgg cttgcaatgc ggcgcacact gctctttggt cggcagaggg ccaggatcctt
4140
agctgatgg ctactgagtc tgcggccttc gacttgcttct gctcttgcag gctgtgctctt
4200
cccagtcttc atggagcacg acggtaacct cggatctgct cttgctagcg ggcggcacgtt
4260
agttcaggt tgcgtcttgg agggtaacct ttctctctta tctcctctct cttccacgctg
4320
ccagcgcctc ttcctcccagt gacagcactt cctcccaaat tctggagcct ttttaacatt
4380
ctgacacaa atctctattg tacgtgcttg gttcctgagt ttcctcttcc tctaacccca
4440
agggaggtt ttcctctattc tgtgtgtttg cctcagcattcg tttcatcagaag tttcaagcg
4500
caactgatg catsacagag acctctccat ctctcagcct ctggactatc
4560
tggagcattg ctcctctttg cttcctgctg gctcagagga cggagagcagc
4620
agggaaacta ggccgaggaat gtttttagc atccacccct tctagaagagc
4680
agagaatcata caaatgtctgg ggtttagct ccttacccct cctagaagagc
4740
agcctcctc actgaggcct gcaagagtact catataacct tattctagt cactatgccctt
4800
tttatagctt tttttatccct ctgctgtgctg atcataactc ctaattttgc gaaagggaggc
4860
ctgctgcttg ggtctgtctc ccgtgcagct ggcagctcctt cagcaagcttt ggaagagctg
4920
agcagacact ttcgctagct ttcgctagct cttctcagct ggcagctcctt cagcaagcttt
4975
-continued

SEQ ID NO: 356
LENGTH: 4627
TYPE: DNA
ORGANISM: Homo sapiens

SEQUENCE: 356

tcacttgct gatatttcaag gcgcagcaca acgtggggct ttcctotaac g 60
tgacagcog tccagcaaca tggccacagc cggggtctgc cggggtctgc 120
tcagccggc cggccagcgg ccggccagcgg ccggccagcgg ccggccagcgg 180
ggggtgctgc agtctgggct gcggccagcgg ccggccagcgg ccggccagcgg 240
ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 300
tgatagcgc agctgctgtgc ttcagagccgc acgtgggtgac ggtgctgctg 360
aacctgctgc ctgagggctgc tttttggctgc gggagccgggc ttcctotaac 420
ggaggagctgc cggaggtggtgc agtggtgggtgc acgtgggtgac ggtgctgctg 480
ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 540
ctatagtgat cggaggtggtgc agtggtgggtgc acgtgggtgac ggtgctgctg 600
saagctgaccc gcgcagcaca acgtggggct ttcctotaac g 660

4926

tgagccaca tggccacagc cggggtctgc cggggtctgc 720
ttacccagc atggctgatc atggctgatc atggctgatc atggctgatc 780
cggggagtgc tggaggtggat cggggagtgc tggaggtggat cggggagtgc 840
ttcagagccgc ctgagccatc ctgagccatc ctgagccatc ctgagccatc 900
agctgctgc ttcagagccgc acgtggggct ttcctotaac g 960
aatagctgc agctgctgc ttcagagccgc acgtggggct ttcctotaac g 1020

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1080
saagctgaccc gcgcagcaca acgtggggct ttcctotaac g 1140
	atagctgc cttggaggtgc tggaggtggat cggggagtgc tggaggtggat 1200

ttcagagccgc ctgagccatc ctgagccatc ctgagccatc ctgagccatc 1260
agctgctgc ttcagagccgc acgtggggct ttcctotaac g 1320
saagctgaccc gcgcagcaca acgtggggct ttcctotaac g 1380

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1440
saagctgaccc gcgcagcaca acgtggggct ttcctotaac g 1500

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1560

9827
	ttctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1620
	ttctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1680

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1740

6144

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1800

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1860

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1920

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 1980

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 2040

ttcctgctgc gtcaggtcgc ctgagccatc ctgagccatc ctgagccatc 2100
tcagggagca tctgttttctt tctctagttt atcctggcact gctctagggtt aaacaggaagcc
2160
atataaagct atggtgtacc tctggagcact cggggcgccg gactgtcctgt catgyttacct
2220
catgcaatt tattctgtgt gcaatcacaacctccttgctcttgctcttgcaaatgtctttctgc
2280
ttaatttcag gaaatagcagc gaaatagcagc gacttgccctttatgcttctatagctctgtttg
2340
tatatatatt cattatactat catctgtctatttggttctgtgttactgcctactatgacttgactgtatgacttctgtttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
aacataaagt tggattggcat gcaatattaa gtaacttatt tgactatgaa tattatccga

ttacctgaatt gtatcatttt ttttttttct taaaactgtc aaatcttatt tattatatt
gattatgag gagaatatag ataatattaa agatatattt aatatatttt attttttttt
ggagattaa aaaaaataaa ataatattaa atgactatgaa cacttgcag tcaaatcttt
cactgac

cagtac
<210> SEQ ID NO: 357
<211> LENGTH: 2634
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE:

gggtagagagc tggattgcgc tttttttttt ttagggggaa ggattttatt tattatatt

tagattgtggc gacagggggttt ttgggggttt ttatatttttt tatattattat gatattttttt

tttttttttttttttttttttt

<420> ACCESSION: Y11448.1
<421> ORGANISM: Homo sapiens
<422> SOURCE: Chromosome 20
<423> LOCUS: TTN
<424> FEATURES: "antisense"
<425> LOCATION: (1790574..1791907) on chromosome 20 short arm (p)
<426> STRAND: minus
<427> QUALITY: 73
<428> DATABASE: GenBank
acacoagt gtagtatctg aactaactaa agctaattaa gaagagacga tgtttggaac 1800
"<210> SEQ ID NO 358
<211> LENGTH: 1246
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 358
"gaacacgctta cacagccttg cactgtgtac cacacccccg tcacagcctg cocagaagcg 60
cgccccccac acagtgcct caagttgagt ccagtgagt ttcagtcag ttttaaaactt ggcttgtagct 120
aalagggag gcacaggtct gacacactt gatgctggta ttaaactgat ttccctctgt 180
cgcagagctct ttagctgctg gacagcactt gttggttgct gacagtctact 240
agcttactg acaagctag ttttagag caagttgctg cactgtgtat ccagtgagt 300
cgtgtgtactc gccactgct ttagctgctg gccactgctg gccttctctct ccagtgagt 360
"gacacgctta cacagccttg cactgtgtac cacacccccg tcacagcctg cocagaagcg 420
cgccccccac acagtgcct caagttgagt ccagtgagt ttcagtcag ttttaaaactt ggcttgtagct 480
aalagggag gcacaggtct gacacactt gatgctggta ttaaactgat ttccctctgt 540
cgcagagctct ttagctgctg gacagcactt gttggttgct gacagtctact 600
agcttactg acaagctag ttttagag caagttgctg cactgtgtat ccagtgagt 660
cgtgtgtactc gccactgct ttagctgctg gccactgctg gccttctctct ccagtgagt 720
aalagggag gcacaggtct gacacactt gatgctggta ttaaactgat ttccctctgt 780
cgcagagctct ttagctgctg gacagcactt gttggttgct gacagtctact 840
"gacacgctta cacagccttg cactgtgtac cacacccccg tcacagcctg cocagaagcg 900
cgccccccac acagtgcct caagttgagt ccagtgagt ttcagtcag ttttaaaactt ggcttgtagct 960
aalagggag gcacaggtct gacacactt gatgctggta ttaaactgat ttccctctgt 1020
cgcagagctct ttagctgctg gacagcactt gttggttgct gacagtctact 1080
aalagggag gcacaggtct gacacactt gatgctggta ttaaactgat ttccctctgt 1140
---continued---

```
taxaatagata cacaatgcata caaatatttta aactcaagttg taacttatata tgaagcaaat

tgatagcata ataaaaaatattcgcagatgt cnaaaaaaaa aaaaaaa

<210> SEQ_ID: NO 359
<211> LENGTH: 2360
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 359

getaogcgcc caacgctgct ggtggtctg aacctagggc goggggtcgg gogggcgccg

60

ggggcggcct gtgagcaag acaagacct caaaggaagc gagtctcgcg ggtgtcccg

120

ocagcctgag aggccgaggg gggcgggaga cgggtccccg cggattctcc cggatcccg

180

cggccagctg tgcgtcagcc gggcggtggc ccgcgtcctg caagtgcctg gtgcctcgcc

240

gctctcagct ggtgcgggac caagaagctg cttgctgcgg tttgccgtga gtttttctct

300

gtgcgcagtggctgcagcgc gagaagcctgg aggccgcttcg agggcgcctt caggggctta

360

gtgcgcgagcc caaggcaacc ggcaacctccc acgtattcoga cggccagctg gttccctcca

420

gagagccgcccc gggccgagcc agtccctggac ttcagcaggg cagacattga ccttttcga

480

gtccatctct ccaggcaccgc acagcccagct gcaccgccca ccaacaggca gcaagggaa

540

gagagagaaa aagggcaggg gtcaggggag agagggcagcc cttgctcctg caaattaag

600

gctcctgcac cctgccttctg agctacaatat tggagggcc ccctggctcgc ctcctgctac

660

tgcgcaccctgg ttgacccattt gattgctg gctgctgctg atgcacagtc gggagctgtg

720

ttatactctt atgcccaac ccacatccttg gcctggcgtg cttggcgttcg gttcatctgc

780

tgtcgtggg tctcttgccg gttttcatgt tttaggcttc ataggaggg aagttcatgat

840

tggagggcctt gagagagaaag gaaatgggag gcctcctatgt ctcagcaggtc gaaaatcog

900

cggcgcccccc ctcagcctgc cggagcctg gggctgtgct cgggtgtgcc ctgtgcgtgc

960

gagggaaggg cccttcggtct cagcaagcc cccttcgctg cctgctgttcg agaatgtgtg

1020

tgggccccct ataatgctt tgcacaataa ccagacctcc caagttcgaac caagatgtatg

1080

tgcgctggtct cctgcttcag aagggggaac cctgatgtgg ctcctgtgcc cccctgttcc

1140

tcccaacacc ccccaaccat ccccaatgaa ctggggtttaa acctcaattc tctggatattg

1200

tttcctctt gatgggtgtt gttgattgctt cttataattt atatatttatt

1260

gagagagaaa aaggggcttc ctggggttga cccttggttg gttttgcggc

1320

atctaatcc gccgcctctt ctacagccgcccc ctcttcttctt ctttcctttt

1380

gtgatctgctt ccaaccaac ccacatccttg ggtgctgccg gacagaggtg cggaggtgtct

1440

caccctctct atatgggttc gcagagaatg ctcgtattat ctaggtgtgc

1500

tatattactt atatgttgctt ccttcctctt atatgctcctt atatgctcctt

1560

ttagtggtgatc ctgtagctgt gctcttctg gtttcttctg gttttgttct

1620

tgcgcccattctg ccacagcttc ccacagcttc ccacagcttc ccacagcttc ccacagcttc

1680

gtagctgttg ctgctgctg gtaaagagaa ccaggtggtgc gataactgtg cggaggtgtct

1740

tcgctgctg gtttcttctt ccttcctctt ctttcctctt ctttcctctt ctttcctctt

1800

tctgctgctg gtttcttctt ccttcctctt ctttcctctt ctttcctctt ctttcctctt

1860

cctgctgctg gtttcttctt ccttcctctt ctttcctctt ctttcctctt ctttcctctt

1920
```
attccggttag ctttcctgtgt ccttcacccc ccgctgccca ggccacagcg tgggaactca 1980
cctcccccttg tcggcaacaa tttctctacc cctgcttatttt cctctctgctg tacccagc 2040
aggggtcagc gcacagcagg acaggtctgg gaaggttatca gcaagacaaa agggctagaa 2100
gggacaggggc accttgcggcg tggcctttct cggtaacctga ttcccttgccc aattgctaccg 2160
agaaggtttgg aggtggggga ggctttgtat acctccaccc acctccacsa acagatgaag 2220
gtacggctgc atggctctct ccggaagttt tcggcgcatt tcctgaactg ttaacaaccttg 2280
tattcccaccc cctgtcttct cttatatactt tgcaacatcc caataagtaaa ccctttttccc 2340
staaaanasa aaaaaaanan 2360

<210> SEQ ID NO 360
<211> LENGTH: 1433
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 360

atcggggygag agggagggag ccagaagcgg gaggagggct cccgctcgca gggcctgqca 60
cctgcgogcc gcgcggtctg cttcgctgcg ctcgctgacg cgctgctgac gcgcctgctg 120
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 180
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 240
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 300
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 360
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 420
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 480
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 540
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 600
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 660
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 720
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 780
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 840
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 900
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 960
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 1020
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 1080
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 1140
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 1200
cctcgggcygag gcgcggtctg cttcgctgcg ctcgctgacg gcgcctgctg 1260
ncggtgtggtg gcgcctgctg cttcgctgcg ctcgctgacg gcgcctgctg 1320
ncggtgtggtg gcgcctgctg cttcgctgcg ctcgctgacg gcgcctgctg 1380
ncggtgtggtg gcgcctgctg cttcgctgcg ctcgctgacg gcgcctgctg 1433
<210> SEQ ID NO: 362
<211> LENGTH: 2756
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 362

atgctgtct tacgaacgcc gacacggtac gacgacgaga cggcgtaca ggattcactat 60
gcgccttgca cacggcctgc ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 120
agcgcgtaca gccggtttgca cggcgtaca gccggtttgca gccggtttgca ggtgctgcgt 180
gcgccttgca cacggcctgc ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 240
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 300
gtcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 360
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 420
gtcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 480
gtcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 540
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 600
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 660
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 720
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 780
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 840
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 900
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 960
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1020
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1080
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1140
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1200
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1260
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1320
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1380
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1440
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1500
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1560
tcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 1620
atggtaattag 1632

