

An inhibitor of ATR kinase for use in a method of treating a hyper-proliferative disease

The present invention covers an inhibitor of ATR kinase, particularly of 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine (in the following called "Compound A"), for use in a method of treating a hyper-proliferative disease in a subject. Preferably the hyper-proliferative disease or the subject is characterized by one or more biomarker(s) selected from

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability.

Background

The integrity of the genome of eukaryotic cells is secured by complex signalling pathways, known as DNA damage response (DDR). Recognition of DNA damage activates DDR pathways resulting in cell cycle arrest, suppression of general translation, induction of DNA repair, and, finally, in cell survival or cell death. Proteins that directly recognize aberrant DNA structures recruit and activate kinases of the DDR pathway, such as ATR. ATR responds to a broad spectrum of DNA damage, including double-strand breaks and lesions derived from interference with DNA replication as well as increased replication stress that is observed in oncogene-driven tumor cells (e.g. Ras mutation/ upregulation, Myc upregulation, CyclinE overexpression).

ATR kinase inhibitors are specifically or generically disclosed in the following publications: J. Med. Chem. 2013, 56, 2125-2138; Exp. Rev. Mol. Med. 16, e10, 2014; WO2010054398A1; WO2010071837A1;

WO2010073034A1; WO2011143399A1; WO2011143419A1; WO2011143422A1; WO2011143423A2; WO2011143425A2; WO2011143426A1; WO2011154737A1; WO2011163527A1; WO2012138938A1; WO2012178123A1; WO2012178124A1; WO2012178125A1; WO2013049719A1; WO2013049720A1; WO2013049722A1; WO2013049859A1; WO2013071085A1; WO2013071088A1; WO2013071090A1; WO2013071093A1; WO2013071094A1; WO2013152298A1; WO2014062604A1; WO2014089379A1; WO2014143240; WO 2014143241; WO 2014143242; ACS Med. Chem. Lett. 2015. 6, 37-41; ACS Med. Chem. Lett. 2015. 6, 42-46, WO 2015085132, WO 2015187451.

2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine (in the following called “Compound A”) is a new ATR kinase inhibitor, which together with more than 400 other ATR kinase inhibitors was described in International Patent Application WO2016020320. Identification of one or more biomarkers that predict sensitivity to Compound A could result in more effective biomarker-driven targeted therapy for hyper-proliferative diseases.

No predictive markers for ATR kinase inhibitors have been identified yet in the clinical setting. However, preclinical evidence suggests a number of candidate predictive biomarkers for ATR kinase inhibitors VE-821, VX-970 and AZD6738; Williamson et al. suggest that ATR kinase inhibitors could have potential as single-agent treatments for ARID1A defective cancers (Nature Communications 7:13837 | DOI: 10.1038/ncomms13837, (2016)). According to Mohni et al. ATR pathway inhibition is synthetically lethal in VE-821 treated cancer cells with ERCC1 deficiency and loss of the structure-specific endonuclease ERCC1-XPF (ERCC4) is synthetic lethal with ATR pathway inhibitors (Cancer Res. 74, (2014), 2835-2845). Strong synthetic lethal relationships with ATR inhibition was also shown for the following genes: ATRIP, RPA, CHEK1, CLSPN, HUS1, RAD1, RAD17, TIMELESS, and TIPIN (Mohni et al., Cancer Res. 74, (2014), 2835-2845). ATR inhibition by VE-821 also seems to synergize with loss of ERCC1, ATM and XRCC1 (Mohni et al., PLOS ONE | DOI:10.1371/journal.pone.0125482 May 12, 2015; Sultana et al, PLoS One, 8(2). (2013), e57098. doi: 10.1371/journal.pone.0057098). According to Hocke et al. (Oncotarget Vol. 7, No. 6, (2016), 7080-7095) POLD1 deficiency might represent a predictive marker for treatment response towards ATR inhibitors. Flynn et al. (Science 347, (2015), 273–277) suggest that ATR kinase inhibitors may be useful for treatment of ALT-positive cancers. According to the data described by Menezes et al. (Mol. Cancer. Res. 13(1), (2015), 120-129) single-agent ATR inhibitors may have therapeutic utility in the treatment of mantle cell lymphoma with ATM loss-of-function. Middleton et al. (Oncotarget, Vol.6, No. 32, (2015), 32396- 32409) suggest that defects in ATM, BRCA2, XRCC3 and XRCC1 and high DNA-PKcs expression conferred sensitivity to VE-821 monotherapy.

According to Jones et al. (Cancer Research (2017), Author Manuscript Published OnlineFirst on October 16, 2017; DOI: 10.1158/0008-5472.CAN-17-2056) in Synovial sarcoma SS18-SSX1 or SS18-SSX2 fusion proteins induce ATR kinase inhibitor sensitivity. Nieto-Soler et al. (Oncotarget. 2016; 7:58759-58767) suggest that expression of EWS-FLI1 (also called EWSR1-FLI1) or EWS-ERG (also called EWSR1-ERG oncogenic translocations sensitizes non-ES cells to ATR inhibitors.

Remi-Buisson et al. (Cancer Res 77(17), (2017), 4567-4578) describe that APOBEC3A and APOBEC3B overexpression confers susceptibility to ATR kinase inhibitors.

Kwok et al (Lancet 26, 385, Suppl 1, (2015), S58. doi: 10.1016/S0140-6736(15)60373-7; Blood 4;127(5), (2016), 582-595. doi: 10.1182/blood-2015-05-644872) showed that AZD6738 sensitized TP53- or ATM-defective primary chronic lymphocytic leukemia (CLL) cells to chemotherapy and ibrutinib.

Ruiz et al (Mol Cell 62(2), (2016), 307-313, DOI: 10.1016/j.molcel.2016.03.006) reported that deficiency in cdc25A confers resistance to ATR inhibitors.

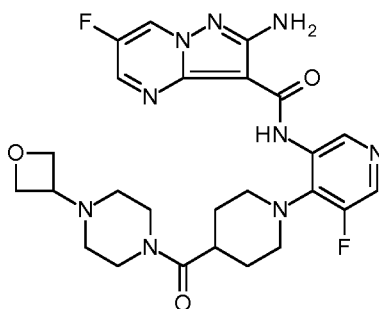
The object of the present invention is to provide one or more biomarker(s) for the treatment of one or more hyper-proliferative disease(s) with an ATR kinase inhibitor, particularly with Compound A as described herein, in a subject.

DETAILED DESCRIPTION of the INVENTION

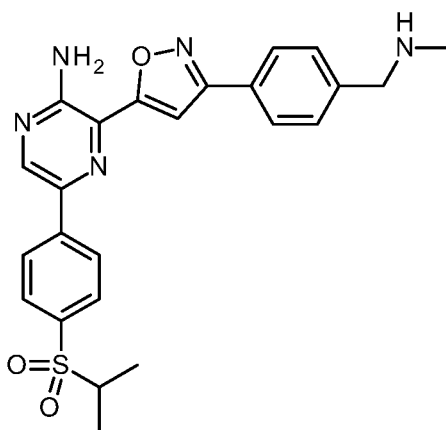
Definitions of terms used in the context of the present invention:

The term “inhibitor of ATR kinase” or the term “ATR kinase inhibitor” as used herein means any compound that inhibits ATR kinase. Examples of ATR kinase inhibitors which may be used in context with the present invention include VX-803, VX-970, AZD-6738 and preferably Compound A (described infra).

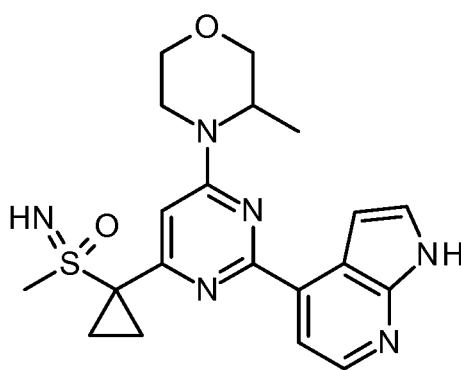
In context with the present invention the term “VX-803” means 2-amino-6-fluoro-N-[5-fluoro-4-(4-[[4-(oxetan-3-yl)piperazin-1-yl]carbonyl]piperidin-1-yl)pyridin-3-yl]pyrazolo[1,5-a]pyrimidine-3-carboxamide. VX-803 has the following structure



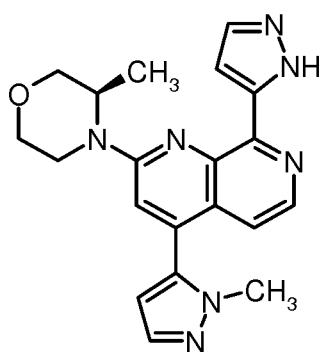
In context with the present invention the term “VX-970” means 3-(3-{4-[(methylamino)methyl]phenyl}-1,2-oxazol-5-yl)-5-[4-(propan-2-ylsulfonyl)phenyl]pyrazin-2-amine. VX-970 has the structure



In context with the present invention the term “AZD-6738” means 4-{4-[(3R)-3-methylmorpholin-4-yl]-6-[1-(S-methylsulfonylmethyl)cyclopropyl]pyrimidin-2-yl}-1H-pyrrolo[2,3-b]pyridine. AZD-6738 has the structure



The term “Compound A” as used herein means 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine of structure:



Compound A.

In particular, the term Compound A refers to 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine.

The expression “gene/protein” means one gene or one protein. The expression “gene(s)/protein(s)” means one or more gene(s) or one or more protein(s). The expression “gene(s)” means one gene or more genes. The expression “protein(s)” means one protein or more proteins.

The term “hyper-proliferative disease” includes but is not limited, e.g., psoriasis, keloids, and other hyperplasias affecting the skin, benign prostate hyperplasia (BPH), as well as malignant neoplasia. Examples of malignant neoplasia treatable with Compound A according to the present invention include solid and hematological tumors. Solid tumors can be exemplified by tumors of the breast, bladder, bone, brain, central and peripheral nervous system, colon, anus, endocrine glands (e.g. thyroid and adrenal cortex), esophagus, endometrium, germ cells, head and neck, kidney, liver, lung, larynx and hypopharynx, mesothelioma, ovary, pancreas, prostate, rectum, renal, small intestine, soft tissue, testis, stomach, skin, ureter, vagina and vulva. Malignant neoplasias include inherited cancers exemplified by Retinoblastoma and Wilms tumor. In addition, malignant neoplasias include primary tumors in said organs and corresponding secondary tumors in distant organs (“tumor metastases”). Hematological tumors can be exemplified by aggressive and indolent forms of leukemia and lymphoma, namely non-Hodgkins disease, chronic and acute myeloid leukemia (CML / AML), acute lymphoblastic leukemia (ALL), Hodgkins disease, multiple myeloma and T-cell lymphoma. Also included are myelodysplastic syndrome, plasma cell neoplasia, paraneoplastic syndromes, and cancers of unknown primary site as well as AIDS related malignancies.