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 362

atgctgtct tacgaacgcc gacacggtac gacgacgaga cggcgtaca ggattcactat 60
gtcttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 120
gtcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 180
ggcgttctgct ggtgctgcgt ggtgctgcgt cggcgtaca gacgacgaga 240
aagaaggaac aaaggtatcc ttatatattc aatacagggt ttgcaagac acggagatta 300
tatgatcagg atttccattaa ggacagggcc aagagttccg aagaccccaaa ctaactgttcct 360
gacagagcc caagtcgcaac cccagaggtat gctggagccag aagagttcga aataattttcgc 420
tatgatcagttctgacagttcctttgctggaagtactggc aagagttcgc 480
gccaaagagctggcagctgtgtgttctgtgaaagtctgtgatgtcgctttgaccatttc 540
gacagagcc caagtcgcaac cccagaggtat gctggagccag aagagttcga aataattttcgc 600
gacacactata ccaacactata cccaaagctctttctgtgtgatctgggacagcagacgctgta 660
gagatatttttgctgtgctctctctctctatcataaaagggtaagagattggattgattggatt 720
tctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggattgattggatt 780
tctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggattgattggatt 840
aaagccactgctaatggtggtgctgtgcttcatctctctatcataaaagggtaagagattggatt 900
aatcagatgagatccgtggctgaataacttgcttttctgggacagcagacgctgta 960
cctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggattgattggatt 1020
atatggattt tcagctgcttaa aataacttctgttttcttttcttcacagcagacgctgta 1080
ctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggattgattggatt 1140
agcttctcttcagcagctgtgtgtgctgtgcttcatctctctatcataaaagggtaagagattggatt 1200
atatggattt tcagctgcttaa aataacttctgttttcttttcttcacagcagacgctgta 1260
atggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 1320
tctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggattgattggatt 1380
aatcagatgagatccgtggctgaataacttgcttttctgggacagcagacgctgta 1440
aacacccacta cccatgttcctt cctggtggtgctgtgcttcatctctctatcataaaagggtaagagattggatt 1500
gtggcagaaactggggtgacgacggtgttcatagcccttttttttcttttcttcacagcagacgctgta 1560
gccagcaac aatcaggtcgctgtgcttcatagcccttttttttcttttcttcacagcagacgctgta 1620
ggtaacaact ttccagagagctgtgcttcatagcccttttttttcttttcttcacagcagacgctgta 1680
gtctggtggtatggtggtggtgctgtgcttcatagcccttttttttcttttcttcacagcagacgctgta 1740
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 1800
tggtggtggtgctgtgcttcatagcccttttttttcttttcttcacagcagacgctgta 1860
aatcagatgagatccgtggctgaataacttgcttttctgggacagcagacgctgta 1920
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 1980
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2040
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2100
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2160
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2220
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2280
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2340
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2400
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2460
atatggggtgacgaggtgttcatagcccttttttttcttttcttcacagcagacgctgta 2520
ctctttttgctg agtatttcg aagaaggtgag ggaagttct gttgtct ctgctg 2580

tcgagagct ttttattgtt cctccttc agagagactc ttttttttttt gtttttaag 2640

cygggac ggggagcc caaagcttct cttttttttttt gttttttttttt gttttttttttt 2700

gttttttttttt gttttttttttt gttttttttttt gttttttttttt gttttttttttt 2750

<210> SEQ ID NO 363
<211> LENGTH: 2768
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 363

tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
---continued---

ggaatataaa aggaggtgac caaatacagt tgaggagga taatttggctt ctcctgccttct
1860
cacggygca actttccao cccagctgag cacatatta tctctcgtag ctcctgataag
1920
agaatcaaca ttttcctaat atttttcctaa caaatatata ataagaaataa gaaatttttt
1980
gatggctcta acmatactat ttatatacca tttttcttct tgggtatatc ataagttttta
2040
tgttaaatca ataaaaaaccttttatata cttatatacg atgttctctc gocataaag
2100
gagaaatca tagaatggaa tggagacccaa caagtaaata tttttgcttc ttttactcac
2160
ttgtggtgctt ggcccccttct cagaactaga atttttatat tatactgttag gaaagtttaa
2220
cacttctgtg ctctccacttt cttttgagat atatatatacct tatttttgtt gccttttttgg
2280
atcagagcaca ctgacgcacq ctcttaactc tagaagttagg gagaagcttg acggytggat
2340
cgctttgaggt cagaggtctca agacagcctgt ggtgtacatg ggtgaaacc ccatcctctact
2400
aaaatccaoa aatattagca gctgggtagg actcgttgtg acttctcota cacagaggcc
2460
tgagcaggg gattaagttgg aacccggygg ggggtgtagg aatggagctag agatggcaac
2520
acctgtactc acgtctgtagt gaggtgagaa ttcttctttaa aasaatagc gcttctttaag
2580
cagctacttt tgtatcacaag gacttaaaac ttattttatc atattttggt tggatatttttg
2640
cccccccttctt ccatccctat ataattctgg gacgacagaa acatgtttttg ctaaacocctt
2700
tgtatctcctt cccacacacc ccacagttcct gttgcaatag aacacttaata aasacaccgct
2760
gatgggctc
2820

<210> SEQ ID NO 364
<211> LENGTH: 2984
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 364

gagagggagc agaagaagaa aaaaagaaaa aaggggaaaa aaataaatac aagaagtcgg
60
agaagaagc aagagttggc ccctttggag aacccagcttgatttatgg acgtatatct
120
tctctccgtt tgtttcattta aatattatatt tgttttttgc aatcttttac ataagctcgtag
180
ccctgcgctc ccctgtgtgc gacgagcttg tctctttctgattcttctggtt cctgcgctctttg
240
gagccatggt ctctccgttgc gaocttgcaag ggtgagagcttg gagaagcttg aagaagctgg
300
gcttctccttc aacagctggag aacggagcggc aagagctgaaag aagagctgtttttgctcttc
360
cctgcggttt ctctgggagc gcttcgccc gacgcagcag gcggctcctcgc ccctttttcgc
420
caagcccccc gccctgttttt gtcctggaga cccagctgagc ggtgcgcgctg gtcctgcttcttctt
480
cagaggggc ggagcggcgg cggaggttag gtagcccttt gttgtaaag cgcacgcagc ctacccctg
540
aacgcttttag cagccagcttc ccacagcagta gtagcgcgatc ctaatttctc gtaagctgactgc
600
cacccgcaaac gtagccttgg gacgttgtcag ctggtctact ctaatttctc gtaagctgactgc
660
tcctcagagc agtgcccagc tgcagcctgtag gcgcacccag cagcctggagc acctggtcctcc
720
tcctcagagc cggagccccct gcctagcactg agagggagc acgtgcatcag agtggtcctcc
780
acagctgtcgg ggcggtcggc gcctggtacct tgcctggtacct tgcctggtacct tgcctggtacct
840
cctgacccccc ggccctgcct gcctggcctgc gtcctggtacct tgcctggtacct tgcctggtacct
900
gacgcgtcgg gcgcgtcggc gcctggtacct tgcctggtacct tgcctggtacct tgcctggtacct
960
ggttctcttc ggccctgcct gcctggtacct tgcctggtacct tgcctggtacct tgcctggtacct
1020
-continued

actctcagtg cagccaggtc ctatcgtggc ccttgaggag accttgagctc tgcagttg 1080
cctgatgct gcctacaaca gatttgctt gtaaaagacg gggcaacgtg acttctcct 1140
gtctcggtgc gcacacgcggc aggtgtgggt ctcctcagggc aaccttaacc tgggcccgt 1200
gagcccttcc ttcgctgccc gcctcagctg ctctgctgca cagacaactct ctctcagt 1260
gtctggccccc agccacaccc tccacatcct gatgcgggtg cagttctatg acagatctc 1320
cctctcggtgc cagcggcagc ccaggtggcg ctccaggaag aacaagaccc ttggctgtca 1380
gtacccaggg tgtgctgaca ctcttccctct gcccaagagg gggccagcgtg atgaccattg 1440
gctctagaca tcaacagctga aataccaaat atacccgagc tctgctcttg gccgtcttctg 1500
gacctaggg caatggggga aotacaggtg tcaagggctc cagagcctca aacctacact 1560
gtctgctacc ccctcgagcc cctctggact ctctgtgtctg gaaagctcttg gggggccagg 1620
cctccgaca agacccgcca octcacaagc tggcctag ctgacccgca tcaacccacc 1680
cgggtcggat cccagagctg gctctgggaa gcacottggtg gtttgacgct gcattccttg 1740
tgctggcat ctatcgtcagtt ctctctctct ctcctctcct gacactgagaacc 1800
tcaagggcata caagctgctc cggccagag aacagctgcat ttcacactct cttgacgggc 1860
tgctggggcc ggcgagcttc acagagcctg gcctggtgacct caaacgggctc gggagctctc 1920
cacgggaaga aacacgctat ctgcgtgctaa gcacacacag cctttggagtg gggggagatg 1980
gaacacgcag aggccacagg gagcggagctg acgatcgctgc ggggaaaaac 2040
tgacagcagtt cggacagagc tccctccgct ctctcctcctg gccgtctctctg aacccacagt 2100
caccaagggc agcagcagcag aggccagagt gcctggtgacct ctcgatctgga 2160
agcccccccag gtctgtgcotct acggcagctg gcacattcccgc ggacgagcagcacc 2220
tggcctctct ccctctcaggg ctgctgctgct gcaccaacctc aacccgacctc 2280
ggctacccag tgcctgctgg ggccggcgag acacccactcg cctgaggtct gcgaatgacg 2340
gagccggggc ccccccttgt acacccattg acaccccacc gccggtcctct acacccgagc 2400
actctgggaa cttctttgctt ctcttoaatc attcgatctata aataactaat gtctotaaac 2460	tttttacaa aacgacacaatta cttctctcata atcaactgatc tagctgagaa aacataacta 2520
gaaagctccct taaaatgcat cacaatgtaa atattaccata aagagctgtaa aacttggaatt 2580
actacaacgc acagagctac gataggggaa aaaaaaanag tggattagtt atgatcccttg 2640
cctttctcata agaasattag gcacgggcac ggtggtccac gctgtgaatt ccagctcttt 2700
gggcagcca gcggccgctc tcaaggtgtc gcagagctcg aacgaaacttg gccacaactct 2760
tgaaacctctg ttcctctcata aaataaanaa attagctgga tgtgggtgcca ggtctctgaa 2820
tgacagcagtt cggccagagc acggcagctg cagagccttc cagaaaggttt 2880
agtcgagccca gtcgacccca cttctctcagtt ccgctgacgc cagagcagctg cccatcctc 2940
attaaaaaaa aaaaaaaaaa aaaaagaagaaaaaaa aaaaaaaa aaaa 2984

<210> SEQ ID NO 365
<211> LENGTH: 3061
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 365

cggccagagc gagcattgggt gggagtagtt ctttggcagc ggccagctgg ggtaccggtg
```
tgctggagcca tgtgagagtgc cggaggttcg aagagatgta cctggaccttc ctggacagc 120
aggyagacacc gggatattat cagagcacaag tctgggagct gatcaagtac aaccaaaccc 180
ggctgattgt caagtgtgatg acagctgaga ggaaagaaca gaagagggtct aacgctgctc 240
tggaaactgc cttgcaagggct ggttggggtt tccagcgggc cttaaacgat gattgtggcc 300
coaatgatgc tacctgacct aacgagatag aagagtctct aagagagtctgg gaaycagct 360
ttggcgtccc gcaacgctgccc cccggctgctt tacctctcgct tgtccccgtc tgtggtgctt 420
gtggtgggg cctggctcc ctaaaccctc taggctgccc ccaagctgctc cgcagcggtccc 480
actacgctgc tgtcatcagc aagcctcagt agcgagcttgt aattttgctcc accacccctcg 540
tgccttccc ctccagcgtcc actatctcctt ccaagagatg gggagacctt ccccttagaga 600
cgaatagtgc ctttctgctt tacagagctccc acacacactt gcaacctcc gagatggccgg 660
agaagcggcccc aagccggcgtc ctcgctccgct tgtgctttctgct gattgtgtgct 720
tggatataac ggaacctggct gacgaggtggt aaggtgtcgga aacacccctct gttcctcttg 780
agaagagggc agctacctggg cttgggacgct tccagagctgt tggtaggctct cttataaggta 840
agcagatgtg cagagctgtg cagccctcct cttctctgtga ggataagcct gaggccaga 900
agttcgaggc aacccatatc cagatctctct cttaacgcgctg acggccccag cggaggcc 960
agtacatcgg gacgatgtgac gtcaagagag atcctctcct cccctggcttg ggcggaggggg 1020
aacggcaaccc aagcaccaccc aagagtactt cttaaacgag cggccacgagc tgggccccaa 1080
aacacccctgg tgcacatgtg cagcctcttg gcgtgtgctcta tgcagcgtgc gcagtcggta 1140
tccacaccc ggccgaggggc ttcgtgtggag gcggctcagt gcacagcgtgc ccccagccagc 1200
agcaagcccc aagggagcgtc cggaggtgct tgtccgactcg ccagctgtgct aagggaggtc 1260
agttgcattg atgatgcagct gcgcgccgct gacgtgctgc ggagacggtg aacagcatct 1320
tgcagaccaag ctggagagcc ctggagatct ccggctggagct gcggtggtgc aatgccccct 1380
gcggtgttttt gcggagctgc aacccatgct gacggaggt aagccagattt gagacacccaa 1440
tggagacacgct tgcgatcag cgtcaacgct tgtgctggag ttcctgtcctc atggagctggc 1500
tggtgctgtg ggcgtgctag cggagttgag agctctccatt cccgtctcctt cggatgccc 1560
gttcagcgtc acctttggtg cagagatggtag atgtatatcc tctttggtgt gcgtgtgatg 1620
ttcgctgccg agatattacc acccttaacc aggaattcaca cgcagccacc cagatattacc 1680
aggaagctgg cagcttccata ctggggcaca agaaaaaaa gggaggtgcgg ggagcttg 1740
caatctgccg cagctttcccactg ctggggccaa aacccatcaac gcggctccct gcagcgatgt 1800
cggccagtaa ctggcttgat gactcagcag cggagtttgc aagacggcgag 1860
acagccgag cgcaggttgc gtctggagcct gcagcggagc aagtcagctcag 1920
cagccccagc cagccgagag atggagcaca atcggaggtc cggacgagcctg 1980
tggaggtgct ttcgtgaagc attcctggag cggagagaga aacgagctggag 2040
aggaagctgg gggagctcgg cagacagcag aagagagag aagagacggc aagagagcggcag 2100
aggaagctgg gagcagagcag aagacgagcag agacaggtgc cagaaggtgg gatcagcag 2160
acacccctgg tccctgagcactagggag aagtccctca aagtcagacact cccctagcgc 2220
cgtcaggaag ggcctggcag gacctcagcag aagtcagctg cagccggagat 2280
catctctgtg gacgcttcct gcggagagct gcggagagctc 2340
```
-continued

attcgctcaca agatccctac aacccggca aacaagaagc acctctctcc gttgagatcc
agyctgtctc gagccagact cagagatgca aatacgctat ggttcttgac gcctctgtctc
attcaatgta aggaggctct gcttctcagac tggggtgttg gacagagagt tgggtctgttg
attcccccact cccaggcttt cgcctctcct ccttacagct gttgatctct
actgaaggygg aggatggttg tgtgtgaagc tgtgctcagag actcggtgaa cccttttggga
attggtcataa aagctcgccca tggggctagg aagaggaga aatctgggga ggaagaacctc
tattgtctct tcatgtcaca aagacttcgtc cattccccttc cctccgcaac ccacacaac
cccaatgtaa atacactatg aacctctgg tttctttggg gcaacctagt tttctctttg
tggttttttt ttgttggggt ccctctctca cgtcacttqcg gtttttggggt
gttcaccatt gttggagagaa gctgctgcag cgtcccaagca gttctgggttg
cagtagacgcc attttgggag aatagttttg gatacaccaca aaacaacccca tatacaggat
aatctcagg atgtgtaata aacttagcaca tttcgtaaca aaaaaggggg goccggtaaa
c

<210> SEQ ID NO 366
<211> LENGTH: 1360
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 366

cgggggtgct ccctctgtgt ccctgctcg ccattacctc ctacagctat gcggactgctg 60
cggggcctgc gcctcctggc gcctctgggc gcggctccgt gcgttctttgg gcgggggtgc
ccttctogc gcocactatt caaggggggt cgcgggctcg cgcgggtatcc gcgtctcctcg 120
ccgcttttgt gcctctctgc gcctcgtggc gcctgctgcgg ctgcgtctcctg
ccggctggcc ccctctgtgt gcggcgggct gcgggggtgc
ccggctggcc ccctctgtgt gcggcgggct gcgggggtgc 240
cgggtcggca cggtggctgt gcggggcagc aagagcaaac gcgggaacagc cgggggctgc
cccggctggc ccctctgtgt gcggcgggct gcgggggtgc
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 300
tggaaactgcc gcactctgac tgggtctgtg tcggcctcct gcggggctgc aggggtgaga 360
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 420
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 480
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 540
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 600
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 660
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 720
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 780
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 840
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 900
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 960
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 1020
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 1080
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 1140
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 1200
tggaaactgcc gcactctgac cagagacgg gcgggtcgcc ctcgctgacgc tacsacccctc 1260
gtcctttgga ggggtcttc tgggtgagg gatgggaaggg aagggacccc tacccccggc 1320
tcttttcg aaccttggca aaaaaattat gytccagggg 1360

<210> SEQ ID NO: 367
<211> LENGTH: 1412
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 367
cgggtgtgca gcaaaagct cagttctgct ctcttcctct ccocgagcaag cagctagcttc 60
acacacgtct caacctcttc caaacactac cagttccctgg gctctgtcaca ggccggcagc 120
taagggcgcgg gggcgtctcag cagcgccgcc agcgtttctag cagggctctgg gggttctctgt 180
ttcgctcttc ctcggccgct ccctggcggc gcaatgtgcttc cgggggctctg 240
gccacagga taccctgggg tttccgagca atggagggca tccaagagca gaagggagc 300
tagcnnacggct cggctgctct taccctgggaa gaggagggag ccttgagacc 360
gagacggac ggggctgagac caaaatctcg gggaccccttg acaagagcag gcocccctgc 420
agagacggca gcctttctct caaacatcct gcaagagcag gttcctccaat ccaccctcag 480
acgtgtgcaag atccaggcaag ctttcttctcc cttggtcccc atgcagatcc ccgcttttggc 540
cttcagcaac agatctgagag cagcttgccgc atgcgccctc ctggggggag ccaaaatctcg 600
tgaacccggag agccattcggct ccctgggctc gtaaacccacgc ctaaatattc ggcagatcag 660
ggggccttca aggagggagc gttttctctgt aagagagcag ccagagaggagc agtccaaggg 720
cctacacggcag ctaagccgagc ctccgggctc accgttgggc tagatgcccc caaatcttcag 780
gaacccggag agatacctggcc gctaaatgtc gcgaatgtcg ccgtgagaccg 840
cgagcatagc caggtgcagc ctcggccgtt cagatgctcg aggagccaccgc aatggcctagc 900
aacagttgct cggagggctg agtgctcttg acagcaatcag cagctgtagc aatgcccgag 960
cagtttctag aatcagcccttt ggcaagacagc aagagatcag aagagagagc 1020
tggagggag gcctggccgct ctaagcccttt gcagagggctg aatgttgtcttt cctccagggc 1080
cactgctgt cggagcccttg ccaaaactgg gcaagacggc aagagccggc ccaatgttctg 1140
gaggtcttgc ctaagcaatcag ggtgctctgg ggcagtctggc aagagcagagc agtggcttttcct 1200
ctggagaggc ggaagcccttg taaagcctgg gttctgtccct gaggagcccatc ctcctgcaat 1260
acagtccagc agacagcagc cagcagcagc gacttgtggtc aagagcagatc aatgtgctttgc 1320
gcagcccaaa ttctgagacc ttaagcagcg aagagagccc taccctttgg ggacagagac 1380
gccataaaaa gcctagcaagt ctatgtgtctgt tc 1412

<210> SEQ ID NO: 368
<211> LENGTH: 1675
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 368
cggcagcccaac atacccgtgag aagggcagcc ggacgagcgg aagggagcgg cagcttggccc 60
ggccggcccac ccgcccgccc cggcgcggcc gaaattgcgcc ccggccgccc ctcctgagcc 120
cgcggagactaagagggagcgcagggc agggcgttgag ggaggctgagcagccggc 180
cgcggcggag ggggctgagcgc tctgcccc gcgcggcgc cgcggcgc 240
geccgagoca gccgcccgag ccccaggggc ggccccaggc ccggcggcg cccgcggcg 300
cacagccgag gccgagcggc ccaggcggcc ccggcggcgc cggcgaggcc 360
tgttctggcc gccgagcgc acggcgagg ccggcgcgg acggcgcgg ccggcgcgg 420
attgacatg tgtgagctcg ggcacaggg ccggcgcgg ccggcggcgc cggcgcgcgc 480
accatttgg caccccttctg agatgagcc ccaacacat caagatta gatactttca 540
aagggctgga gaaagatac aagattatat cagcgcgcgc ccctcctcga gccggttggg 600
tgtctgctga cctgtgagaa aagaggagat atctcttctgc agggagggct ggggggag 660
gcgcagtcgc cttgacatcg tggactcctg ggcacaggg ccggcgcgc cggcgcgcgc 720
agaagagag cctgaacccc agctgacac gttggagctc cttgcagctg cccgggctgc 780
ccagccgta cacagctcct cttcctgag cagatgctct cttgtgagc cttgcgctac 840
agaagacat ccagcgggga cgggagagag ccctctctgc tctgtggcgc cggcgcgcgc 900
ttgtgctgt cgtcggcgc cgggagggcc cccggcacga gttccttcag cttgggagggc 960
cataagcgg ccctccaggg ccctgtgccc acggcgcgc ccggcgcgc ccggcgcgc 1020
tgtgctgag cctgacatcg otctcctgaa acagagctgg taatgacttc atccccc 1075

<210> SEQ ID NO: 369
<211> LENGTH: 1127
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 369
cacagctcgc gctgctgggt ggtctctagg gggactggc ggagttaggc gttcctgagc 60
gacagcagca gttctctgg ggtcaggag gctggaggct gttcctgagc 120
gcgccgcttg tggctctgg ggtcaggag gctggaggct gttcctgagc 180
gggcagctgt gggcagctgt gggcagctgt gggcagctgt gggcagctgt 240
catgtcgcgc gagtgagcgc gggcagctgt gggcagctgt gggcagctgt 300
ggggtggcgc cagccagcgc cagccagcgc cagccagcgc cagccagcgc 360
gggcgccgcc cccctctcgg cagccagcgc cagccagcgc cagccagcgc 420
tcttggcgc ggtgctgggt tggagcagc cagtccagc cagtccagc cagtccagc 480
gggaggggc gcacggaggg ggaggggggg gcacggaggg gcacggaggg 540
ggcgacccag cttgctgggc cagccagcga cttgctgggc cagccagcgc cttgctgggc 600
gttggcgc gcacggaggg gcacggaggg gcacggaggg gcacggaggg 660
gatcgccgcc ggtgctgggc ctcctccagc ctcctccagc ctcctccagc 720
cagggggcc ctcctccagc ctcctccagc ctcctccagc ctcctccagc 780
cacagcagca gttctctgg ggtcaggag gctggaggct gttcctgagc gggcagctgt 840
aaggggtggt ccttcctcgc ccccagagct ggcgagctgc ccccagagct ggcgagctgc 900
gttgagcgc gcacggaggg gcacggaggg gcacggaggg gcacggaggg gcacggaggg 960
cggggcggcc gcacggaggg gcacggaggg gcacggaggg gcacggaggg gcacggaggg 1020
gtcggagcgc cagcttccag ctcctccagc cttgctgggc gcacggaggg gcacggaggg 1080
gggggcggcc ctcctccagc cttgctgggc gcacggaggg gcacggaggg gcacggaggg 1127
cctaatgaga aaggaagaat ctcacttggaac cttttgttgt cctggtgtct tgtctctctt 60
ttagttccct ctttcctact tgaggagaga gctgtaggg aaggaagct cctcttggc 120
tgatctcctt tcttatattt gaaatatcg tacagataa gttaagggac atctgyaat 180
ccttccaaat ttcctccaaa gctctatgctgt ctgctttcag cctctttttt gcctatgac 240
ccatacgtgt gtttcatgga ttagggcggc tgaagggagc accgtgtgtat aatggagag 300
agttttcttg aagagcgaat tcccatcatt cttcacaagat tactaagga cttggaatac 360
tttcogaca tggagagaga tggagagaga ccctgggtttctt ccctcccacaacctgctcgg 420
atgttgagct gggagagagc agcaaggaag accgtgttcac cctgcggcctgt cctctgcctg 480
tgaggtgtg gagaagacac aagctttccc accttctgtgc cctctattct cctgcgtgtg 540
cctctggct tgcgctttt cccagacacc ctttcgttttta caggtacagcagagag 600
attctctcag ccgcttgagaa agatcctcaat aaaaacctcag gattctggaat cccctcggcc 660
tccagcctcc aacatctcctt ctctcctcaactt cccccaatg ccccaacagc 720
tggttaaaa gttctgctctt aacggaatgt acatggagac ggaagatggaa gcaccccaag 780
tcattcctgta gttcaggtga cctcgagacct tatacttggt tgttctgctat aaaaatggac 840
aggaagacag caaccccaaa atcaggtttc atatggcttaa atggtcggcg aactgtacgtg 900
atcatttcttg tgcgcttggct cgaagcacaac gacccctctc tcgagatggg cctctgcctc 960
tgctgagcag cccagaggtc acagtaaag tccagggaga gattgctcat tgaatggcag 1020
gcaccaggg cccctcgcct cagaggttaa gcacccctcc ttcacgtctg gctgtctgagc 1080
acagagatca gcagatcaat gcacgtttcc cccgaggtgt gcccctgtca gtcgacacgt 1140
atacttggt caagatcaact ccctccctcaag aggtgtttga tacaacactaatgtgtgagct 1200
cataacaaaaa ggaagcaacac ataatggccact tactgtcctt ggtgtcagat gacacaccag 1260
atctctctct tccataatct ttgccctctgc gcacccacatt agaagaagct gcaccaactt 1320
agagcttgca caaactctctt cttcttcag cogaagacat atcttcattgc gcagaagac 1380
tgcgctctg gcgagctattt ttttacattc cccacacattt ccagaccttt ttaatgctaat 1440
tcgtgagtga ttttaagagc cttcaatacact cttcaggttaa ccacaggtt gttctcttgcc 1500
caccctcata ccagctggcg ctctctccttgcttgccaat cttagtccag ctggtgtctgc 1560
acatggctctc acgtcgcact ctctctctact gcaagcactc cccactctta tggcttctctc 1620
acagctct ctcactcact ctcctctactt atccctactatta cttgctctct gatggctgtg 1680
gagaggtgtc cagttcaagt cacaatatct tctgagatta ttcatctctg atacacacttg 1740
tattacgcttg catgatgctga aatttttttaat tcagacacttgt gaattcacttgt 1800
taccacagc aaggagctgt aacagatgtat aatcttactgct attctctttt gtaggtgtct 1860
aaatgagcag ccagctgctctt gtaggtgtat 1920
<400> SEQUENCE: 371