Examples of breast cancer include, but are not limited to invasive ductal carcinoma, invasive lobular carcinoma, ductal carcinoma in situ, and lobular carcinoma in situ, particularly with bone metastases.

Examples of cancers of the respiratory tract include, but are not limited to small-cell and non-small-cell lung carcinoma, as well as bronchial adenoma and pleuropulmonary blastoma. Examples of brain cancers include, but are not limited to brain stem and hypophtalmic glioma, cerebellar and cerebral astrocytoma, medulloblastoma, ependymoma, as well as neuroectodermal and pineal tumor. Tumors of the male reproductive organs include, but are not limited to prostate and testicular cancer. Tumors of the female reproductive organs include, but are not limited to endometrial, cervical, ovarian, vaginal, and vulvar cancer, as well as sarcoma of the uterus. Tumors of the digestive tract include, but are not limited to anal, colon, colorectal, esophageal, gallbladder, gastric, pancreatic, rectal, small-intestine, and salivary gland cancers. Tumors of the urinary tract include, but are not limited to bladder, penile, kidney, renal pelvis, ureter, urethral and human papillary renal cancers. Eye cancers include, but are not limited to intraocular melanoma and retinoblastoma. Examples of liver cancers include, but are not limited to hepatocellular carcinoma (liver cell carcinomas with or without fibrolamellar variant), cholangiocarcinoma (intrahepatic bile duct carcinoma), and mixed hepatocellular cholangiocarcinoma. Skin cancers include, but are not limited to squamous cell carcinoma, Kaposi's sarcoma, malignant melanoma, Merkel cell skin cancer, and non-melanoma skin cancer. Head-and-neck cancers include, but are not limited to laryngeal, hypopharyngeal, nasopharyngeal, oropharyngeal cancer, lip and oral cavity cancer and squamous cell. Lymphomas include, but are not limited to AIDS-related lymphoma, non-Hodgkin's lymphoma, cutaneous T-cell lymphoma, Burkitt lymphoma, Hodgkin's disease, and lymphoma of the central nervous system. Sarcomas include, but are not limited to sarcoma of the soft tissue, osteosarcoma, malignant fibrous histiocytoma, lymphosarcoma, and rhabdomyosarcoma. Leukemias include, but are not limited to acute myeloid leukemia, acute lymphoblastic leukemia, chronic lymphocytic leukemia, chronic myelogenous leukemia, and hairy cell leukemia.

In particular, the present invention covers the treatment of lung cancer, colorectal cancer, cervical cancer, bladder cancer, breast cancer, melanoma, B-cell lymphoma, particularly diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma, prostate cancer, gliomas, ovarian cancer, glioblastoma, neuroblastoma, chronic lymphocytic leukemia (CLL), fibrosarcoma, gastric cancer, esophageal cancer, pancreatic cancer, chronic and acute myeloid leukemia (CML / AML), acute lymphoblastic leukemia (ALL), Hodgkins disease, multiple myeloma (MM) and T-cell lymphoma, endometrial cancer, vaginal cancer, and vulvar cancer, as well as sarcoma of the uterus.

Preferably, the present invention covers the treatment of prostate cancer, B-cell lymphoma, particularly diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma, melanoma, particularly malignant melanoma, ovarian, particularly, ovarian adenocarcinoma, colorectal cancer, lung, particularly non-small cell lung carcinoma, cervical cancer, and breast cancer, particularly triple-negative mammary carcinoma, pancreatic cancer, fibrosarcoma.

The term “functional mutation” as used herein means a mutation of a gene which results in an altered function of the gene, its corresponding RNA or its corresponding protein compared to the function of the respective wildtype gene, corresponding wildtype RNA or corresponding wildtype protein. The term “altered function” as used herein means either reduced or increased function of the gene, its corresponding RNA or its corresponding protein compared to the function of the respective wildtype gene, corresponding wildtype RNA or corresponding wildtype protein. The term “altered function” also includes the complete loss of the function or the gain of a new function of the gene, its corresponding RNA or its corresponding protein compared to the function of the respective wildtype gene, corresponding wildtype RNA or corresponding wildtype protein.

The reference nucleotide sequences of the cDNA's of the respective wildtype genes are described in the attached sequence protocol (SEQ ID Nos 1 to 111). The reference amino acid sequences of the respective wildtype proteins are described in the attached sequence protocol (SEQ ID Nos 112 to 222).

The functional mutation can be a “deleterious mutation” or an “activating mutation”.

The term “deleterious mutation” as used herein means a mutation of a gene which has a deleterious effect on the function of said gene or on the function of its corresponding RNA or its corresponding protein. For example, the deleterious mutation of the gene may result in a reduced gene expression level of said gene, a reduced amount or a reduced activity of the protein corresponding to said gene, or it may result in a nonfunctional gene/protein (“loss-of-function”) compared to the respective wildtype gene/protein.

Examples of a deleterious mutation include but are not limited to the following:

The deleterious mutation can be a nonsense mutation, which is a point mutation in the respective gene, resulting in a premature stop codon, or a nonsense codon in the transcribed mRNA, and in a truncated, incomplete, and nonfunctional protein corresponding to the respective gene.

The deleterious mutation can be a missense mutation, which is a point mutation in the respective gene, resulting in the production either of a nonfunctional protein (complete loss of function) or in a protein with partial loss of function compared to the respective wildtype protein.

The deleterious mutation can also result in a frameshift mutation, which is a genetic mutation in the respective gene caused by insertions or deletions of one or more nucleotides in such gene, wherein the number of nucleotides is not divisible by three, and resulting in a (sometimes truncated) nonfunctional protein corresponding to the respective gene.

The deleterious mutation can also be a large rearrangement mutation, for example a deletion of one or more exons disrupting the reading frame or a critical functional domain of the corresponding protein. Another example for a large rearrangement mutation is a duplication of one or more non-terminal exons disrupting the reading frame or a critical functional domain of the corresponding protein.

The deleterious mutation can also be a splice site mutation, which is a genetic mutation that inserts, deletes or changes a number of nucleotides in the specific site at which splicing takes place during the processing of precursor messenger RNA into mature messenger RNA. Splice site consensus sequences that drive exon recognition are located at the very termini of introns. The deletion of the splicing site results in one or more introns remaining in mature mRNA thereby resulting in the production of a nonfunctional protein corresponding to the respective gene.

The deleterious mutation can also be a copy number variant (CNV), particularly a decrease of the gene copy number (e.g. a homozygous or heterozygous deletion) compared to the normal gene copy number of the respective gene.

The term “activating mutation” as used herein means a mutation of a gene which changes said gene, its corresponding RNA and/or its corresponding protein in such a way, that its effects (e.g. the amount of corresponding RNA/protein, or the protein activity) get stronger compared to the respective wildtype gene/RNA/protein. The term “activating mutation” also includes a mutation of a gene, in which the protein corresponding to said gene gets a new function compared to the function of the corresponding wildtype protein. Examples of activating mutations include but are not limited to the following:

The activating mutation can be a substitution of one amino acid residue by another that confers a new or higher activity upon the protein.

The activating mutation can be a copy number variant (CNV), particularly an increase of the gene copy number compared to the normal gene copy number of the respective gene.

The activating mutation can also be a fusion gene or fusion protein, e.g. occurring as a result of translocation, interstitial deletion or chromosomal inversion.

The term “stratification method” as used herein means the method by which one or more of the functional mutation(s) as defined herein, particularly of the deleterious mutations and the activating mutations, the

activation of the ALT pathway and/or the microsatellite instability is (are) determined.

Preferably, the stratification method is an in-vitro method. Examples of stratification methods, which can be used in context with the present inventions, are described infra.

The term “activation of the ALT pathway” as used herein refers to cancer cells which overcome replicative senescence by activating the Alternative Lengthening of Telomeres (ALT) pathway.

The term “microsatellite instability” (“MSI”) as used herein is the expansion or reduction in the length of repetitive DNA sequences (known as microsatellites) in the DNA of a sample, e.g. a tumor sample, compared to normal cells.

MSI testing can detect an abnormal number of microsatellite repeats, which indicates that the cancer may have arisen from cells with defective mismatch repair genes. A microsatellite is a tract of tandemly repeated (i.e. adjacent) DNA motifs that range in length from one to six nucleotides, and are typically repeated 5-50 times. For example, the sequence TATATATATA is a dinucleotide microsatellite, and GTCGTCGTCGTCGTC is a trinucleotide microsatellite (with A being Adenine, G Guanine, C Cytosine, and T Thymine). Repeat units of four and five nucleotides are referred to as tetra- and pentanucleotide motifs, respectively. Microsatellites are distributed throughout the genome. Many are located in non-coding parts of the human genome, however they can also be located in regulatory regions and within the coding region. MSI tumors may result from inactivating germline mutations in one or more genes, including MLH1, MSH2, MSH6 and PMS2, and epithelial cell adhesion molecule (EPCAM), such as occurs in patients with Lynch syndrome, for whom more than 90% of colon cancers test MSI positive. MSI also occurs sporadically in several cancer types, including colorectal, endometrial, ovarian, and gastric cancers. In contrast to Lynch syndrome, sporadic MSI is often due to somatic promoter hypermethylation of MLH1 in the absence of gene sequence mutations.

The term “sample” as used herein means the sample from the subject, preferably an in vitro sample, which is used in the stratification method (as defined herein), e.g. a sample of tumor cells or of tumor tissue, a blood sample, particularly a sample of tumor tissue containing tumor cells.

Aspects of the present invention:Use(s) of the present invention

The present invention covers an inhibitor of ATR kinase, particularly of 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine (in the following called “Compound A”) or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof, particularly Compound A, for use in a method of treating a hyper-proliferative disease in a subject.

Particularly, the present invention covers an inhibitor of ATR kinase, particularly of Compound A, or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof, particularly Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject or the hyper-proliferative disease is characterized by one or more biomarker(s) defined herein.

Particularly, the present invention covers an inhibitor of ATR kinase, particularly of Compound A or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof, particularly Compound A, for use in the treatment of a hyper-proliferative disease in a subject, wherein said subject or the hyper-proliferative disease is characterized by one or more biomarker(s) defined herein.