agtcagacct gcgccccagc agtgctgggt gccggaggtc cggccgggacc 60
actgaggggac agtgagcagg gacgaagctg gaaagcggga gaggagaggag 120
agggctctct aggagccggc tctgggcac ccggcgagcg ctcagaggtg 180
cgggctgcc ctactgacgc atgaaaagaag tggcctggg aggagccggc 240
cgggcaatag gttggtgctc ggcotagaga ggcgagatgg gaaagcaggag 300
gaatgacact aggggaggag ccggccggtg gcccagatctcgctctgtg cggagatgg 360
tgagtttccg gcgcaggagg gcacgggcacg aacgccgccc gacgctgatg 420
gagccagaca agagccgggg cggcctggcg gccgcagactg cgcgcctct 540
cccgcgaaa cttgagcgcg gtcagacggt ggagcgccgg gccgagctgg 600
gtggtgactcc gttggagggc cccgtctgag aaggtgcggc cggcagagaa 660
tgtgcctcaag ggtattctcg agagagcagcc ctgcgcctcc gcggccagcaag 720
ctcaccgaggg gcggcgagc gttggagagc ctggagctgt ctggagcagagc 780
agctcccccgt ccggaggggg ctggagggag ggtgggctgg ctggagcgtt 840
tccgctctgg tgggccgagc gctctctgag tccggagcgg cggcctgggg 900
tgcacaggttt cgggcgaggg ctggagcagagc cggcctgggg cggcctgggg 960
gtaaggggaa cggcgtggctg ccggctggag ccccgagagc cggcctgggg 1020
agcgcgttgc gcggccgagag ggtcgagag ggcggccagag cggcctgggg 1080
ccttcctact ccgggagaga ggtctgcctg ccggctggag ccccgagagc 1140
acacccacac gccgttcctgc gcccggcagc acgccgagag cggcctgggg 1200
agctccggcc cggcgtgcgg ctggagctgt ccggctggag ccccgagagc 1260
goatcctgg cggcctgggg ccggctggag ccccgagagc cggcctgggg 1320
agaagggcgt cgcccctggc cggcctgggg ccggctggag ccccgagagc 1380
agcgtcgcgg ccggcctggg cggcctgggg ccggctggag ccccgagagc 1440
gggtgatcc ttaactcagct tcgaggagc aacgagcagc acgccgacag 1500
cgtaaattg cttggggttg ggcacaggcc gcggcagcag cagccgagag 1560
ccatcggtc ggtcttggag cgcagagcag cagccgagag 1620
tcggagcttc gcggagcag cgcagagcag cagccgagag 1680
acagcggcgg ccggagcag cgcagagcag cagccgagag 1740
gccggcgg ccggagcag cgcagagcag cagccgagag 1800
ccaggtggcc ccggagcag cgcagagcag cagccgagag 1860
ccaggtggcc ccggagcag cgcagagcag cagccgagag 1920
ccaggtggcc ccggagcag cgcagagcag cagccgagag 1980
acagcggcgg ccggagcag cgcagagcag cagccgagag 2040
acagcggcgg ccggagcag cgcagagcag cagccgagag 2100
acagcggcgg ccggagcag cgcagagcag cagccgagag 2160
acagcggcgg ccggagcag cgcagagcag cagccgagag 2220
-continued

cagagctgt ccccgagggc gccgctccc ctcactttac cagcggcagc ctcggccccgg 2280
tgcttcctt ttgcttgag atgcagatcc ttccctcagc tctgcctgag aaccccaaca 2340
tcaagacaa gccaccagtc tggggtgagg tcggcacaag aggccccgat ggagccagtt 2400
tgagcttcgc gttgacaggc cgttttcgcg aggagtcgtt atctccctctg gggaaggttc 2460
tgacacact ggcgcgcggtg tgggctata ggcctctgcat cttctccctct ctcgtcctg 2520
tctgagagaa gcagcttgag ctctgccccg cccgogact tgcggcagaa aaggaagccg 2580
tgagagcagc gttcttgagcg cccagggca aagggacgttt ttcagcagcc atcaggtgctt 2640
actagcaag cggctcggcc cctctggccc ccgggccccg cgggcttcgc gcctctcggcg 2700
gcgggctgcc ggcgggtggac acgggctgtg gcgtggctcccc ggagtctgacg 2760
goctcggga cccctccagct ctcgagcccc ccgggctgcgc agagggcttc cggcggttcgc 2820
agatgagggg cgcctggcgc cgtgggcccc ccgggggtcg cgggggcccc cgggcttcgc 2880
agagcgctgc caggtgtcct gaaacgcagc ccggagcctg gacgagcgcg cgggccccgg 2940
agagcagctg ccccttctctg ggcgcgagtt ccagcagggc aaggggctcgc gcctctcggcg 3000
cctcctgtgt ggcgcgagtt ccagcagggc aaggggctcgc gcctctcggcg 3060
tggccagctt cttctcctgt gcgcgcgagtt ccagcagggc aaggggctcgc gcctctcggcg 3120
cacccctagtt gcgggctggc gcgggctggc gcgggctggc gcgggctggc gcgggctggc 3180
ccagcagctgc aagctctcct acgcgcgagtt ccagcagggc aaggggctcgc gcctctcggcg 3240
ccgacacaga tggcgggctgc cgcggggcag cgggggtgcg cggggggcag cgggggtgcg 3300
tcccctggc gccggggcag cgggggtgcg cggggggcag cgggggtgcg cgggggtgcg 3360
cggggggcc cggacgagc acgcgcgagc ccggggggcc cggacgagc acgcgcgagc 3420
egggacgcgc cggacgagc acgcgcgagc ccggggggcc cggacgagc acgcgcgagc 3480
agagcgctgc caggtgtcct gaaacgcagc ccggagcctg gacgagcgcg cgggccccgg 3540
agagcagctg ccccttctctg ggcgcgagtt ccagcagggc aaggggctcgc gcctctcggcg 3600
cogatgtcct gggcgcgtcc gcggggctgc cggggggcag cgggggtgcg cggggggcag 3660
cacctgtaacct ggcgcgtcc gcggggctgc cggggggcag cgggggtgcg cggggggcag 3720
agaggtgtgg ttccttgctg aacaggtcag gcggggtgcg cggggggcag cgggggtgcg 3780
ccggggggcc cggacgagc acgcgcgagc ccggggggcc cggacgagc acgcgcgagc 3840
agactgacgt gcgcgcgagc acgcgcgagc ccggggggcc cggacgagc acgcgcgagc 3900
gcggggtcc cggggtggtc gcggggggcc cggacgagc acgcgcgagc ccggggggcc 3960
gggggttgg gcggggtggtc gcggggggcc cggacgagc acgcgcgagc ccggggggcc 4020
tctgtgcgtag ccgagctgtg gcggggtgcg cggggggcag cgggggtgcg cggggggcag 4080
cctctagggcc agcgcggcgt gcgcgcggcgt gcgcgcggcgt gcgcgcggcgt gcgcgcggcgt 4140
caccaaccgaga ggcgacccgct ctcactagctt ctcctagtgg ctttcggtgc ctcactagctt 4200
agcgagcgttt ctccttcagcg tggcagccgc tggcagccgc tggcagccgc tggcagccgc 4260
cacgctggag ccagcgcgagc acgcgcgagc ccggggggcc cggacgagc acgcgcgagc 4320
agcggcggcc ctccttcagcg tggcagccgc tggcagccgc tggcagccgc tggcagccgc 4380
cttccagacag gcggggtggtc gcggggggcc cggacgagc acgcgcgagc ccggggggcc 4440
ccccgccag cctccttcagcg tggcagccgc tggcagccgc tggcagccgc tggcagccgc 4500
-continued

tgccacaccc gctccaggg gcaagaggtc atgagggctt tctggtcaga goccocctgc  4560
gtgccacag aagcccccag cccaccacac ggtggagtct cactctgtag tccagctctg  4620
gttggcnaag caaccaccac acagaatagc caacaggcag ccaagcgcct tggccggcct  4680
cagggcctcg gccggccg ccctccgagg gcgtccgcgg tgggaccgga tctgggcttg  4740
ccttgagaa gocctgagct aocctggggt gctgggctgg ttcttggaag acagtttccc  4800
cagaaccttc ccctgctcct gtgcgtggat gtgtacacag gggcaccaca gcttgacccc  4860
tccacggga ggaggagacc ccctgctgca cgtgctcat ttaatcagg ggacaggtct  4920
cctcaaaa gctgcgtgca gtcgcc  4946

<210> SEQ ID NO: 372
<211> LENGTH: 1743
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 372

cagttgctct ctcgcaaaaa ctaacaaaaa tcttgttagc caataactca gcccotctca  60
tatttacgt caaggttttt atcctcaatt tcaacgtgtg gagagcgatt gcccccgggtc  120
cacacctag aagccggagaacctggtt gcccccgggc aatctgggga cagcgtcttg  180
atcacaactc ctcagctcag ggccagacca gcctcttcac caacagcggg cccccccc  240
ggagaagct gttgcggcga ggagttcag tctgagaggg gggagagggg gcccccattc  300
gggggtogact gctctttcct caaaaactgt gggaggagtc cctggggcga caaaaactgc  360
tctctctgca gcggtagagc agagagacag tcgcgcggcc cgcccaaacg gagctgcctt  420
cggagacatgg gcctacggcc gctgctgcag ctgctgctgc tggcctacagct gcctgcttc  480
cgccctttgg gcctctctgct ctcgctctgc aacagcaccag ggtgtgctcg tgggacagq  540
tggcccttc gcagacggct ctggacggac acagatcgtc gcttggtggga aagagagag  600
gagtgtggcct gttggagaga agtatctacca ccctcgagga acacacacac gagcctgac  660
tactggattc gttgaagat cacacgcttt acagagttgtc tgtgtggggt aagatgttgtc  720
aaccggcaca ccctgctggt ccgctgcttc cttccggcta gcctgatcct gcagttcttt  780
tctggtctg cactacgcat gatogtggag gggggcggg acacagacc gcggctgctg  840
aaccgctggag gcacgtccttc gcggttggtg acacctgcca cctaaaagag tgaggggagg  900
cgtccaaag atgggaccca ccctcggtgg tgtgtccatcc ttcogccctc cccggttcct  960
aatggtcttc acaacaccag ccttcctca actctgaat gctgcaacac ccctcactcg 1020
aaccggcc ccattctgga gttgaatt aacccgcaag cagccggaca ggtgtac ago 1080
tgacacggct gaccagcaca cttgtgctcc tctgtgaga cttctcagat gcagcctgca 1140
ggccccagca atcatgtctt ggtggcacc gcacatcag acaccaaaaa ccaagactat 1200
gatggaaag gtctgctgcac cgcctctagtg tgaacatcaag cccacctggg tgcagctctc 1260
agcctgacac acctttgtgt ctctctgctgt actaaaaatgt gctgtacacc cccaggcctg 1320
gctttcagtc acaccaggg gcgtcctctg cagctggtcct gctgctcatc cagctgcaac 1380
atcccatct gcagctttcg gcagacccct ccctctgctgct tctctgctgat ctaacactta 1440
aaccgccttc tctgtgcttc ctggatcgg ccgagcccttc tggccttccc cctgctgcctc 1500
gcctacag cttctgctgt gcacaaggg ccgctctctg cgcctcttgg gcctctgttt 1560
-continued

tccacgttat gaanaacgct atcctcaaca gttgtgtgaa gcagaagaga aagcttgag 1620
gggcgggtg gcagattgga gacgtcttgt tatatatatt atgtgtggyct gtgtgtgtt 1680
gttgagtata attatatttc atattatatc tttatatctt acataagaat ttttgtaccag 1740
tgg 1743

<210> SEQ ID NO 373
<211> LENGTH: 5061
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 373

atggtcgaag ttctttagga cagcggtatt aaagaagttg catcttctca tocgggaacca 60
acaaagacga aagagttggc gaaagaaaga gacattccag tgcagcagca ggttttagca 120
aacttgacaa aggtgaacaattagttgac aceaagacag gctttgatgt gtcgaagacg 180
acccagaaaa aacagacggt tttatctcag gacggattgt atctccaggt gttctcctgaa 240
tcgctttccc asaasagacg attagatatt ggtatctgaa aagcttccacc aagcttgaattt 300
gagaaactc tgtgtgattg caggttctag actaaaatca caactgtatc aaacagttact 360
cctatgtgag gcctctcttt ttccagcagc ctctatattta gcacactatct tcagagagaa 420
cagccgcccc atctggatctc cgtaatctctc acctctgtt catctctct caataaaccctc 480
atctcagata aacccagcgc attccaaattg gcattaactac caagaacctgc ccccttttaa 540
tccagagata cttatatattg tgcctttcag ggggttactc atatctttctt cagactcttg 600
aacccctcr caaacctctcct ccccaagagtg aacatctatata attaatcttg 720
acggaaagaa gtttaggtgga gataacaaga ataatctgtg ggttatcctcc 780
aaagctcagtt gcagagttgta gataacaaga ataatctgtg ggttatcctcc 840
tggaggttaa gggtgtattg agaccattag gacagaaaaa gttgcatcaag ttttgaagaa 900
aaggtctatt gggagctggt cttcctttcg gagaatctgc gacagagatc cagaaatttt 960
aagtaaaggta cttcttctgt gcacatttga cacaagctgc tctattttctt 1020
atttgtaacc ctctctctct ccccaagagtg aacatctatata attaatcttg 1080
aacccctctc gcaccgctgg aagattctct cttccagagc ttggaatcag gaaagagggg 1140
atggctcgttt ttcctcaggt cattacaaga ttgagcaaca aatcttcttt ctcaccctacc 1200
cgcaaaacc ggctctattc gtcagacgca aagagacact atgcaaggaag 1260
aagctcagtt gcagagttgta gataacaaga ataatctgtg ggttatcctcc 1320
tgccatgta tcctctctgt gcacatttga cacaagctgc tctattttctt 1380
gccctgctg aagtaaaggta cttcttctgt gcacatttga cacaagctgc tctattttctt 1440
tgcagcttt atatttctgt gcaggtcttc ggtataaacc atgcaaggaag 1500
acgagaaatc gacactcact cttgtccttc agtcagattg gtcgaacctc gcaccgcaaa 1560
gacaagagat acagacacgg ccctgttttg gagaactctct tagattctttt ctcaaccacc 1620
tgtgagctgc cccttctctt gaaacagag aacccctctc 1680
gtgagaagg cttctgctgg aacaataaacc atgcaaggaag 1740
gtgagaagg cttctgctgg aacaataaacc atgcaaggaag 1800
---continued---

aagagtaaga aagagcaggt taaaatccca aagagtaaaa ctgtctgtgt gacctctttg 1860
ttttgaggag aagacactag cagagttccaa aacctgggct caccctaccc tggaaacctt 1920
gttcaagtaa gataaaacca attaactgca gcaatattag actcatctca gctcagacga 1980
aatctgtyga gggtgtccac aacctgtgccc caaagggcca aaggtgctca ggaagccttg 2040
actcacaacg agcagotcga tttaactatg ttggtgtcgtc atggaatatt accaatttcc 2100
gtaccaactt atgaaaaata ctactggtata ttgtaactgt ctcacaatgg aacagatcct 2160
ttttacata ttcataccaa gagaatttgc atttgaattg caatctcatt 2220
tttttttttttt tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
caaacactag aagccagaac tacaattttt ccatacttgg ttaattgaaac aagtttggga 4140
gacctgca caaatttgg cccttcccatt ccaaccttgc ctcacgctgc cttttctcgtg 4200
ccttcttta atgcgaggacc catocttccaa tttttacctta aacaactcct cttttgaaca 4260
ggtgctgcta ccacccagggt ctctgttttct atcattcacta aagaaatca accatcataa 4320
cattaatgt aatgacactag aatattcgtgg gaagacagca tgaacactc atttgtcttc 4380
cgaacatgttg tcaatttccag gaaacctcag aataagctca gttatatttt tocaaccttg 4440
aagttaccccg gctttctcaac taggatgtggc ctgggaagca caacacattaa agatgtgaga 4500
gcacaacagga aatscagttg aaacagcttac cttaagacttt tgtgaattgc ttcacacgat 4560
gtacagattg gtagcaatggt ttgtctttct ttccaaacctc ttaaccagtt taagaaagtct 4620
gagggagatag ctaagcttctg agatgcaaggt ttctctcaagt ctctcctcagc ccacatagga 4680
g gagacacgtg aatatacactc ctctttcaagc aatgtactct tttttatcactg ggtgatcagt 4740
atcaaacatg ttgtaactga agatgaggtg gaacacattac catatgtccaa aacatacacta 4800
cctcagata acccaaccacg atcaccagct aacacacaaa ttttatcagaa aacggacaa 4860
cgacacttca atgaacagtct tgtacactgt ggtatacagc aagaaactca aagacacgga 4920
gacacctcag taaggaactaac ctgagcagaa ttaacttccgg aagaaaaataa cttggttggaa 4980
gtacaccttg ctgtggagaag cttaccttgg aagacctagc aagtttaatg gtatcagctg 5040
actgcagcaca ctcctctgtga a 5060

<210> SEQ ID NO 374
<211> LENGTH: 6802
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 374
cggcccaaca aacaccaagac gataggaggg gcgggcggag gaggggagag aaccggggaa 60
ccgggagctgt gttggtggtgt gcggggtggag atgtagaggac tgtcgagcog cggccggcg 120
ccggcgaagt taoaggagcct ctcgagcgcgt ttgaacacgg aacgcgggctgtg ctcagctggt 180
gagggcgcggct cctcccaggggc gcgcgctcgc gcgcacaccca tcogtcgcagc ccaagccccc 240
gggggggggc gcggggcggc gcggggggggt gcgggggagc gcggggggcgg gcggggggccc 300
gggggcggcgc gcgggggggc gatccggggt gcggggggag gcgggggggg gcgggggggc 360
gggggggccc gcgggggccc gcggggggcc gcgggggggg gcgggggggg gcggggggggc 420
ccggggggcc gcggggggcc gcggggggcc gcggggggcc gcggggggcc gcggggggccc 480
cctcggggcc gcggggggcc gcggggggcc gcgggggggg gcgggggggg gcggggggggg 540
cctcggggggcc gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcggggggggg 600
cggggggggcc gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcggggggggg 660
aagggagggc aagaggagagtg tgttctagac aagagagagtg tgttctagac aagagagagtg 720
ccggggggcc gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcggggggggg 840
gggggggggcc gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcggggggggg 900
tttttttttttt cttttttttttt cttttttttttt cttttttttttt cttttttttttt cttttttttttt 960
tttttttttttt cttttttttttt cttttttttttt cttttttttttt cttttttttttt cttttttttttt 1020
tgtatagctc agtttggata atggtcaaa caatttttta tcocattact aataatgtaa 1080
caatgccc atgaagaaa aatacaaaaa gtggaaatat gatattttoc cotttttatat 1140
tgcatctgtcgt gttaccagtg gaagcttacc tagagcattg atctttttcga cggatctgtctgt 1200
ttattttgaa cagatgttta aataggtca tttttttttc taaattttgt tcaatggat 1260
cataactacct gaaaatttctc agctagatgt ttaatctctc caaattgtacc actattttctc 1320
atggcattgct tttttttcttttctttttaa ctaaaaatgc agatattatattc ccatcccc 1380
catatatttc ataaattcagt gcagtttctct atgtaggttt ttataaaaaa gcttctggtta 1440
aatggccgta aagtttccag cttaaggtcta aatattttttttt caaacacatatatc catttttttc 1500
caagctatgag atcaagctctc talgtattgat gggacaggtta ttttctagctc tggccaggttcc 1560
attggcattgt atcaagctctt atcattttctt ttttttttctg gaaaattttttttt tggctggtg 1620
atgtggttttc gcagtttctct ggcgagctcg gatgctgcttt tttttttttt ctaaaaatgc 1680
ttttttttttcttaa gaaaatttttttta ctaaaaatgc gaggagccag gcagtttctttt atcaatgctt 1740
aaataaattt ccttctgtgt cctccatttta aaaaattttttta tttttttttttt cttaaaatgc 1800
aataattttt ctttattttttt atcaaggttta tttttttttttttt ttttttttttttt ttttttttttttt 1860
aataatcata tttcctttttttt aataatcatttggctgatc tttttttttttttt gaaacatcata 1920
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 1980
ttttttttctt ttttttttctt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2040
aataatcata ttttattttttt ttttttttcttt tttttttttttttt tttttttttttttt ttttttttttttt 2100
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2160
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2220
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2280
aataatcata ttttattttttt ttttttttcttt tttttttttttttt tttttttttttttt ttttttttttttt 2340
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2400
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2460
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2520
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2580
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2640
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2700
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2760
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2820
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2880
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 2940
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3000
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3060
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3120
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3180
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3240
aataatcata ttttattttttt tttttttttttttt tttttttttttttt tttttttttttttt ttttttttttttt 3300
gtgtgggccc ctgggtctct acocgtgcga gatctatactg acotgccgcg gggatgggga 840
ggacacagcc caagacggtgg agctcgggtga ggaccgcccc gcaggggaty gaaacctcaca 900
gaacgctgga cagctctgcttg tcgcttgaga agaggacgcg aagataacgt gcaagtgcga 960
gcactagggg ctgcgtccag cctcctgcgt gcagttggaag cagtctccccc tgcagcccat 1020
cocactctgg ggtatctggtg tgcggtgtgg tcgctttgca ggtgtactca cttgagctgc 1080
ggtcgtcgtct gtcgctggga gaaagaagag ctcgagagaa aagagagggg gcatctcota 1140
ggggcacagt agatatgaag aaggtgagct cctgagatcc tgggatctt gttggtgga 1200
gccactgggg aggctacccca cccacatact cttctctgcttg gcctctcctct ggtggtctct 1260
gacacaggtg cttgctttgg caatacctag cagctgacagt gcccaggggt ctgatagtgctc 1320
tctcagctgt tggataagtt acaaccocggg gggctgtgag tgggtggtgg ttgaggggga 1380
acacagggaca tgcctgtctct atgcggttcct ttcagctcata atgatctgag catgctgatg 1440
gcggtttaa atgtgcctcc tcaacgtaaac tgtgctcttgg aatatttttc taaatatcgtc 1500
stggtgtaga aacotcgcgcc cttggtggaac ttgatgtccgg gcctccgggt gacccctaaga 1560
acccgctacc cctccttctgc agagacagcc ccaccccgtg ggcocacctg aggccctctcc 1620
tcctgcagct cttcccctct ttcacccatca cttcctctgc tccacacaca gggctggatt 1680
gttccgcttc gtcgctcmaa ttgttgtgcc actgactat maacctacctc gtaatatan 1740
ttataactctg atgtaaattc taaattttta aatataatct cggagagaat gatgygttaa 1800
ttaagggaga agatcctctga aatgtagag aaaaaatataa 1840

<210> SEQ_ID NO 376
<211> LENGTH: 6754
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 376

gtgcgacgct ggccggcgcc gccggcggcg cttggcggag tttttccgat ttaaagctga 60
gtgcgacgga aaatgagcc ggaggacaca aacatctgg gcctgtgag acactagagac 120
cacaaatctt ccatcctctt gcctgtgag gcctctctct gcctgatctt gcctgatctt 180

gagatgctg gattgagaga aaaatcggat ccaoctgcccc gcctctctct gcctgatctt 240
tgctctctgc atatctccag acacctcttc acatgcctct gcggaaacac gcagagggga 300
caggaatgta aatataatctg agatgtggct gctggctat ctcgaactct ctgtgtctt 360
gcctagcta atgctcaac cc ataataatct gcgtgttttta agaacaagct cagccgttt 420
cgctctctct ccacccgctg agtatttttta agaggtatct ctggagcgtct tttatcttct 480

gctcgcgtgt gtttctcttt tgtgtgtat cggagcggga acatgtttat ttatgttcgaa 540
gacatacgc acatctgcga atataagcca ggaagctgag ttaaaccaag tgttttctataat 600