In one embodiment of the invention said one or more biomarker(s) is (are) selected from

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or

- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment of the invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, BLM, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

The term “POLB/POLL” as used herein means a double mutation comprising one or more deleterious mutation(s) in POLB gene/protein and one or more deleterious mutation(s) in POLL gene/protein.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In a preferred embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from selected from BRCA1, ATM, FANCD2, H2AFX, RAD17, UBE2N.

In another preferred embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from selected from BRCA1, ATM, BLM, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another preferred embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another preferred embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention said one or more biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

Particularly, the present invention covers an inhibitor of ATR kinase, particularly of Compound A, or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof, particularly Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject or the hyper-proliferative disease is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) in one or more of the gene(s)/protein(s) defined herein.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from the activation of the ALT pathway.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from microsatellite instability, particularly high microsatellite instability (herein also referred to as “MSI-high”).

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by

- a) one or more biomarker(s) selected from one or more functional mutation(s) in one or more

gene(s)/protein(s) selected from APC, ARID1A, ATG5, ATM, ATR, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC5, FANCA, FANCB, FANCD2, FANCE, FANCI, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP4, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD17, RAD18, RAD50, RAD51, RAD54B, RAD54L, RB1, REV3L, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOP2A, TOP2B, TOPBP1, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2 and/or XRCC3 gene/protein; and/or

b) microsatellite instability, particularly high microsatellite instability.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by

a) one or more biomarker(s) selected from one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATG5, ATM, ATR, ATRX, BARD1, BLM, BRAF, BRCA1, BRCA2, CCND1, CCNE1, CCNE2, CDC7, CHEK1, CHEK2, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC5, FANCA, FANCD2, FANCI, FANCM, HDAC2, KRAS, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PIK3CA, POLA1, POLN, POLQ, PRKDC, PTEN, RAD17, RAD18, RAD50, RAD51, RB1, REV3L, SLX4, TDP2, TMPRSS2, TMPRSS2-ERG, TOP2A, TOP2B, TOPBP1, TP53, TP53BP1, TRRAP, UBE2N, USP1, WDR48, WRN, XPA, XRCC1, XRCC2 and/or XRCC3 gene/protein; and/or

b) microsatellite instability, particularly high microsatellite instability.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATG5, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRCA1, BRCA2, BRIP1, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, H2AFX, HDAC2, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, NBN, PALB2, PARP1, PARP2, PARP3, PARP4, PMS2, POLA1, POLB, POLH, POLN, POLN, POLQ, PRKDC, PTEN, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RAD9A, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1,

XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein, wherein the functional mutation is (are) a deleterious mutation(s).

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATG5, ATM, ATR, ATRX, BARD1, BLM, BRCA1, BRCA2, CHEK1, CHEK2, DCLRE1C, ERCC2, ERCC3, ERCC5, FANCA, FANCD2, FANCI, FANCM, HDAC2, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, NBN, PALB2, POLA1, POLN, POLQ, PRKDC, PTEN, RAD17, RAD18, RAD50, RAD51, RB1, REV3L, SLX4, TDP2, TP53, TP53BP1, TRRAP, UBE2N, USP1, WDR48, WRN, XPA, XRCC1, XRCC2 and/or XRCC3 gene/protein, wherein the functional mutation is (are) a deleterious mutation(s).

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATR, ATRIP, BRAF, CCND1, CCNE1, CCNE2, CDC7, DYRK1A, EGFR, ERBB2, ERBB3, HRAS, KRAS, MYC, NRAS, PCNA, PIK3CA, TMPRSS2, TOP2A, TOP2B, TOPBP1 and/or TP53 gene/protein, wherein the functional mutation is (are) an activating mutation(s).

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATR, BRAF, CCND1, CCNE1, CCNE2, CDC7, DYRK1A, EGFR, ERBB2, ERBB3, KRAS, MYC, NRAS, PIK3CA, TMPRSS2, TOP2A, TOP2B, TOPBP1 and/or TP53 gene/protein, wherein the functional mutation is (are) an activating mutation(s).

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from ATM, BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious

mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In a preferred embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, FANCD2, H2AFX, RAD17, UBE2N.

In a preferred embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX,

PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) of the gene(s)/protein(s), particularly deleterious and/or activating mutations, as described in Table 1 and/or in Table 2 infra:

Table 1: Deleterious mutations - examples

Gene	Short insertions/ deletions (INDELs)	Substitution-Nonsense	Substitution-Missense
APC	p.T1556fs*3 / c.4666_4667insA	p.R24* / c.70C>T	p.A2D / c.5C>A
ATG5	p.K235fs*4 / c.704delA/	p.R9* / c.25C>T	p.K58M / c.173A>T
ARID1A	p.S186fs*209 / c.557_570del14	p.Q605* / c.1813C>T	p.Q561H / c.1683G>C
ATM	p.E26fs*7 / c.73_76delAAAG	p.R250* / c.748C>T	p.R2832C / c.8494C>T
ATR	p.I774fs*5 / c.2320delA	p.E91* / c.271G>T	p.R1015Q / c.3044G>A
ATRIP	p.L63fs*8 / c.186_189delGCTT	p.E338* / c.1012G>T	p.S310F / c.929C>T
ATRX	p.D275fs*13 / c.824delA	p.G1304* / c.3910G>T	p.L192S / c.575T>C
BAP1	p.K3fs*1 / c.6_7insT	p.R60* / c.178C>T	p.G185R / c.553G>A
BARD1	p.D172fs*40 / c.513delA	p.S142* / c.425C>A	p.E268K / c.802G>A
BLM	p.N92fs*37 / c.271delA	p.W934* / c.2801G>A	p.P30L / c.89C>T
BRCA1	p.E23fs*17 / c.66_67delAG	p.Q94* / c.2801G>A /	p.M1V / c.1A>G
BRCA2	p.K437fs*22 / c.1301_1304delAAAG	p.E97* / c.289G>T	p.M1I / c.3G>A
BRIP1	p.Y313fs*25 / c.937delT	p.R261* / c.781A>T	p.D184Y / c.550G>T
CDK12	p.L21fs*10 / c.60_61delTT	p.K172* / c.514A>T	p.R890H / c.2669G>A
CHEK1	p.L355fs*1 / c.1061delT	p.R453* / c.1357A>T	p.G361D / c.1082G>A
CHEK2	p.R132fs*29 / c.394delA	p.L303* / c.908T>A	p.S428F / c.1283C>T
DCLRE1A	p.K346fs*7 / c.1038delA	p.S45* / c.134C>A	p.P137S / c.409C>T
DCLRE1B	C149fs*28 / c.443_444insT	p.W44* / c.132G>A	p.D96N / c.286G>A
DCLRE1C	p.Y99fs*7 / c.295_296insT	p.G70* / c.208G>T	p.Q137H / c.411G>C
ERCC2	p.E294fs*40 / c.880delG	p.S74* / c.221C>G	p.V231M / c.691G>A
ERCC3	p.W493fs*7 / c.1475_1476insT	p.R452* / c.1354C>T	p.D60N / c.178G>A
ERCC4	p.K916fs? / c.2743delA	p.Q5* / c.13C>T	p.T809M / c.2426C>T
ERCC5	p.E164fs*6 / c.485delA	p.C12* / c.36C>A	p.M222I / c.666G>A
FAM175A	p.E204fs*1 / c.609_610insT	p.E142* / c.424G>T	p.Y219C / c.656A>G
FANCA	p.L72fs*6 / c.215delT	p.Q1389* / c.4165C>T	p.M415I / c.1245G>A
FANCB	p.F25fs*43 / c.74delT	p.Q512* / c.1534C>T	p.L27F / c.81G>T
FANCC	p.N152fs*6 / c.455delA	p.R174* / c.520C>T	p.R245W / c.733C>T
FANCD2	p.L446fs*17 /	p.R408* / c.1222C>T	p.R1299H / c.3896G>A

Gene	Short insertions/ deletions (INDELs)	Substitution-Nonsense	Substitution-Missense
	c.1332_1333delCT		
FANCE	p.L173fs*15 / c.515_516insC	p.E235* / c.703G>T	p.Q285H / c.855G>T
FANCF	p.D27fs*54 / c.79delG	p.S18* / c.53C>G	p.R10C / c.28C>T
FANCG	p.S387fs*16 / c.1158delC	p.R102* / c.304A>T	p.L589P / c.1766T>C
FANCI	p.H1218fs*2 / c.3654delC	p.Q208* / c.622C>T	p.G119V / c.356G>T
FANCL	p.S351fs*2 / c.1051_1052delAG	p.W57* / c.170G>A	p.M74I / c.222G>A
FANCM	p.L57fs*9 / c.166_167insT	p.S1618* / c.4853C>G	p.S1665F / c.4994C>T
FBXO18	p.L116fs*1 / c.345delC	p.Q643* / c.1927C>T	p.R754Q / c.2261G>A
FBXW7	p.T165fs*4 / c.493delA	p.S294* / c.881C>G	p.R465C / c.1393C>T
FEN1		p.Q54* / c.160C>T	p.P89L / c.266C>T
GEN1	p.M44fs*1 / c.124delA	p.C117* / c.351C>A	p.R93W / c.277C>T
H2AFX	p.P49fs*13 / c.146_147delCA		p.E42K / c.124G>A
HDAC2	p.T459fs*>30 / c.1375delA	p.E185* / c.553G>T	p.V154A / c.461T>C
LIG4	p.K424fs*20 / c.1271_1275delAAAGA	p.R37* / c.109A>T	p.D165G / c.494A>G
MDC1	p.D580fs*36 / c.1738_1739delGA	p.E14* / c.40G>T	p.E149A / c.446A>C
MLH1	p.K196fs*6 / c.583delA	p.R226* / c.676C>T	p.E172K / c.514G>A
MLH3	p.N434fs*4 / c.1295_1296insA	p.Q173* / c.517C>T	p.M181I / c.543G>A
MRE11A	p.G114fs*31 / c.341delG	p.Q97* / c.289C>T	p.D86N / c.256G>A
MSH2	p.F85fs*1 / c.252_253delTT	p.Q215* / c.643C>T	p.G221V / c.662G>T
MSH3	p.D190fs*1 / c.562_563insT	p.Y227* / c.681C>G	p.N365H / c.1093A>C
MSH6	p.L290fs*1 / c.867delC	p.Q4* / c.10C>T	p.V474A / c.1421T>C
NBN	p.R466fs*18 / c.1396delA	p.R43* / c.127C>T	p.I35T / c.104T>C
PALB2	p.N186fs*4 / c.552_553insA	p.Q552* / c.1654C>T	p.Q479H / c.1437G>C
PARP1	p.P359fs*22 / c.1076delC	p.E297* / c.889G>T	p.K59N / c.177G>T
PARP2	p.R13fs*10 / c.36_37ins14	p.R395* / c.1183C>T	p.R241W / c.721C>T
PARP3	p.A300fs*29 / c.894_897delGCAG	p.Q340* / c.1018C>T	p.R524H / c.1571G>A
PARP4	p.K629fs*19 / c.1885_1888AAAG>GA	p.E83* / c.247G>T	p.G1003S / c.3007G>A