gcttcatac gcacacacg aagagctttt ttaagagatt acccactactc cagcgttata 660

ggagttctct gcacacagg cccaaaagga ccccaagcggc atacattcaaa ttcgctgcaga 720

ttctgaagaa cccacatgta tacattgaca gacacagacgc cctgctccga aatgagctcc 780

gatctacata cagttccttc cttggcctg gctctctcgc acctgacgcg cagagggagt 840

ggaagatctct cttccctgctct ccctcctctc gcctctctct gcctgatctt gcctgatctt 900

gagagggaga atcctctgat gatctgtgag gcacgctgcca ttacacagg aaaaaaataa 960
tttccatcaaa aacctgagag cttttatacc agacatgatc tttccaggaat ggttgcatt 1020
atagatacata attcaatgag atctctcattg aggcttggct tggagagat aacaaggaag 1080
tgatattcaga gatttttaat tctaatagtct ggccagctat gttccacaga agctgcatat 1140
cacgaagtag ctcatgtgagatatctc ccaacagcgtt atcttattgag ccatgagcas 1200
aaccagatgatt atgagatcctcg tagggctgat ggaactcatag tgaotcgaca gacaaaggca 1260
aaacgcttcc gaaatcgtgt aacaatgatc agatagcgct tggtcctcacc caaactcttt 1320
cagagagac agaagtggata tgcacccaac ccaaacacttg tgggaagaggg gattagacca 1380
cctatggctg gatccacagc ttcgtagggc ggcaatgatg tgcctggaaca ccaagctcta 1440
cagagcgcga gcagacggg acataggttgg gcaacacca tgaacacgg gcaagatgag 1500
ggagttcgat tgggagggct ccatgacatg tgcactatga gactctgctg gaggctgct 1560
tgacatctt catccacaga caaagcactat agctctacga ttcagactcgg ccaactcttg 1620
agtcggtgct tcgcaccaaca ccagcagact actacatttt ctoctgctgta tctgggaggt 1680
ccagagtagc tgcacatcctg gttttcactct gttgccagttg tgcacacattc ctaggcctct 1740
tgcgcaata ctgggaccaac cagctttgtcc aagcaotcctc tcaagccctt gcaagccacc 1800
agttggtggt tgggccacccctt ttttttctac actcgtctct ccaacggccccc ccaaatggat 1860
aacctcaca atagatgact taccaaccaac agtaagagtta gcaatggagc tttcaagact 1920
tctcgtgctc ttatttcagc ccaacagcagc gttggtagtta caagtctgcag gtcgacattg 1980
agagatcaac tcaagtccaga aagaaagtgag gagaagcttg tggaggggag agagaatcaaa 2040
aggggctttc tggggaccaac agtcgacctaa ccacattgctgc aagtttcacctg ctggttctct 2100
gagagcgagg tgcacatcctc ttcggacacac cccccctctctact actcgagttg ccgacatctct 2160
tctgtatgtgc tccgagaccc tccccctgtct cttcctctct cacttgaggg gtagatctctct 2220
aacatcaaa tggcaggttg acaagtttaaca gagaagccac gcgttttggca caagttgctg 2280
cagatgacca atccagacag tggagtaggac aaagatttctg cagaagccac gcgtggaagac 2340
acagagatga taacctcttg tgaggagcag aagttgtgca acaagagccaga gtaagactcg 2400
aggagagg agaagatgcct actctcttgg acaagcctgca acaagttttttg acggctgctg 2460
gcactcctga aagaaatccag goccaagagtg ggaaggtgg aataaaaaat gacgctgctg 2520
acacgccatt ccactcctctg ctaagcagaa gagaagcccc ccasaacttgaa gacagacaga 2580
agtgaagagg cagctgggag cttggtaaat ctagtcgata ttcttggag ttagaactag 2640
tctgactcct acaaaatcct cctacccgca gtaattcctt actcgagggc tggacacacag 2700
gttcttcag gcacacttcg tctggatttc aaggtcctg caagttgcct gctctctcttg 2760
cctccatata agccggaggtg tggcttgatga acgctgattt tgcgggtgct aagtcctcctc 2820
gtccaaaaaa taacagctcctt cccatgcgc acaagacaccc cctagtggg tgggaagac 2880
agaagatgggt atagcagagc cccagttggc tcaagacgagc gagactgggg cttaccacaa 2940
tcaggccggc gcagactgatc accagatcct tcagacccagc tgggaagacc agggagagat 3000	ttagatcct tttctcctctg aacgctgactg gcgttctgtgt tcocctctgg ccctttcagtc 3060	tcctacacgt ctaagcttggc gacgacagta ttcggacaca gcagcagagat ggcctcaagt 3120
agggctgttg aacacattcag ggaagtgggg gcctacacctc tgcggccagc agagcctctt 3180
aaccagactgg gttccacggcc gcgtgctagtt tggtcctggag caaagcattg ctcagaccact 3240
-continued

caatagggc ctctctcttag gatcctccttg gtagctttg ggcttgacc cctccacccagt 3300
gacgacagag tcgacgaag agcaattattg ggacagctgc acocctctctc cagcaacaca 3360
gacctgacagc gcctggaaga aatgcgaact ctttgggca gctctggaact ctgtcaacagt 3420
ggtcgcgttc tgaagccag cagcagcttg cagcagctgg cagcagctgg cagcagctgg 3480
gatcagagga caggtataa tgcacagaca cacccagaccag cagcaggtcc aatgcgaagga 3540
ggtttcaato tcctagggca atcaccacat tttaacctctca ttgagatctca gtagcaccag 3600
cagcactt ctcctctcca aggcatacca ccacccagcc caacatcag gacccgaca 3660
aacacccccca aagcacattc aatgcaagtt ccagcaggcc tcgagccgta gcagtttggtg 3720
aatcagacgc gacacgacact tgaacctgaa aatgcaaacact ctcctcgttg tgcgaagcgtg 3780
gctgctggag ctatgtagcct gccctccagcc ggttcttctta aatgcaccat ggtgcaacc 3840
cgccagcag cagcagctacag ctgcacgacc aatgccacagt gccctccagcc 3900
cagcagcagc gcagcagccg gcagcagccg gcagcagccg gcagcagccg 3960
cagcagcagc gcagcagccg gcagcagccg gcagcagccg gcagcagccg 4020
atggtggcc tttttgaggg aacccaacatg ccaccaaggc ttcgccaca gtttcctag 4080
caccaaaact ttcaggtgg gacaccacacc gcacccagcc gtcggcggct tgcagcttctc 4140
ccccctgcc tttgctgtgc aagaaaaaga ccccttccct ccctccctggt ccacccagcc 4200
cgttgccgt ccacccggtc ttcgccaca gttcgaagct gcagccgacgg aatgctgctgac 4260
agggacctgg tgtcttctcc gccctccagcc ggttgagagg aagtaagggg ccccttgtggc 4320
atggctgcc cagcctgcgc cagcctgcgc cagcctgcgc cagcctgcgc 4380
atgtcctgcct cccctgagag ccttgcagtct cccctgagag ccttgcagtct cccctgagag 4440
tctcagggg ctcctgaggga gcacccgagg cttccctctt aatccctggttt tttcttcttt 4500
tcagacggcc acacccggtact aactccaggg gcacccggtact aactccaggg gcacccggtact 4560
gattgtcgtg gacccggtact gatgtaattag cccgagggcc aatctctctg gtcttcttctt 4620
ccttcttcgt cctggtaccc ttgctgtcct cttccacaa aacagaaact ttcttggcct cttccaccc 4680
ctgtgatata ttatgtggta ttcgtgacttc ttcgtgacttc ttcgtgacttc ttcgtgacttc 4740
ttttgtcgtg ttcgctttcc cttatcatgagct cttccagcgtc ttcgctttcc cttatcatgagct 4800
aatcagacgc gacactccat tagtagcttc cagcagctgc aagacgtctct cagcagctgc 4860
ccacccctgc agcctcgtga gttcccttctt cagcagctgc aagacgtctct cagcagctgc 4920
aatcagacgc gacactccat tagtagcttc cagcagctgc aagacgtctct cagcagctgc 4980
tcctagtac gtttgagag atctctctctg cttccagcgtc ttcgctttcc cttatcatgagct 5040
ccctgcgtgc actccggtgtc gcacccggtact aactccaggg gcacccggtact aactccaggg 5100
tcgacgacagc gcctggaaga aatgcgaact ctttgggca gctctggaact ctgtcaacagt 5160
aatcagacgc gacactccat tagtagcttc cagcagctgc aagacgtctct cagcagctgc 5220
ctctgttcct gcacccggtact aactccaggg gcacccggtact aactccaggg gcacccggtact 5280
tctattctatg aactccaggg gcacccggtact aactccaggg gcacccggtact aactccaggg 5340
tcgcccggtgc gcctggaaga aatgcgaact ctttgggca gctctggaact ctgtcaacagt 5400
ctctacttacta attacttacttacta attacttacttacta attacttacttacta 5460
atttacttacta attacttacttacta attacttacttacta attacttacttacta 5520
<210> SEQ ID NO: 377
<211> LENGTH: 757
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 377

```
tgagctacc atgggaagag acatcccttg acttttgtgg cctggggag gggtagtgca 5580
cocacgcttt tcttccoca ccccacgcc ttgatgctct cgtctctttc aatotcttta 5640
tctaaattct ttataaagag atattattgt tagaagttagg cattttaatt tttaaaaaat 5700
tctctccac gactacgcca ctttatgtct tggggggga aagtaagata tgggggaaa 5760
aacataaaa aaactccgga aattaattaaa aacgagcact tggasaagaa actttttgtgat 5820
tttaagcgct cgcaaatatat gcaaatctcat ggctctgtgtg ttgtgtgtgtta tyggtgtgtg 5880
tgttgtgtgt tgtttttatttt tttttttt tttttttt tttttttt tttttttt tttttttt 5940
gggtcatgag aacgtgaccc catattgagatt tctgagatg comeccguttg gttttctcg 6000
gtgtgtcgcc atggctccccc tccacgtatctc atcagagcag tctcatcctt tagctgtgga 6060
tttagaattg cctcgttaatt ttttaagtta ccatgggtcc aaaaaacgcc gttttttcgt 6120
gatcggagcc cccgctctac cggacagcct tttttttttttt tttttttttttt tttttttttt 6180
accccttttt tgtggcctct ttttgagcct tttttttttttt tttttttttttt tttttttttt 6240
tcacttttct ttacatttttgtt tttttttttttt tttttttttttt tttttttttttt tttttttttt 6300
gtttctgatg ccctggttacctt cccagcgaccc tttttttttttt tttttttttttt tttttttttt 6360
```

```
taactggctc cccgctctac tttttttttttt tttttttttttt tttttttttttt tttttttttttt 6420
cctagagttag aaccccgctt ctctttttttttt tttttttttttt tttttttttttt tttttttttttt 6480
tccgacgaccc cccgctctac tttttttttttt tttttttttttt tttttttttttt tttttttttttt 6540
cccttttttt cttccaccct cttttttttttt tttttttttttt tttttttttttt tttttttttttt 6600
catcggagcc cccgctctac tttttttttttt tttttttttttt tttttttttttt tttttttttttt 6660
ccoagccgg ctgctcttttt tttttttttttt tttttttttttt tttttttttttt tttttttttttt 6720
gttggagga atgaagatttt gtctttttttttt tttttttttttt tttttttttttt tttttttttttt 6780
```
<210> SEQ ID NO: 378
<211> LENGTH: 476
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 378

```
taaggccaa gacggttttt cctttaatgta cacatttga gacgtaaaaa ccagctoaca   60
taaatcaac gacgcaagtgg taaataaat ctttaacctat attgctgctt tggctccttg   120
agaacagtga ttacattcat ttagttttct aacggtattt ttcaagagca gocatagga   180
atataataa ccttttccac ccaagaacca tctgtosea ataatactga aagttcaca   240
ttaggaacac gtoagacact aacgaaggtgc agagctgctg ccacagcaca gtaatcact   300
agasaagggc acgcgggttc ccagqgtgggg gtcgtcagc acgcagctcc ctccctcggc   360
cgccccccc acgggtaccc tctggtcctc ttcgccccaa gagggtcagc toaatagcag   420
atttccttc cacgctgcct ttagtgctc cacaaacgag tgcagcgcaca ggtcag   476
```

<210> SEQ ID NO: 379
<211> LENGTH: 2518
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 379

```
ggttgctgcct gcggctggtgc cacocttcttg gtcggggggg gctctgcgct cgcggcacc   60
aatcctccaa cagccttccgc gcgggctgggc gcgggctggc gcgggctggc gcgggctggc   120
gggacgccct gcgcgggtgc gcgggctggc gcgggctggc gcgggctggc gcgggctggc   180
ggcggcgctt cgctgggtgc acgggttgcc acgggttgcc acgggttgcc acgggttgcc   240
tgacacgcc ggcgctgcgg ctgcaagagcc cccgcgggctc ccgcgggctc ctgggctgc   300
gccaggttc gcggagttcgt gcgggctgct gcgggctgcct gcgggctgcct gcgggctgcct   360
ccgggctgcc ctgggttgcc ccgctgcggc gcgggctgcct gcgggctgcct gcgggctgcct   420
ggttgctgcct gcggctggtgc cacocttcttg gtcggggggg gctctgcgct cgcggcacc   480
aacctcctc cagctcttgct cagctcttgct cagctcttgct cagctcttgct cagctcttgct   540
cgctgcgctgc cgcgggctgca ctggggccac acggggtggg cggggggggg cggggggggg   600
cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg   660
aatccgaggct ctcgggggtcc cGGGTTGTTT tttttaaagg ggacatagcg ggtggagag   720
cgctgcgctgc cgcgggctgca ctggggccac acggggtggg cggggggggg cggggggggg   780
cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg   840
ggggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg   900
atccgaggct ctcgggggtcc cGGGTTGTTT tttttaaagg ggacatagcg ggtggagag   960
gggttgata ttttttaagg atctgcgccc acggggtggg cggggggggg cggggggggg 1020
tcgggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg 1080
tcgggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg 1140
ttttttaagg cttgatggg cttgatggg cttgatggg cttgatggg cttgatggg 1200
gcgggggggg cgggggggg cgggggggg cgggggggg cgggggggg cgggggggg 1260
aatccgaggct ctcgggggtcc cGGGTTGTTT tttttaaagg ggacatagcg ggtggagag 1320
```
ggtacctg ggasgctctgc gcagggccaa cagctctgcc ctcctcttgg ggctcattgg ggtgtttg 1380
attgggaatgc agggcccaga gctgtcttga ttcagcactg gctgctctct ggcatacttc 1440
gagaggatct aaagagaact tggagatagcgc acaaggcagag cctgctcagtg gggctacctg 1500
agaattgctt ccagctcgta gatagcgacag tggcgggacg ggcgggtggc acacgaagg 1560
agagggggttg cttgctgtat acctctcatt ccaagggcctgtggcggcttg 1620
gagggaaat cattcactgg attggctgag atgggacctat cactgcaatg ttcggggtg 1680
ttacgccg gactgagcgc cagcgtctcc acgctgttgg gctctgttctgg cacaaggctg 1740
taggcaaggg actcgcaccct gtctgtggag ccgagtgcgtct tcgggactgtg ggcgttgccc 1800
agagggatgc atagcactgtc cttgctgtat acctctcatt ccaagggcctgtggcggcttg 1860
agagggatgc atagcactgtc cttgctgtat acctctcatt ccaagggcctgtggcggcttg 1920
tggtgggtgg cacagggact ggcagactgc gacggttgctt gccgggttgg cagaaggtgc 1980
tacggcctct tctgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2040
actcgcggcc ggtgtctggg ctgctgtcctgg cagggctgtct ggggtgtgct ggtgccgtct 2100
tgctagagcg ctggtctgtg gacaggtgct cttgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2160
ggtgggtgg cacagggact ggcagactgc gacggttgctt gccgggttgg cagaaggtgc 2220
gctggctgct ccggccagcc ccaggggccg cccggggtgc gggctcgccgg cggggtgctt gcaggtgct cttgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2280
cctgctgcttc ccggccagcc ccaggggccg cccggggtgc gggctcgccgg cggggtgctt gcaggtgct cttgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2340
tctactacta ctctctcactg gacagggact ggcagactgc gacggttgctt gccgggttgg cagaaggtgc 2400
ccggccagcc ccaggggccg cccggggtgc gggctcgccgg cggggtgctt gcaggtgct cttgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2460
ccggccagcc ccaggggccg cccggggtgc gggctcgccgg cggggtgctt gcaggtgct cttgctgtcc ctggctgcttc ttcaggtcct gttgaggggct ggcgtggcgt 2520