Gene	Short insertions/ deletions (INDELs)	Substitution-Nonsense	Substitution-Missense
PMS2	p.E109fs*3 / c.325delG	p.K647* / c.1939A>T	p.H24Y / c.70C>T
POLA1	p.Q32fs*4 / c.93delC	p.E276* / c.826G>T	p.E89V / c.266A>T
POLB	p.N128fs*5 / c.378delA	p.Q159* / c.475C>T	p.P251S / c.751C>T
POLH	p.F18fs*12 / c.48delT	p.Q543* / c.1627C>T	p.M14V / c.40A>G
POLL	p.C198fs*2 / c.587_588insT	p.R549* / c.1645C>T	p.A285T / c.853G>A
POLN	p.F332fs*14 / c.996delT	p.E599* / c.1795G>T	p.G419D / c.1256G>A
POLQ	p.K1068fs*2 / c.3204delA	p.R602* / c.1804C>T	p.R375W / c.1123C>T
PRKDC	p.L65fs*13 / c.194_195insT	p.E84* / c.250G>T	p.Q16K / c.46C>A
PTEN	p.K6fs*4 / c.16_17delAA	p.L25* / c.74T>A	p.I101T / c.302T>C
RAD17	p.N51fs*6 / c.147delA	p.K107* / c.319A>T	p.K370N / c.1110A>C
RAD18	p.K345fs*28 / c.1035delA	p.E152* / c.454G>T	p.K52T / c.155A>C
RAD50	p.N320fs*5 / c.954_955insA	p.W25* / c.75G>A	p.E387D / c.1161G>T
RAD51	p.Y54fs*11 / c.159_160insG	p.Q30* / c.88C>T	p.E258D / c.774G>T
RAD52	p.V105fs*7 / c.313delG	p.Q221* / c.661C>T	p.R46K / c.137G>A
RAD54B	p.P18fs*10 / c.51_52insA	p.E75* / c.223G>T	p.L528F / c.1582C>T
RAD54L	p.L113fs*10 / c.336_337insT	p.R75* / c.223C>T	p.F163L / c.489C>A
RAD9A	p.K96fs*6 / c.284delA	p.Q205* / c.613C>T	p.R150W / c.448C>T
RB1	p.I124fs*6 / c.370_371delAT	p.E54* / c.160G>T	p.V654M / c.1960G>A
REV3L	p.N639fs*16 / c.1916delA	p.E1707* / c.5119G>T	p.K1512N / c.4536A>C
RPA1	p.F222fs*3 / c.662delT	p.R586* / c.1756C>T	p.V27F / c.79G>T
RPA2	p.V207fs*26 / c.620_621delTG	p.Y97* / c.291T>A	p.G204D / c.611G>A
SLX4	p.L470fs*8 / c.1406_1407insC	p.E53* / c.157G>T	p.K301N / c.903G>T
TDP1	p.P359fs*21 / c.1073delC	p.K177* / c.529A>T	p.K292E / c.874A>G
TDP2	p.K24fs*35 / c.71delA	p.W52* / c.156G>A	p.E176D / c.528A>C
TP53	p.L35fs*8 / c.102_103insT	p.R213* / c.637C>T	p.R175G / c.523C>G (Ref 1)
TP53BP1	p.N419fs*67 / c.1256delA	p.Q106* / c.316C>T	p.F307L / c.919T>C
TRRAP	p.F468fs*52 / c.1400delT	p.R1650* / c.4948C>T	p.S722F / c.2165C>T
UBE2N	p.I75fs*6 / c.223delA	p.Q100* / c.298C>T	p.R7S / c.21G>T
UIMC1	p.T189fs*2 / c.565_566insA	p.W183* / c.549G>A	p.S44F / c.131C>T
USP1	p.N21fs*14 / c.57_58insA	p.R180* / c.538C>T	p.E250G / c.749A>G
WDR48	p.W195fs*13 / c.580_581insT	p.G107* / c.319G>T	p.R235C / c.703C>T
WRN	p.M497fs*60 / c.1485delA	p.E48* / c.142G>T	p.W85L / c.254G>T

Gene	Short insertions/ deletions (INDELs)	Substitution-Nonsense	Substitution-Missense
XPA	p.C153fs*1 / c.459_460delTG	p.E84* / c.250G>T	p.E106K / c.316G>A
XRCC1	p.G61fs*3 / c.180_181insT	p.Q134* / c.400C>T	p.R350W / c.1048C>T
XRCC2	p.K267fs*>14 / c.801delA		p.R91Q / c.272G>A
XRCC3	p.T77fs*28 / c.228_229insC		p.S23L / c.68C>T
XRCC4	p.C128fs*25 / c.380delT	p.E295* / c.883G>T	p.P14A / c.40C>G
XRCC6	p.L41fs*17 / c.116delT	p.R80* / c.238C>T	p.G28E / c.83G>A

Ref 1: Xu Y, Induction of genetic instability by gain-of-function p53 cancer mutants. *Oncogene*. 2008 27(25):3501-7.

Table 2: Activating mutations - examples

Gene	Missense alteration	Fusion
BRAF	p.V600E / c.1799T>A	AKAP9{ENST00000356239}:r.1_3551_BRAF{ENST00000288602}:r.1202_2480
EGFR	.L858R / c.2573T>G	
ERBB2	p.S1050L / c.3149C>T	
ERBB3	p.Q1239H / c.3717G>C	
PIK3CA	p.H1047R / c.3140A>G	
TMPRSS2		TMPRSS2{ENST00000332149}:r.1_79_ERG{ENST00000442448}:r.312_5034
DYRK1A	p.R559C / c.1675C>T	
PCNA	p.I88V / c.262A>G	
NRAS	p.Q61L / c.182A>T	
MYC	p.P57S / c.169C>T	
KRAS	p.G12D / c.35G>A	
HRAS	p.Q61L / c.182A>T	
CDC7	p.E25K / c.73G>A	
CCNE2	p.W327R / c.979T>C	
CCNE1	p.R240C / c.718C>T	
CCND1	p.D240H / c.718G>C	
TOP2A	p.R268H / c.803G>A	
TOP2B	p.H977Y / c.2929C>T	

Gene	Missense alteration	Fusion
TOPBP1	p.F699C / c.2096T>G	
TP53	p.R273H / c.818G>A (Ref 1)	

Further examples of deleterious/activating mutations of the gene(s) mentioned herein are described in publically available databases, such as e.g. ClinVar (Landrum MJ, Lee JM, Riley GR, et al., “ClinVar: public archive of relationships among sequence variation and human phenotype”, Nucleic Acids Res. 2014;42:D980–5; <https://www.ncbi.nlm.nih.gov/clinvar>), HGMD (the Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk/ac/index.php>; Stenson PD, Mort M, Ball EV, et al., “The human gene mutation database: 2008 update.”, Genome Med. 2009;1:13) or in “The Human Variome Project” (<http://www.humanvariomeproject.org>; Timothy D Smith and Mauno Vihinen, “Standard development at the Human Variome Project”, Database 2015, 2015), which has curated a gene-/disease- specific databases to collect the sequence variants and genes associated with diseases.

Further examples of deleterious/activating mutations of the gene(s), which may be used in context with the method(s)/use(s)/kit(s)/pharmaceutical composition(s) of the present invention, are described in COSMIC database (www.cancer.sanger.ac.uk; “COSMIC: exploring the world's knowledge of somatic mutations in human cancer”, Forbes et al., Nucleic Acids Res. 2015, Jan; 43 (Database issue):D805-11. doi: 10.1093/nar/gku1075. Epub 2014 Oct 29), particularly in release 79 of COSMIC (COSMIC v79), which was released on 14th November 2016.

Examples of relevant functional mutations of the TMPRSS2-ERG fusion gene/protein are described for example in Tomlins et al. (Science (New York, N.Y.) 2005; 310(5748):644-648); Soller et al. (Genes, chromosomes & cancer 2006; 45(7):717-719); Clark et al. (Oncogene 2007; 26(18):2667-2673); Wang et al. (Cancer research 2006; 66(17):8347- 8351); or in Tu et al. (Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc 2007, 20(9):921-928). In another embodiment of the present invention the subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from one or more functional mutation(s) of the gene(s)/protein(s) which are described in the Experimental Section infra.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the APC gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ATG5 gene.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ARID1A gene.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ATM gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker invention comprise(s) one or more functional mutation(s), particularly deleterious or activating mutation(s), of the ATR gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious or activating mutation(s), of the ATRIP gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ATRX gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BAP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BARD1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more

biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BLM gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the BRAF gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BRCA1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BRCA2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the BRIP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the CCND1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the CCNE1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the CCNE2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly

activating mutation(s), of the CDC7 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the CDK12 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the CHEK1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the CHEK2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the DCLRE1A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the DCLRE1B gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the DCLRE1C gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the DYRK1A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the EGFR gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the ERBB2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the ERBB3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ERCC2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ERCC3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ERCC4 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the ERCC5 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FAM175A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCA gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCB gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCC gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCD2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCE gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCF gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCG gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCI gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCL gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more

biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FANCM gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FBXO18 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FBXW7 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the FEN1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the GEN1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the HDAC2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the H2AFX gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the HRAS gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly

activating mutation(s), of the KRAS gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the LIG4 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MDC1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MLH1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MLH3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MRE11A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MSH2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MSH3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the MSH6 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the MYC gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the NBN gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the NRAS gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PALB2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PARP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PARP2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PARP3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PARP4 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PCNA gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the PIK3CA gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PMS2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLA1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLB gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLH gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLL gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLN gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more

biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the POLQ gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PRKDC gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the PTEN gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD9A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD17 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD18 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD50 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD51 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly

deleterious mutation(s), of the RAD52 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD54B gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RAD54L gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RB1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the REV3L gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RPA1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the RPA2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the SLX4 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the TDP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the TDP2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the TMPRSS2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious or activating mutation(s), of the TMPRSS2-ERG gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the TOPBP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the TOP2A gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly activating mutation(s), of the TOP2B gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious or activating mutation(s), of the TP53 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the TP53BP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more

biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the TRRAP gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the UBE2N gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the UIMC1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the USP1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the WDR48 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the WRN gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the XPA gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the XRCC1 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly

deleterious mutation(s), of the XRCC2 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the XRCC3 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the XRCC4 gene/protein.

In another embodiment the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), wherein the biomarker comprise(s) one or more functional mutation(s), particularly deleterious mutation(s), of the XRCC6 gene/protein.

In another embodiment of the present invention, the subject is chemotherapy-naïve.

The term “chemotherapy-naïve” as used herein means that the subject, prior to the treatment with Compound A according to the present invention, has not received a chemotherapy.

In another embodiment of the present invention, the subject has received a chemotherapy prior to the treatment with Compound A. The term “chemotherapy” as used herein means a category of cancer treatment that uses one or more chemotherapeutic agents as part of a standardized chemotherapy regimen. Chemotherapeutic agents are rather non-specific agents including but not limited to alkylating agents, anthracyclines, taxanes, epothilones, histone deacetylase inhibitors, inhibitors of topoisomerase I, inhibitors of topoisomerase II, nucleotide analogues, platinum-based agents, vinca alkaloids.

The present invention also covers an inhibitor of ATR kinase, particularly Compound A, for use in a method of treating a hyper-proliferative disease in a subject, said method comprising the steps:

- a) determining if one or more of the biomarker(s) defined herein are present in a sample, preferably in an in vitro sample, of the subject;
- b) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step a) is (are) determined positively.

The present invention also covers an inhibitor of ATR kinase, particularly Compound A, for use in a method of treating a hyper-proliferative disease in a subject said method comprising the steps:

- a) determining if one or more of the biomarker(s) selected from
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
 are present in a sample, preferably in an in vitro sample, of the subject;
- b) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps a)(i), a)(ii) and/or a)(iii) is (are) determined positively.

The present invention also covers an inhibitor of ATR kinase, particularly Compound A for use in a method of treating a hyper-proliferative disease in a subject said method comprising the steps:

- a) determining if one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein are present in a sample, preferably in an in vitro sample, of the subject;
- b) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step a) is (are) determined positively.

In another embodiment of the use of an inhibitor of ATR kinase, particularly of Compound A, in a method of treating a hyper-proliferative disease in a subject according to the present invention said method comprises the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined herein are present in the sample;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step (b) is (are) determined positively.

Particularly, the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in a method of treating a hyper-proliferative disease in a subject according to the present invention said method comprises the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined in (i), (ii) and/or (iii) are present in the sample:
 - (i) the one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of

Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps (b)(i), (b)(ii) and/or (b)(iii) is (are) determined positively.

In another embodiment of the use of an inhibitor of ATR kinase, particularly of Compound A, in a method of treating a hyper-proliferative disease in a subject according to the present invention said method comprises the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein are present in the sample;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step (b) is (are) determined positively.

In context with the present invention the term “determined positively” means that the presence of said functional mutation, said activation of the ALT pathway and/or microsatellite instability, particularly high microsatellite instability, in the sample, preferably in samples of tumor cells or tumor tissue, was confirmed, particularly by one or more of the stratification method(s) described herein.

In another embodiment the present invention covers an inhibitor of ATR kinase, particularly of Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject is selected by having one or more biomarker(s) defined herein.

In another embodiment the present invention covers an inhibitor of ATR kinase, particularly of Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject is selected by having one or more of the biomarker(s) selected from

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) as defined herein;
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers an inhibitor of ATR kinase, particularly of

Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject is selected by having one or more of the biomarker(s) selected from

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers an inhibitor of ATR kinase, particularly of Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject is selected by having one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein.

The present invention also covers an inhibitor of ATR kinase, particularly of Compound A, for use in a method of treating a hyper-proliferative disease in a subject, wherein said hyper-proliferative disease is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) as defined herein;
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention also covers an inhibitor of ATR kinase, particularly of Compound A, for the use in a method of treating a subject diagnosed with a hyper-proliferative disease,

said method comprising the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined herein are present in the sample;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step b) is (are) determined positively.

In another embodiment the present invention covers an inhibitor of ATR kinase, particularly of Compound A, for the use in a method of treating a subject diagnosed with a hyper-proliferative disease, said method comprising the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined in (i), (ii) and/or (iii) are present in the sample:
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMRSS2, TMRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps

(b)(i), (b)(ii) and/or (b)(iii) is (are) determined positively.

In another embodiment the present invention also covers an inhibitor of ATR kinase, particularly of Compound A, for the use in a method of treating a subject diagnosed with a hyper-proliferative disease, said method comprising the steps:

- a) assaying a sample, preferably an in vitro sample, from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein are present in the sample;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step b) is (are) present in the sample.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, for the preparation of a medicament for treating a hyper-proliferative disease in a subject, wherein said subject is characterized by one or more biomarker(s) defined herein.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, for the preparation of a medicament for treating a hyper-proliferative disease in a subject, wherein said subject is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) as defined herein;
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, for the preparation of a medicament for treating a hyper-proliferative disease in a subject, wherein said subject is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA,

FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or

- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for treating a hyper-proliferative disease in a subject, wherein said hyper-proliferative disease is characterized by one or more biomarker(s) defined herein.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for treating a hyper-proliferative disease in a subject, wherein said hyper-proliferative disease is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) as defined herein; and/or
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for treating a hyper-proliferative disease in a subject, wherein said hyper-proliferative disease is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17,

RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or

- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for treating a hyper-proliferative disease in a subject, wherein said hyper-proliferative disease or said subject is characterized by one or more biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein.

In another embodiment of the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for treating a hyper-proliferative disease in a subject according to the invention the one or more functional mutation(s), the activation of the ALT pathway and/or the microsatellite instability is (are) determined by one or more of the stratification method(s) described herein.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for a method of treating a hyper-proliferative disease in a subject, said method comprising the steps:

- a) assaying a sample from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined herein are present in the sample;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step (b) is (are) determined positively.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for a method of treating a hyper-proliferative disease in a subject, said method comprising the steps:

- a) assaying a sample from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined in (i), (ii) and/or (iii) are present in the sample:
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
- c) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps (b)(i), (b)(ii) and/or (b)(iii) is (are) determined positively.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for a method of treating a hyper-proliferative disease in a subject, said method comprising the steps:

- a) determining if one or more of the biomarker(s) defined herein are present in a sample of said subject;
- b) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step a) is (are) determined positively.

In another embodiment the present invention covers the use of an inhibitor of ATR kinase, particularly of Compound A, in the manufacture of a medicament for a method of treating a hyper-proliferative disease

in a subject, said method comprising the steps:

- a) determining if one or more of the biomarker(s) selected from
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
 are present in a sample of said subject;
- b) administering a therapeutically effective amount of the inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps a)(i), a)(ii) and/or a)(iii) is (are) determined positively.

Method(s) of the present invention

In another embodiment the present invention covers a method for the treatment of a hyper-proliferative disease in a subject using an effective amount of an inhibitor of ATR kinase, particularly of Compound A, wherein said subject or said hyper-proliferative disease is characterized by one or more biomarker(s) defined herein.

In another embodiment the present invention covers a method for the treatment of a hyper-proliferative disease in a subject using an effective amount of an inhibitor of ATR kinase, particularly of Compound A, wherein said subject is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) as defined herein; and/or

- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers a method for the treatment of a hyper-proliferative disease in a subject using an effective amount of an inhibitor of ATR kinase, particularly of Compound A, wherein said subject is characterized by

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMRSS2, TMRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
- b) the activation of the ALT pathway; and/or
- c) microsatellite instability, particularly high microsatellite instability.

In another embodiment the present invention covers a method for the treatment of a hyper-proliferative disease in a subject using an effective amount of an inhibitor of ATR kinase, particularly of Compound A, wherein said hyper-proliferative disease or said subject is characterized by one or more biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein.

In another embodiment of the method for the treatment of a hyper-proliferative disease in a subject using an effective amount of an inhibitor of ATR kinase, particularly of Compound A, the one or more functional mutation(s), the activation of the ALT pathway and/or the microsatellite instability is (are) determined by one or more of the stratification method(s) described herein.

The present invention also covers a method of treatment of a subject diagnosed with a hyper-proliferative disease comprising the steps

- a) assaying a sample from the subject, preferably an in vitro sample from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined herein are present in the sample;
- c) administering a therapeutically effective amount of an inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step b) is (are) determined positively.

The present invention also covers a method of treatment of a subject diagnosed with a hyper-proliferative disease comprising the steps

- a) assaying a sample from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) defined in (i), (ii) and/or (iii) are present in the sample:
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMRSS2, TMRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - (ii) the activation of the ALT pathway; and/or
 - (iii) microsatellite instability, particularly high microsatellite instability;
- c) administering a therapeutically effective amount of an inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by any one of steps

(b)(i), (b)(ii) and/or (b)(iii) is (are) determined positively.

The present invention also covers a method of treatment of a subject diagnosed with a hyper-proliferative disease comprising the steps

- a) assaying a sample from the subject, preferably an in vitro sample from the subject, particularly by one or more of the stratification method(s) described herein;
- b) determining if one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein are present in the sample;
- c) administering a therapeutically effective amount of an inhibitor of ATR kinase, particularly of Compound A, to the subject, if one or more of the biomarker(s) determined by step b) is (are) determined positively.

The present invention also concerns a method for identifying a subject having a hyper-proliferative disease disposed to respond favorably to an inhibitor of ATR kinase, particularly of Compound A, wherein the method comprises the detection of one or more of the biomarker(s) defined herein in a sample of said subject, preferably in an in vitro sample of tumor cells or of tumor tissue.

The present invention also concerns a method for identifying a subject having a hyper-proliferative disease disposed to respond favorably to an inhibitor of ATR kinase, particularly of Compound A, wherein the method comprises the detection of one or more of the biomarker(s) selected from:

- (ii) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2,

TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or

(ii) the activation of the ALT pathway; and/or

(iii) microsatellite instability, particularly high microsatellite instability;

in a sample of said subject, preferably in an in vitro sample of tumor cells or of tumor tissue.