<210> SEQ ID NO: 380
<211> LENGTH: 4160
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 380

ggcttgctgg ggggttgtgc tgaagacact cttattttatt gtccacccact gtggtggga 60

ttggctgttg cacattggat gccctgtctgg ccggccacgct tgaagacact cccgggtccc 120
tgaatagacc gcgcgccgct ccagccccttg ccgtcactctt ttaggggaa 180
gacatcagacc ccctcgcccc ccaacccccg gcagctgctctt ttaggggaa 240
tgatggagc gcggttcctgg ggcgggtggg gggtttgtgg ggcgggtggg 300
cgggtgcctg ccttacctgg gcagctgctctt ttaggggaa 360
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 420
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 480
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 540
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 600
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 660
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 720
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 780
cgggggtttg gcgggttctgc gatgtctccct ccaaggtcttt gggtttgttc 840
aagccagtc gggaagcccc agctgtggag gagaacgtta cctccagccc agggactcct 900
gctctcct cgtgctcttc aggccatac agaggacctg cagttgacg cgytagtttg 960
atgtgtggtc tgttggggtg caagcttta caagctgag aagcttctcc ttactgtgaa 1020
gctctgctg cgtcaggttg tggggccttc gactgtgctt acagaaagtc acaagcactt 1080
ggggtcagg ccactgtctc caactgata gttgatgctc tggagcagag caaggagccc 1140
gaccaagggaa tgggaagcgg gcagcgagca gacgaacacg tgtcataaag ttgtgcctcc 1200
ggggttcag acgtctcttg gcaacgggaa ggcgtgagaa ggttccagag gaggaggtg 1260
cgtggcttcag ccaatgaccgg tcgtgctcgc gactgtggct gacaggttcct cgtgactttt 1320
gcagcctgtg tgcctgtttgg ctcctgggag cctgctcaatg ggaagttggg cctcaatggc 1380
aatctgattaa atggctcctag gactgctgca gcggggccga gggacacttt gtaacagattg 1440
cgctgattaa ggtgctcacaac aacogggagaa atgggtatcttg tccacatcctt gctgagctgg 1500
tgggagagct cggggagagag acgtcagcag cagaaattgg aggacacttt gttgagtctt 1560
gggacagcag tgcagctttg gacgtctgga aacgacggcc gtgcctcacc gttctctatc 1620
tcagagctgc acacccctcc cgggctatc ttttctgtga aaaaagccaa atgggacactt cttgagtctca 1680
gtacagagct cgaagctactgtt gacggctact ggaagactct tcccctctcca 1740
acatacagcc ggtggctggaa ctcctgttcc atttttcctt gcacgctgcc tatttctata 1800
agatcaagtg gatgatgctt accctgagaga atcctgagag ctcgctgtct ctgcggcttc 1860
tggagagttt gctggggagt tcacctggtt caacagactt titctctcta atgtaaatgc 1920
itatatatc tatatatatg acagtaagtt cccacactcc atgtaatatg ttcagcattc 1980
cgcgtacta atcagatatga accctttttg gtaggggggt gtagggcact gtaaccgtgc 2040
tccagagtgc ggtgctcaagt gcagatctcttg gcgctctaatg agcagctgctt tccagctgcc 2100
aacaggtctct cccactctgg cccactcaatg agctgtggcg acaggtgctg acacccgacc 2160
cgggtactaa gtttgatttt tggtagtatata gtaaggtttg cctagttgct tggagcctgtc 2220
tcagatgctgc atggctgcag cggctcactt cccagccccg gactgatgc 2280
gctgactgccc cccctccctt ccttcttcatc cactatcactt cttctcaagttt 2340
tatagcatg aacaaacctt tagtaatttt ctctctttaca ttctctcttc taataatatag 2400
atagcactctt aaggaactcg cactacactg ggaagagagc ggaagcccc 2460
cactgttggc gtttcagccc tccctcagat gtaacgsaggc aggaaacctt 2520
gacagggcct gtttctccacca tgtataggtt ctagatatta ctttttagta 2580
gtgcttcaca tatgagtagg gaccaaaaag aagctgagac acaggtgactg cttgagttgg 2640
cgtgagtc cctgtgcagc tctatggttg gcaagccttc ccagctctgtt acaggtctgg 2700
acagagggcg atgggttaga taacacctaa caagtttgct ttagctacttc ctaaagctgaa 2760
cctgtgactcct gttgtctttcc ctagatggag ctgctgagc cgtttcaagg 2820
gctaacaggt tgtctctcag ggagctctgg cggcctgtta acaggtgactg ccagctctgtt 2880
atgtgtaagtgc cggcttacat ctccttttgg acagaaactc acgatgctac cggagctgctc 2940
gggacagctg ttggggaatgc gcgtctcttg acaggtgactg ccagctctgtt 3000
aatagcacttt cccactctgg cccactcaatg acaggtgactg ccagctctgtt 3060
cctttggtgt ctctctctat cttctttta acagctcagc aaccatctca ggtctttctc 3120
agcaaaacctc tctccatactg atttcgatca tgggaaggatc atattcgac gtagtcacct 3180
caggaagttt tggcttcttccc tgcctaaagg cttgtgctgt cttgtgctgg cactgtggttg
3240
gttggaaccg aatctgaaat gctctgcaat cacactactc acactgcttg tctgtgtgagt
3300
ttttagagag gggggttccca gttggctgttc cttcacattc acactgctgcc aatctgactc
3360
ggcttcacc gcggggctctgc gtcacattt gtagaagcag ggggtgctgt gcctctgtgtt
3420
catatgcttt gctctgcccc ctctggtcagg ctgctgagcg cgctggtcagt gctctgagct
3480
tccgctcctc gctctgctgt cccttcacaca gcggcttcag tgcagtagt gcctgagctc
3540
tcaagcctct caggtcttac catctctcttc tttttatttctt cgcagacttc cgcagacttc
3600
actctctctgc cagagtgcag gctgggctctc gccggtgccag ctttgagatc gctgggctctc
3660
cgctgcttcc gcaggctgat gctgctgcag cctctgagat cttgggctctc agggccaccg
3720
cacactctcc gcaggccagt tgcagagtct gcagctgcag cgttgctggt gcctgagctc
3780
gtgggtcggc gcctgctgctg ccctgctgctg gcctgctgctg gcctgctgctg gcctgagctc
3840
cggctggag cccagctggc gcggacctgt ttcctctttat gcagctgcct gcagctgcct
3900
aatcatttgc ggggtgtttg gtatagatgg attggtcttt cgcagacttc cgcagacttc
3960
actgagtgc accgcgagag ctaatgttag tgcaggtcag gcctgctgctg gcctgagctc
4020
cattgcttcc gttctgagct gttctgagct gttctgagct gttctgagct gttctgagct
4080
tgcagatc ccaatgttag attggtcttt cgcagacttc cgcagacttc cgcagacttc
4140
taaaaaata aaaaaaata aaaaaaata aaaaaaata aaaaaaata aaaaaaata aaaaaaata
4160

<210> SEQ ID NO 381
<211> LENGTH: 1295
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 381

gtgcgtagtt tggctgttcc ggggttagaa ggtgagcctg cgagcgcggg ggaagagttc 60
cggcttttaa ttggctgttcc tttgttgcag ctctgagctg ggggtgtagc 120
cggcttggg gcggctgttcc ggtgagcctg ggagatctga cggctgttcc 180
ggagacagttt gctggttccag ccaagctgacg ctggtgctgt gcgtggtgctg 240
ggagctgacg ataatgttag gcgctgctgct ggtcagagct ggtcagagct 300
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 360
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 420
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 480
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 540
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 600
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 660
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 720
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 780
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 840
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 900
ggagctgacg gcggctgttcc tggagatcag ctcgctgctgct ggtcagagct 960
ctggtgcta caaactgctt aclotcttcata tgggaacccggaa agccaaaggctt ggcacacttg
1020
tacaacgac ccacgcaaa ggaatgcaac gtaagttttttc aacttacccgct tcaaggtttt
1080
agagatagtt cgtagttcct cgctggtttgct tgggaacccag aagggctttt ggcacacccgg
1140
aaacggtcgc ggacagccggaa ggaaaactattt cacacccggaa caaacttccagaa tcatgtc
1200
attatatgt gttggaataa cagatgtaac aagggacctt gttacottaag atatataata
1260
aaacactat ggcagataag aaaaaaa aaaaa
1295
<210> SEQ ID NO 382
<211> LENGTH: 2210
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 382

cgggccccttc tctcctcctctg gacotggtcg gtcgggcctg tgcggggtc cggagtgccc ctttaaagg
60
cagctatttc tgcgggagag gggcggggtt ggggcggagc ggggctagg gcccgggccc
120
ctcctccctc ctccctctgt cccggtctcg ctgcgctgct ttggtgctcg cggccggtgc
180
cgtcggggcg gacotgctcg cgctcagttt cgccgctgtc tcggggcgtg gcgggtgagg
240
gggcgggtgt ctcgggctcg ggcgggtcgct gcggggcgttc aggggctccg gcgggtgtgc
300
cggcgggtgc gggcgggttc cggcgggttc cggcgggttc cggcgggttc cggcgggttc
360
gagtaagggct gccggggcgc aagagactgc ggcgggtgagg cctgggggc cggcgggttc
420
caggtagaag tgtccaggaag gctgggggaag gcaagatgaga ccaggagattc gcggtttgg
480
catctgcttc agatctttcc caagccgctgc gacgagggcc cccggtgcct caaggtgcct
540
cggatcatt ctcggccttt tgggacgccgc cggagttgac tgggacgtttt gggctgcggt
600
cgctcgggag cattgcttgc gctatgctgg gggccgcttc gcgggtgagtt cggcgggtgc
660
aacacagctc cggctgcttt gattacac cggaggtgtg tggacagcgg gctgtgagcat
720
attgcaggt tctcgccggt cttccccggag cgggttgccag ccggggtttt gccggggtgc
780
aggtgggt gcgtggtgca tggacagcgc gacaagagatc cttccggcag ccggggtttt
840
aggtgggt gcgtggtgca tggacagcgc gacaagagatc cttccggcag ccggggtttt
900
gagctagct aaggggcttc gcgggttagt cttccggctgg ccgggttagt ccgggttagt
960
tttccgcttc aacgcttcag ggcacacgc cgggggtttt gccggggtttt gccggggtttt
1020
agcagcttc accggggcct cgggtggagat cttccggcag ccgggttagt cttccggcag ccgggttagt
1080
aacacccctc cgtctgcttc tgggtggcct tgggtggac cagagggcag ccgggttagt
1140
agaagagc cgtgctgggt cttccggcag ccgggttagt cttccggcag ccgggttagt
1200
ggctctgcgg ccacggccgg ccagagtttt cgcagctggt cgggggtttt gccggggtttt
1260
ggctctgcgg ccacggccgg ccagagtttt cgcagctggt cgggggtttt gccggggtttt
1320
gggctgcttc ccgggtgcttc ccgtgcttgg cgggtgtgct cgggtgcttc ccgtgcttgg
1380
gcgggtgcttc ccgggtgcttc ccgtgcttgg cgggtgcttc ccgtgcttgg cgggtgcttc
1440
aacggggtgc cgggtgcttc ccgtgcttgg cgggtgcttc ccgtgcttgg cgggtgcttc
1500
gtggtgccgg ccagagggcag ccgggtgcttc ccgtgcttgg cgggtgcttc ccgtgcttgg
1560
gggtgcttc ccgtgcttttc cgggtgcttc ccgtgcttttc cgggtgcttc ccgtgcttttc
1620
gggtgcttc ccgtgcttttc cgggtgcttc ccgtgcttttc cgggtgcttc ccgtgcttttc
1680
ccggagaaact tcactgtgag ggtccatga cggcgagggc ccggcagctt ccgggctgtg 1740
gagaaacaco tctccgcaaa aaaaatgggg ttcttttttttt ttcttttttttt ttctttttc 1000
tctttttagg acgaagagga aatacaaggg acagccacag acgaggggcc gtcttccccac 1860
eggagattgcgt tcctttttttac gggccaggag cggccaggtt cggccaggtt cggccaggtt 1920
cgtggagtgt ctgcttcttc tagttagggt tcgtggccct ctgcttctgc cggctttgtg 1980
gcctgctcgc aggcgggccc tgggggtgctg ccagccacac ccagccacac ccagccacac 2040
cggcagggcc gggcggaggt cctccttttct tcgccgagagc ttggtggttg gcccagagac 2100
cggctctagc tcgtctccgg cggccagagc ttggttgcct gttggttacct aatcctcaga 2160
gttcaaaaac aataagaaac aaaaaaaaaa gtaaaaaact ctctaaaaaa 2210