The present invention also concerns a method for identifying a subject having a hyper-proliferative disease disposed to respond favorably to an inhibitor of ATR kinase, particularly of Compound A, wherein the method comprises the detection of one or more of the biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein in a sample of said subject, preferably in an in vitro sample of tumor cells or of tumor tissue.

In another embodiment the one or biomarker(s) is (are) determined by one or more of the stratification method(s) described herein.

The present invention also concerns a method for identifying a subject with a hyper-proliferative disease who is more likely to respond to a therapy comprising an inhibitor of ATR kinase, particularly of Compound A, than other subjects, the method comprising

- a) determining in a sample from said subject one or more of the biomarker(s) defined herein;
- b) identifying those subjects for whom in step a) one or more of the biomarker(s) is (are) determined positively.

The present invention also concerns a method for identifying a subject with a hyper-proliferative disease who is more likely to respond to a therapy comprising an inhibitor of ATR kinase, particularly of Compound A, than other subjects, the method comprising

- a) determining in a sample from said subject the biomarker(s) selected from:
 - (i) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4,

ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or

(ii) the activation of the ALT pathway; and/or

(iii) microsatellite instability, particularly high microsatellite instability; and

- b) identifying those subjects for whom one or more of the biomarker(s) of any one of a)(i), a)(ii) or a)(iii) is (are) determined positively.

The present invention also concerns a method for identifying a subject with a hyper-proliferative disease who is more likely to respond to a therapy comprising an inhibitor of ATR kinase, particularly Compound A, than other subjects, the method comprising

a) determining in a sample from said subject one or more biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein;

b) identifying those subjects for whom in step a) one or more of the biomarker(s) is (are) determined positively.

The present invention also concerns a method of determining whether a subject having a hyper-proliferative disease will respond to the treatment with an inhibitor of ATR kinase, particularly with Compound A, wherein the method comprises the detection of one or more of the biomarker(s) defined herein in a sample of said subject. Preferably the sample is a sample of tumor cells or of tumor tissue of said subject. Particularly, the biomarker(s) is (are) determined by one or more of the stratification method(s) described herein.

The present invention also concerns a method of determining the likelihood that a subject with a hyper-proliferative disease benefits from treatment with an inhibitor of ATR kinase, particularly with

Compound A, the method comprising the detection of one or more of the biomarker(s) defined herein in a sample of said subject and identifying the subject being more likely to respond to said treatment with the inhibitor of ATR kinase, particularly with Compound A, when the one or more biomarker(s) is (are) determined positively.

The present invention also covers a method of predicting whether a subject with a hyper-proliferative disease will respond to the treatment with an inhibitor of ATR kinase, particularly with Compound A, wherein the method comprises the detection of one or more of the biomarker(s) defined herein in a sample of said subject.

The present invention also covers the use of one or more of the biomarker(s) defined herein for identifying a subject with a hyper-proliferative disease who is disposed to respond favorably to an inhibitor of ATR kinase, particularly to Compound A.

Kit(s) and pharmaceutical composition(s) of the present invention

The present invention also covers a kit comprising an inhibitor of ATR kinase, particularly Compound A, together with means, preferably a detecting agent, to detect in a sample from a subject one or more of the biomarker(s) selected from:

- a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMPRSS2, TMPRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
- b) the activation of the ALT pathway; and/or

c) microsatellite instability, particularly high microsatellite instability.

The present invention also covers a kit comprising an inhibitor of ATR kinase, particularly Compound A, together with means, preferably a detecting agent, to detect, particularly in a sample from a subject, one or more of the biomarker(s) defined herein.

The present invention also covers a kit comprising an inhibitor of ATR kinase, particularly Compound A, together with means, preferably a detecting agent, to detect in a sample from a subject one or more biomarker(s) comprising one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N and/or XPA gene/protein.

In another embodiment the present invention covers a pharmaceutical composition comprising an inhibitor of ATR kinase, particularly Compound A, together with one or more pharmaceutically acceptable excipients for use in any of the method(s)/use(s) for treating a hyper-proliferative disease in a subject described herein.

The inhibitor of ATR kinase, particularly Compound A, can act systemically and/or locally. For this purpose, it can be administered in a suitable manner, for example by the oral, parenteral, pulmonal, nasal, sublingual, lingual, buccal, rectal, dermal, transdermal, conjunctival, otic route, or as an implant or stent. The inhibitor of ATR kinase, particularly Compound A, can be administered in administration forms suitable for these administration routes.

Suitable administration forms for oral administration are those which deliver Compound A in a rapid and/or modified manner, and contain Compound A in crystalline and/or amorphous and/or dissolved form, for example tablets (uncoated or coated tablets, for example with enteric or retarded-dissolution or insoluble coatings which control the release of Compound A, tablets or films/wafers which disintegrate rapidly in the oral cavity, films/lyophilizates, capsules (for example hard or soft gelatin capsules), sugar-coated tablets, granules, pellets, powders, emulsions, suspensions, aerosols or solutions.

Parenteral administration can be accomplished with avoidance of an absorption step (for example by an intravenous, intraarterial, intracardial, intraspinal or intralumbal route) or with inclusion of an absorption (for example by an intramuscular, subcutaneous, intracutaneous, percutaneous or intraperitoneal route). Suitable administration forms for parenteral administration include injection and infusion formulations in

the form of solutions, suspensions, emulsions, lyophilizates or sterile powders.

For the other administration routes, suitable examples are pharmaceutical forms for inhalation or inhalation medicaments (including powder inhalers, nebulizers), nasal drops, solutions or sprays; tablets, films/wafers or capsules for lingual, sublingual or buccal administration, films/wafers or capsules, suppositories, ear or eye preparations (for example eye baths, ocular insert, ear drops, ear powders, ear-rinses, ear tampons), vaginal capsules, aqueous suspensions (lotions, shaking mixtures), lipophilic suspensions, ointments, creams, transdermal therapeutic systems (for example patches), milk, pastes, foams, dusting powders, implants, intrauterine coils, vaginal rings or stents.

Compound A can be converted to the administration forms mentioned. This can be done in a manner known per se, by mixing with pharmaceutically suitable excipients.

These excipients include carriers (for example microcrystalline cellulose, lactose, mannitol), solvents (e.g. liquid polyethylene glycols), emulsifiers and dispersing or wetting agents (for example sodium dodecylsulphate, polyoxysorbitan oleate), binders (for example polyvinylpyrrolidone), synthetic and natural polymers (for example albumin), stabilizers (e.g. antioxidants, for example ascorbic acid), dyes (e.g. inorganic pigments, for example iron oxides) and flavour and/or odour correctants.

Pharmaceutically acceptable excipients are non-toxic, preferably they are non-toxic and inert. Pharmaceutically acceptable excipients include, inter alia: fillers and excipients (for example cellulose, microcrystalline cellulose, such as, for example, Avicel®, lactose, mannitol, starch, calcium phosphate such as, for example, Di-Cafos®),

- ointment bases (for example petroleum jelly, paraffins, triglycerides, waxes, wool wax, wool wax alcohols, lanolin, hydrophilic ointment, polyethylene glycols),
- bases for suppositories (for example polyethylene glycols, cacao butter, hard fat)
- solvents (for example water, ethanol, Isopropanol, glycerol, propylene glycol, medium chain-length triglycerides fatty oils, liquid polyethylene glycols, paraffins),
- surfactants, emulsifiers, dispersants or wetters (for example sodium dodecyle sulphate, lecithin, phospholipids, fatty alcohols such as, for example, Lanette®, sorbitan fatty acid esters such as, for example, Span®, polyoxyethylene sorbitan fatty acid esters such as, for example, Tween®, polyoxyethylene fatty acid glycerides such as, for example, Cremophor®, polyoxethylene fatty acid esters, polyoxyethylene fatty alcohol ethers, glycerol fatty acid esters, poloxamers such as, for example, Pluronic®),
- buffers and also acids and bases (for example phosphates, carbonates, citric acid, acetic acid, hydrochloric acid, sodium hydroxide solution, ammonium carbonate, trometamol, triethanolamine)

- isotonicity agents (for example glucose, sodium chloride),
- adsorbents (for example highly-disperse silicas)
- viscosity-increasing agents, gel formers, thickeners and/or binders (for example polyvinylpyrrolidone, methylcellulose, hydroxypropylmethylcellulose, hydroxypropylcellulose, carboxymethylcellulose-sodium, starch, carbomers, polyacrylic acids such as, for example, Carbopol®, alginates, gelatine),
- disintegrants (for example modified starch, carboxymethylcellulose-sodium, sodium starch glycolate such as, for example, Explotab®, cross-linked polyvinylpyrrolidone, croscarmellose-sodium such as, for example, AcDiSol®),
- flow regulators, lubricants, glidant and mould release agents (for example magnesium stearate, stearic acid, talc, highly-disperse silicas such as, for example, Aerosil®),
- coating materials (for example sugar, shellac) and film formers for films or diffusion membranes which dissolve rapidly or in a modified manner (for example polyvinylpyrrolidones such as, for example, Kollidon®, polyvinyl alcohol, hydroxypropylmethylcellulose, hydroxypropylcellulose, ethylcellulose, hydroxypropylmethylcellulose phthalate, cellulose acetate, cellulose acetate phthalate, polyacrylates, polymethacrylates such as, for example, Eudragit®),
- capsule materials (for example gelatine, hydroxypropylmethylcellulose),
- synthetic polymers (for example polylactides, polyglycolides, polyacrylates, polymethacrylates such as, for example, Eudragit®, polyvinylpyrrolidones such as, for example, Kollidon®, polyvinyl alcohols, polyvinyl acetates, polyethylene oxides, polyethylene glycols and their copolymers and blockcopolymers),
- plasticizers (for example polyethylene glycols, propylene glycol, glycerol, triacetine, triacetyl citrate, dibutyl phthalate),
- penetration enhancers,
- stabilisers (for example antioxidants such as, for example, ascorbic acid, ascorbyl palmitate, sodium ascorbate, butylhydroxyanisole, butylhydroxytoluene, propyl gallate),
- preservatives (for example parabens, sorbic acid, thiomersal, benzalkonium chloride, chlorhexidine acetate, sodium benzoate),
- colourants (for example inorganic pigments such as, for example, iron oxides, titanium dioxide),
- flavourings, sweeteners, flavour- and/or odour-masking agents.

The present invention further covers the use of pharmaceutical compositions which comprise the inhibitor

of ATR kinase, particularly Compound A, together with one or more, preferably inert, nontoxic, pharmaceutically suitable excipients, for use in any of the method(s)/use(s) for treating a hyper-proliferative disease in a subject described herein.

Based upon standard laboratory techniques known to evaluate compounds useful for the treatment of hyper-proliferative diseases by standard toxicity tests and by standard pharmacological assays for the determination of treatment of the conditions identified above in mammals, and by comparison of these results with the results of known active ingredients or medicaments that are used to treat these conditions, the effective dosage of the compounds of this invention can be determined for treatment of each desired indication. The amount of the active ingredient to be administered in the treatment of one of these conditions can vary widely according to such considerations as the particular compound and dosage unit employed, the mode of administration, the period of treatment, the age and sex of the patient treated, and the nature and extent of the condition treated.

The total amount of the active ingredient to be administered will generally range from about 0.001 mg/kg to about 200 mg/kg body weight per day, and preferably from about 0.01 mg/kg to about 50 mg/kg body weight per day. Clinically useful dosing schedules will range from one to three times a day dosing to once every four weeks dosing. In addition, "drug holidays" in which a patient is not dosed with a drug for a certain period of time, may be beneficial to the overall balance between pharmacological effect and tolerability. A unit dosage may contain from about 0.5 mg to about 1500 mg of active ingredient, and can be administered one or more times per day or less than once a day. The average daily dosage for administration by injection, including intravenous, intramuscular, subcutaneous and parenteral injections, and use of infusion techniques will preferably be from 0.01 to 200 mg/kg of total body weight. The average daily rectal dosage regimen will preferably be from 0.01 to 200 mg/kg of total body weight. The average daily vaginal dosage regimen will preferably be from 0.01 to 200 mg/kg of total body weight. The average daily topical dosage regimen will preferably be from 0.1 to 200 mg administered between one to four times daily. The transdermal concentration will preferably be that required to maintain a daily dose of from 0.01 to 200 mg/kg. The average daily inhalation dosage regimen will preferably be from 0.01 to 100 mg/kg of total body weight.

Of course the specific initial and continuing dosage regimen for each patient will vary according to the nature and severity of the condition as determined by the attending diagnostician, the activity of the specific compound employed, the age and general condition of the patient, time of administration, route of administration, rate of excretion of the drug, drug combinations, and the like. The desired mode of

treatment and number of doses of a compound of the present invention or a pharmaceutically acceptable salt or ester or composition thereof can be ascertained by those skilled in the art using conventional treatment tests.

In spite of this, it may be necessary to deviate from the amounts specified, specifically depending on body weight, administration route, individual behaviour towards the active ingredient, type of formulation, and time or interval of administration. For instance, less than the aforementioned minimum amount may be sufficient in some cases, while the upper limit mentioned has to be exceeded in other cases. In the case of administration of greater amounts, it may be advisable to divide them into several individual doses over the day.

For example, an inhibitor of ATR kinase, particularly Compound A, may be combined with known antihyperproliferative, cytostatic or cytotoxic substances for treatment of cancers. Examples of suitable antihyperproliferative, cytostatic or cytotoxic combination active ingredients include:

131I-chTNT, abarelix, abiraterone, aclarubicin, adalimumab, ado-trastuzumab emtansine, afatinib, aflibercept, aldesleukin, alectinib, alemtuzumab, alendronic acid, alitretinoin, altretamine, amifostine, aminoglutethimide, hexyl aminolevulinate, amrubicin, amsacrine, anastrozole, ancestim, anethole dithiolethione, anetumab ravtansine, angiotensin II, antithrombin III, aprepitant, arcitumomab, arglabin, arsenic trioxide, asparaginase, atezolizumab axitinib, azacitidine, basiliximab, belotecan, bendamustine, besilesomab, belinostat, bevacizumab, bexarotene, bicalutamide, bisantrene, blinatumomab, bortezomib, buserelin, bosutinib, brentuximab vedotin, busulfan, cabazitaxel, cabozantinib, calcitonine, calcium folinate, calcium levofolinate, capecitabine, capromab, carbamazepine, carboplatin, carboquone, carfilzomib, carmofur, carmustine, catumaxomab, celecoxib, celmoleukin, ceritinib, cetuximab, chlorambucil, chlormadinone, chlormethine, cidofovir, cinacalcet, cladribine, clodronic acid, clofarabine, cobimetinib, copanlisib, crisantaspase, crizotinib, cyclophosphamide, cyproterone, cytarabine, dacarbazine, dactinomycin, daratumumab, darbepoetin alfa, dabrafenib, dasatinib, daunorubicin, decitabine, degarelix, denileukin diftitox, denosumab, depreotide, deslorelin, dianhydrogalactitol, dexrazoxane, dibrospidium chloride, dianhydrogalactitol, diclofenac, dinutuximab, docetaxel, dolasetron, doxifluridine, doxorubicin, doxorubicin + estrone, dronabinol, eculizumab, edrecolomab, elliptinium acetate, elotuzumab, eltrombopag, endostatin, enocitabine, epirubicin, epitioestanol, epoetin alfa, epoetin beta, epoetin zeta, eptaplatin, eribulin, erlotinib, esomeprazole, estradiol, estramustine, ethinylestradiol, etoposide, everolimus, exemestane, fadrozole, fentanyl, filgrastim, fluoxymesterone, floxuridine, fludarabine, flutamide, folinic acid, formestane, fosaprepitant, fotemustine, fulvestrant, gadobutrol, gadoteridol, gadoteric acid meglumine, gadoversetamide, gadoxetic acid, gallium nitrate, ganirelix, gefitinib, gemcitabine, gemtuzumab, Glucarpidase, glutoxim, GM-CSF, goserelin, granisetron,

granulocyte colony stimulating factor, histamine dihydrochloride, histrelin, hydroxycarbamide, I-125 seeds, lansoprazole, ibandronic acid, ibritumomab tiuxetan, ibrutinib, idarubicin, ifosfamide, imatinib, imiquimod, improsulfan, indisetron, incadronic acid, ingenol mebutate, interferon alfa, interferon beta, interferon gamma, iobitridol, iobenguane (123I), iomeprol, ipilimumab, itraconazole, ixabepilone, ixazomib, lanreotide, lansoprazole, lapatinib, Iasocholine, lenalidomide, lenvatinib, lenograstim, lentinan, letrozole, leuporelin, levamisole, levonorgestrel, levothyroxine sodium, lisuride, lobaplatin, lomustine, lonidamine, masoprocol, medroxyprogesterone, megestrol, melarsoprol, melphalan, mepitiostane, mercaptopurine, mesna, methadone, methotrexate, methoxsalen, methylaminolevulinate, methylprednisolone, methyltestosterone, metirosine, mifamurtide, miltefosine, miriplatin, mitobronitol, mitoguazone, mitolactol, mitomycin, mitotane, mitoxantrone, mogamulizumab, molgramostim, mopidamol, morphine hydrochloride, morphine sulfate, nabilone, nabiximols, nafarelin, naloxone + pentazocine, naltrexone, nartogastim, necitumumab, nedaplatin, nelarabine, neridronic acid, netupitant/palonosetron, nivolumabpentetreotide, nilotinib, nilutamide, nimorazole, nimotuzumab, nimustine, nintedanib, nitracrine, nivolumab, obinutuzumab, octreotide, ofatumumab, olaratumab, omacetaxine mepesuccinate, omeprazole, ondansetron, oprelvekin, orgotein, orilotinod, osimertinib, oxaliplatin, oxycodone, oxymetholone, ozogamicine, p53 gene therapy, paclitaxel, palbociclib, palifermin, palladium-103 seed, palonosetron, pamidronic acid, panitumumab, panobinostat, pantoprazole, pazopanib, pegaspargase, PEG-epoetin beta (methoxy PEG-epoetin beta), pembrolizumab, pegfilgrastim, peginterferon alfa-2b, pemetrexed, pentazocine, pentostatin, peplomycin, Perflubutane, perfosfamide, Pertuzumab, picibanil, pilocarpine, pirarubicin, pixantrone, plerixafor, plicamycin, poliglusam, polyestradiol phosphate, polyvinylpyrrolidone + sodium hyaluronate, polysaccharide-K, pomalidomide, ponatinib, porfimer sodium, pralatrexate, prednimustine, prednisone, procarbazine, procodazole, propranolol, quinagolide, rabeprazole, racotumomab, radotinib, raloxifene, raltitrexed, ramosetron, ramucirumab, ranimustine, rasburicase, razoxane, refametinib, regorafenib, risedronic acid, rhenium-186 etidronate, rituximab, rolapitant, romidepsin, romiplostim, romurtide, roniciclib, samarium (153Sm) leixidronam, sargramostim, satumomab, secretin, siltuximab, sipuleucel-T, sizofiran, sobuzoxane, sodium glycididazole, sonidegib, sorafenib, stanozolol, streptozocin, sunitinib, talaporfin, talimogene laherparepvec, tamibarotene, tamoxifen, tapentadol, tasonermin, teceleukin, technetium (99mTc) nofetumomab merpentan, 99mTc-HYNIC-[Tyr3]-octreotide, tegafur, tegafur + gimeracil + oteracil, temoporfin, temozolomide, temsirolimus, teniposide, testosterone, tetrofosmin, thalidomide, thiotepa, thymalfasin, thyrotropin alfa, tioguanine, tocilizumab, topotecan, toremifene, tositumomab, trabectedin, trametinib, tramadol, trastuzumab, trastuzumab emtansine, treosulfan, tretinoin, trifluridine + tipiracil, trilostane, triptorelin, trametinib, trofosfamide, thrombopoietin, tryptophan, ubenimex, valatinib,

valrubicin, vandetanib, vapreotide, vemurafenib, vinblastine, vincristine, vindesine, vinflunine, vinorelbine, vismodegib, vorinostat, vorozole, yttrium-90 glass microspheres, zinostatin, zinostatin stimalamer, zoledronic acid, zorubicin.

Biomarker(s) of the hyper-proliferative-disease or subject

In another embodiment of the use(s)/method(s)/pharmaceutical composition(s)/kit(s) of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, BLM, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the use(s)/method(s)/pharmaceutical composition(s)/kit(s) of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, BLM, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

The expression "POLB/POLL" as used herein means a double mutation: one or more deleterious

mutation(s) in POLB gene/protein and one or more deleterious mutation(s) in POLL gene/protein.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, BLM, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from ATM, BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In another preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, FANCD2, H2AFX, RAD17, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, FEN1, H2AFX, PCNA.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BLM, BRCA1, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, REV3L, TP53BP1, UBE2N.