<210> SEQ ID NO: 383
<211> LENGTH: 4604
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 383

ggaacagtct tcgcaacggcc cggcagagcc agaaoccttt ggctcttgag tgcttgtgag 60
acccaaacaga aagccacocccc tggggcaccct tttttgctttt cagggaggtg 120
ggacatggcg cgccccgcaact cccctgctga cccctgctga tttttgcccc cccctgctga 180
gactctgggg cccttgctga cccctgctga tttttgcccc cccctgctga tttttgcccc 240
agagtggcag cggctcctca gggagaccat gggagaccat gggagaccat gggagaccat 300
cacccggggcg tggccctgca cccgggaccc cggccagagc cggccagagc cggccagagc 360
agaggttgct gagccgaggtg tgggactcgg acagaggttgct gagccgaggtg tgggactcgg 420
gcggccgctc atgcctggtg acgagccgctc atgcctggtg acgagccgctc atgcctggtg 480
cagctctttt ccgtggtggt gcccctgccc cccctgctga tttttgcccc cccctgctga 540
cctggtgcct cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 600
tggacaggg gccttgcaag cggccagagc cggccagagc cggccagagc cggccagagc 660
tcccctctct tcgctggccc cccctgctga tttttgcccc cccctgctga tttttgcccc 720
tcgggctctga cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 780
ccgctctctc cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 840
tgggcttgtc cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 900
gcgtgcttac tggcctctag atgtagcctt gcgctgcttac tggcctctag atgtagcctt 960
cacagctgtg cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1020
agccgcccag acggctctgg cggccagagc cggccagagc cggccagagc cggccagagc 1080
gctggtggtg cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1140
tgggtctgg gcggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1200
acgctgctgtg gcggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1260
gcggggcttg gcggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1320
ctgctctctg gcggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1380
tgggcttgtc cggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1440
agggccagct gcggccagagc cggccagagc cggccagagc cggccagagc cggccagagc 1500
gttcagcccc ttcctgtgca cctctcctat ccacccagcc cattctctct cctgtcaacc 1560
cctasocctt ttttcggtgg tttttccgyc tcctctgaga cctcagccat gaggatgycg 1620
tggtgtggtg aacaagaaac ccagtttgagg gcagagggcca gaggagggag gcaagcccgtg 1680
cccgagagtgt ccctccttct ctcctgacttg ggtgctctgc gtagtgtgag gcaagcacag 1740
gagaagccat ctcctcctct agggagcag agacagcagcct ctcocccccc tgcagggccc 1800
gctagtcagc gctctagtag cgtctcctct ctcctcctct tctcagatata ataatagccc 1860
cacaggcccc atccaccccc atccagctgc ccccaaatc ctcattcccct ggttatattc 1920
tccggtggag ctagctggtg gaggaggttg tctctctctct cactgagcag ccagagccgag 1980
acacccagtg tgtgcctcag cca ggcaggggga cagggagag aacagcacaac aacagctcag caggaagaata 2040
tgtgagcctg ctcctccca aagctcagacctt ctcctgcctg ttggtctaaat gggttggtgg 2100
aggtggyccag cctcctcctag caggtacctt tagctccctt gaatggctcc caggtgcagcag 2160
aacaacatc ggcagctctgc tctcccccct cgcagctgcttc aacagggacgc 2220
caccccccct gccctccctc gctcctcctccgc caggtgtgctc ccoccaaccc 2280
tccccccaggg agaggtgctg cggagcagctt gcgcctgatct cctctctgct ccaggctggg 2340
gatgggagag aagactctctt gggcagcagcc ggtctgtctgca tgcctcctcctt gccgctccat 2400
ttggtcctg cccagcctgtgt gtcctccagct gctcgagcct actggcgcagc acoccaagag 2460
gtggggaggg agccgtccggt ggtctcgttct ttgtagctagcc cagaggtgcttg tggggtctcc 2520
gtgattaggg ttcagagccc tgtctgctt cactcacaac cccaaagacta aaaaagttgac 2580
tgtgagctgc ggggaggggt ctaacacacac atggctacctgc ttctccgacy cggagccaggc 2640
cocaccgcctt ggtacggttat ctcggcttctt ctcagttctc cctctctgtttc gcttgcgg 2700
gttaggggtg ggggctcagg ctacagccagcc cttgctgctct cggagcctgc cccatctctc 2760
agatcctcgtg ccggacagcgc atctcctagct tctcagagcgc tcagacagcgc cggagaggtg 2820
agttgaggagc aagcctctctt gcgtgctctc ggtacagcct caagtagtcc atctggccct 2880
cocctctctt ccctggccttt ctcagctctct tcagctctgtta asacccacctg ccctctgctt 2940
aatagcctaa cacaagtcttt tttattgattct cagctctctt tctcaagttca ttttgcgagt 3000
accatctctt tttacagcactt ccgctcctct cggcctcagtgt cctttgggatt ctcgagttct 3060
attggggggc gacgccttacct cagcatcacttt cctgctctactttt cttattttgct cagctctgcttt 3120
tttacagggg ttcaggtag cggagggcagttt ggtcttggttt tcagatctgcc cccctgtcttc 3180
gtgtgcctct cccagccagc aaaaagcatc aggcccttga ccccttccctt ccccttccctt 3240
cagagtcagc ggcagagggg ggtgtctctcc ccatcctgct cttcgatgggc cccctcctcct 3300
ccggagcaac cggacagggc ggtgtgctgac gctctagctt cccctgcttgg cctctgcttt 3360
ggttaggtgag cggcgcttcct cctccggcctt cctccggcctt ggcggccttgg ctccgctgctt 3420
tcagcctcttc cctccggcctt cctccggcctt ggcggccttgg ctccgctgctt 3480
tctctctactt cccctctcctt cccctctcctt cccctctcctt cccctctcctt cccctctcctt 3540
gccttaactc acagctctcc agggagctgc gcgagccgct ggcctccggc cctccggcctt 3600
tccctgaggag ccgctccttcct cctccggcctt cctccggcctt ggcggccttgg ctccgctgctt 3660
tctctctactt cccctctcctt cccctctcctt cccctctcctt cccctctcctt cccctctcctt 3720
tccctctcctt cccctctcctt cccctctcctt cccctctcctt cccctctcctt cccctctcctt 3780
1-31. (canceled)

32. A method of predicting the likelihood of long-term survival of a breast cancer patient without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising determining the expression level of an RNA transcript or its product in a breast cancer tissue sample obtained from said patient, normalized against the expression level of all RNA transcripts or their products in said breast cancer tissue sample, or of a reference set of RNA transcripts or their products, wherein the transcript is the transcript of Bcl2, wherein overexpression of Bcl2, indicates an increased likelihood of long-term survival without breast cancer recurrence.

33. The method of claim 32 further comprising determining the expression level of one or more RNA transcripts or their products, normalized against the expression level of all RNA transcripts or their products in said breast cancer tissue sample, or of a reference set of RNA transcripts or their products, wherein the transcript is the transcript of one or more genes selected from the group consisting of: FOXM1, PRAME, STK15, CEP51, Ki-67, GSTM1, CA9, PR, BBC3, NME1, SURY, GATA3, TFRC, YB-1, DPYD, GSTM5, RPS6KB1, Src, Chk1, ID1, EstR1, p27, CCNB1, XIAP, Chk2, CDC25B, IGF1R, AK055699, P13KC2A, TGF13B, BAG1, CYP3A4, EpCAM, VEGFC, pS2, hENT1, WISP1, HNF3A, NFkBP65, BRCA2, EGFR, TK1, VDR, Contig51037, pENT1, EPHX1, IF1A, DIABLO, CDH1,
HIF1α, IGFBP3, CTSB, and Her2, wherein overexpression of one or more of FOXM1, PRAME, STK15, Ki-67, CA9, NME1, SURV, TFR3, YB-1, RPS6KB1, Src, Chk1, CCNB1, Chk2, CDC25B, CYP3A4, EpCAM, VEGFC, hENT1, BRCA2, EGFR, TK1, VDR, EPHX1, IGF1R, Contig51037, CDH1, HIF1α, IGFBP3, CTSB, Her2, and pE2T1 indicates a decreased likelihood of long-term survival without breast cancer recurrence, and the overexpression of one or more of CEPG1, GSTM1, PR, BBC3, GATA3, DPYD, GSTM3, ID1, EsrR1, p27, XIAP, IGF1R, AK055699, P13KCA, TGFβ3, BAG1, p52, WISP1, HNF3A, NFKBp65, and DIABLO indicates an increased likelihood of long-term survival without breast cancer recurrence.

34. The method of claim 32 wherein the breast cancer is invasive breast carcinoma, further comprising determination of the expression levels of one or more transcripts or their products wherein the transcript is the transcript of one or more genes selected from the group consisting of: FOXM1, PRAME, STK15, CEPG1, Ki-67, GSTM1, PR, BBC3, NME1, SURV, GATA3, TFR3, YB-1, DPYD, CA9, Contig51037, RPS6KB1 and Her2.

35. The method of claim 32 wherein said breast cancer is characterized by overexpression of the estrogen receptor (ER).

36. The method of claim 35 further comprising determination of the expression levels of the transcripts of one or more transcripts or their products wherein the transcript is the transcript of one or more genes selected from the group consisting of: PRAME, FOXM1, DIABLO, EPHX1, HIF1α, VEGFC, Ki-67, IGFIR, VDR, NME1, GSTM3, Contig51037, CDC25B, CTSB, p27, CDH1, and IGFBP3.

37. The method of claim 33 wherein the expression level of 2 or more RNA transcripts is determined.

38. The method of claim 32 wherein said RNA is isolated from a fixed, wax-embedded breast cancer tissue specimen of said patient.

39-40 (canceled)

41. A method of predicting the likelihood of long-term survival of a patient diagnosed with invasive breast cancer, without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising the steps of:

1. determining the expression levels of the RNA transcripts or the expression products of a gene set selected from the group consisting of:

(a) Beit2, cyclinG1, NFKBp65, NME1, EPHX1, TOP2B, DR5, TERC, Src, DIABLO; and

(b) Beit2, PRAME, cyclinG1, FOXM1, NFKBp65, TS, XIAP, Ki67, CYP3A4, p27;

in a breast cancer tissue sample obtained from a said patient, normalized against the expression levels of all RNA transcripts or their products in said breast cancer tissue sample, or of a reference set of RNA transcripts or their products;

2. subjecting the data obtained in step (1) to statistical analysis; and

3. determining whether the likelihood of said long-term survival has increased or decreased.

42. A method of predicting the likelihood of long-term survival of a patient diagnosed with estrogen receptor (ER)-positive invasive breast cancer, without the recurrence of breast cancer, following surgical removal of the primary tumor, comprising the steps of:

1. determining the expression levels of the RNA transcripts or the expression products of a gene set selected from the group consisting of:

(a) Beit2, cyclinG1, NFKBp65, NME1, EPHX1, TOP2B, DR5, TERC, Src, DIABLO; and

(b) Beit2, PRAME, cyclinG1, FOXM1, NFKBp65, TS, XIAP, Ki67, CYP3A4, p27;

in a breast cancer tissue sample obtained from a said patient, normalized against the expression levels of all RNA transcripts or their products in said breast cancer tissue sample, or of a reference set of RNA transcripts or their products;

2. subjecting the data obtained in step (1) to statistical analysis; and

3. determining whether the likelihood of said long-term survival has increased or decreased.

43. The method of claim 41 wherein said statistical analysis is performed by using the Cox Proportional Hazards model.

44. (canceled)

45. (canceled)

46. The method of claim 32, wherein said RNA is fragmented RNA.

47. The method of claim 32, wherein said RNA is isolated from a fine needle biopsy sample.

48. The method of claim 32, further comprising creating a report summarizing the data obtained by the determination of said gene expression levels.

49. The method of claim 48, wherein said report is a report summarizing the data obtained by the determination of said gene expression levels.

50. The method of claim 49, wherein said report includes recommendation for a treatment modality of said patient.

51. The method of claim 42, wherein said statistical analysis is performed by using the Cox Proportional Hazards model.

52. The method of claim 32 further comprising the step of subjecting the expression level of said gene product to analysis using the Cox Proportional Hazards model.

53. The method of claim 32 wherein said expression level is determined using RNA obtained from a formalalin-fixed paraffin-embedded tissue sample.

54. The method of claim 32 wherein said expression level is determined by reverse transcriptase polymerase chain reaction (RT-PCR).