In a preferred embodiment of the use/method/pharmaceutical composition/kit of the invention described

herein the hyper-proliferative disease or the subject is characterized by one or more biomarker(s), wherein the biomarker(s) comprise(s) one or more deleterious mutation(s) in TP53 gene/protein and one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, FEN1, H2AFX, PCNA.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in BLM gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in BRCA1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in ERCC5 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in FEN1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in FANCD2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in FANCG gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described

herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in H2AFX gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in PARP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in PCNA gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in POLL gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in RAD9A gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in RAD17 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in RAD52 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in REV3L gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious

mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in TDP2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in TP53BP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in UBE2N gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TP53 gene/protein and by one or more deleterious mutation(s) in XPA gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in ATM gene/protein and/or by one or more deleterious mutation(s) in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in POLL gene/protein and by one or more deleterious mutation(s) in POLB gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in POLN gene/protein and by one or more deleterious mutation(s) in POLQ gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in POLH gene/protein and by one or more deleterious mutation(s) in REV3L gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention described herein the hyper-proliferative disease or the subject is characterized by one or more deleterious mutation(s) in TDP1 gene/protein and by one or more deleterious mutation(s) in TDP2 gene/protein.

Prostate Cancer

In another embodiment of the use/methods/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is prostate cancer.

The term “prostate cancer” as used herein means any histology type of prostate cancer including but not limited to acinar adenocarcinoma, ductal adenocarcinoma, transitional cell (or urothelial) cancer, squamous cell cancer, carcinoid, small cell cancer, sarcomas and sarcomatoid cancers, particularly acinar adenocarcinoma, castration resistant prostate cancer (CRPC), particularly stage M0 castration-resistant prostate cancer (M0 CRPC) or stage M1 castration-resistant prostate cancer (M1 CRPC).

The terms “M0” and “M1” (including M1a, M1b, M1c) are used in accordance with the “TNM staging system” for prostate cancer developed by the American Joint Committee on Cancer as further described in “TNM CLASSIFICATION OF MALIGNANT TUMORS”, 7th edition Edited by James D. Brierley, Mary K. Gospodarowicz, Christian Wittekind, Published by UICC 2011.

According to said TNM classification and as used herein the term “M0 CRPC” means, that there are no distant metastases and that the CRPC has not spread to other parts of the body. The term “M1 CRPC” as used herein means that there are distant metastases and that the CRPC has spread to distant parts of the body.

In another embodiment of the present invention, the castration resistant prostate cancer (CRPC) is stage M0 castration resistant prostate cancer (M0 CRPC) or stage M1 castration-resistant prostate cancer (M1 CRPC).

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATM, ARID1A, ATG5, ATR, ATRX, BARD1, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CDC7, CHEK2, DCLRE1C, DYRK1A, EGFR, ERBB3, ERCC3, ERCC5, FANCA, FANCB, FANCD2, FANCI, GEN1, HDAC, KRAS, LIG4, MLH1, MLH3, MSH2, MSH3, MSH6, MYC, NBN, PALB2, PARP4, PIK3CA, PMS2, POLA1, POLL, PRKDC, PTEN, RAD18, RAD50, RAD51, RB1, REV3L, SLX4, TMRSS2, TMRSS2-ERG, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, USP1, WDR48, WRN, XPA, XRCC1 and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ARID1A, ATG5, ATR, ATRX, BARD1, BRAF, BRCA2, CCND1, CDC7, DCLRE1C, DYRK1A, EGFR, ERBB3, FANCA, FANCD2, FANCI, KRAS, MSH2, MSH3, MSH6, MYC, PIK3CA, POLA1, PRKDC, PTEN, RAD50, RAD51, RB1, REV3L, SLX4, TMPRSS2-ERG, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, USP1, WDR48, WRN, XPA, XRCC1 and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by microsatellite instability, particularly high microsatellite instability.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ARID1A, ATM, ATR, ATRX, BARD1, BRAF, BRCA2, CCND1, CDC7, DCLRE1C, EGFR, ERBB3, FANCA, FANCD2, MSH2, MSH3, MSH6, MYC, PIK3CA, POLA1, PTEN, RAD50, RAD51, REV3L, SLX4, TMPRSS2-ERG, TOP2A, TOP2B, USP1, WDR48, and/or WRN gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ARID1A, ATR, ATRX, BARD1, BRAF, BRCA2, CCND1, CDC7, DCLRE1C, EGFR, ERBB3, FANCA, FANCD2, MSH2, MSH3, MSH6, MYC, PIK3CA, POLA1, PTEN, RAD50, RAD51, REV3L, SLX4, TMPRSS2-ERG, TOP2A, TOP2B, USP1, WDR48, and/or WRN gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) of the ATM gene/protein, particularly by a deleterious mutation of the ATM gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-

proliferative disease is prostate cancer, particularly M0 CRPC or M1 CRPC, and the subject or the prostate cancer is characterized by one or more functional mutation(s) in at least five gene(s)/protein(s) selected from APC, ATM, ARID1A, ATG5, ATR, ATRX, BARD1, BRAF, BRCA1, BRCA2, BRIP1, CCND1, CDC7, CHEK2, DCLRE1C, DYRK1A, EGFR, ERBB3, ERCC3, ERCC5, FANCA, FANCB, FANCD2, FANCI, GEN1, HDAC, KRAS, LIG4, MLH1, MLH3, MSH2, MSH3, MSH6, MYC, NBN, PALB2, PARP4, PIK3CA, PMS2, POLA1, POLL, PRKDC, PTEN, RAD18, RAD50, RAD51, RB1, REV3L, SLX4, TMRSS2, TMRSS2-ERG, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, USP1, WDR48, WRN, XPA, XRCC1 and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is prostate cancer and the subject or the prostate cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is prostate cancer and the subject or the prostate cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the prostate cancer cell lines, particularly the subject or the prostate cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Ovarian Cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is ovarian cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATM, ATR, BRAF, BRCA1, BRCA2, CDC7, CHEK1, ERBB2, ERBB3, FANCA, FANCM, FBXW7, KRAS, MLH1, MRE11A, MSH3, MSH6, MYC, PALB2, PARP4, PIK3CA, POLH, POLN, POLQ, PRKDC, PTEN, RAD50, REV3L, TDP1, TP53, TOP2A, TOP2B and/or TOPBP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or

more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATM, ATR, BRAF, BRCA1, CDC7, CHEK1, ERBB2, ERBB3, FANCM, KRAS, MLH1, MSH6, MYC, PIK3CA, POLN, POLQ, PRKDC, PTEN, RAD50, TP53, TOP2A, TOP2B and/or TOPBP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more functional mutation(s), particularly deleterious mutations, in one or more gene(s)/protein(s) selected from APC, ARID1A, ATR, BRAF, BRCA1, CDC7, CHEK1, ERBB3, FANCM, MLH1, MSH6, PIK3CA, POLN, POLQ, PRKDC, PTEN, RAD50, TOP2A, TOP2B and/or TOPBP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more functional mutation(s), particularly deleterious mutations, in one or more gene(s)/protein(s) selected from APC, ATR, BRAF, CDC7, FANCM, PRKDC, TOP2B and/or TOPBP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more functional mutation(s) of the ATM gene/protein, particularly by a deleterious mutation of the ATM gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by microsatellite instability, particularly high microsatellite instability.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is ovarian cancer and the subject or the ovarian cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the ovarian cancer cell lines,

particularly the subject or the ovarian cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Colorectal Cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is colorectal cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATM, ATRX, BLM, BRAF, BRCA2, CDK12, CHEK2, ERBB3, ERCC3, ERCC5, FANCA, FANCM, FBXW7, FBX018, GEN1, KRAS, MLH1, MSH2, MSH3, MSH6, MYC, NBN, PIK3CA, POLH, POLN, POLQ, PRKDC, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53, USP1, WRN and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ARID1A, ATM, ATRX, BLM, BRAF, BRCA2, CHEK2, ERBB3, ERCC5, FANCA, FANCM, KRAS, MLH1, MSH2, MSH3, MSH6, MYC, NBN, PIK3CA, POLN, POLQ, PRKDC, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53, USP1 and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in at least three gene(s)/protein(s) selected from APC, ARID1A, ATM, ATRX, BLM, BRAF, BRCA2, CDK12, CHEK2, ERBB3, ERCC3, ERCC5, FANCA, FANCM, FBXW7, FBX018, GEN1, KRAS, MLH1, MSH2, MSH3, MSH6, MYC, NBN, PIK3CA, POLH, POLN, POLQ, PRKDC, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53, USP1, WRN and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ARID1A, ATM, ATRX, BLM, BRAF, BRCA2, CHEK2, ERCC5, FANCA, FANCM, KRAS, MLH1, MSH2, MSH3, MSH6, MYC, NBN, PIK3CA, POLN, POLQ, PRKDC, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53, USP1

and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ARID1A, ATM, ATRX, BLM, BRAF, BRCA2, CHEK2, ERCC5, FANCA, KRAS, MLH1, MSH2, MSH3, MSH6, MYC, NBN, PIK3CA, PRKDC, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53, USP1 and/or XRCC2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATRX, BRAF, BRCA2, ERCC5, FANCA, MLH1, MSH3, MSH6, MYC, PIK3CA, RAD50, REV3L, SLX4, TOP2A, TOP2B, TP53 and/or USP1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by microsatellite instability, particularly high microsatellite instability.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is colorectal cancer and the subject or the colorectal cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the colorectal cancer cell lines, particularly the subject or the colorectal cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Lung cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is lung cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ATR, BARD1, BRCA1, BRIP1, FANCD2, FANCI, CCNE1, CDK12, KRAS, MDC1, MSH3, MYC, NBN, NRAS, PIK3CA, PMS2, PRKDC, RAD50L, REV3L, SLX4, TOP2B, TP53 and/or XRCC3 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ATR, CCNE1, KRAS, MSH3, MYC, NRAS, PIK3CA, PRKDC, SLX4, TOP2B, TP53 and/or XRCC3 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ATR, CCNE1, NRAS, SLX4, TOP2B, TP53 and/or XRCC3 gene/protein

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, CCNE1, KRAS, MSH3, MYC, NRAS, PRKDC, SLX4, TOP2B and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATR, CCNE1, MSH3, PRKDC and/or KRAS gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is lung cancer and the subject or the lung cancer is characterized by one or more

biomarker(s) described in the Experimental Section for one or more of the lung cancer cell lines, particularly the subject or the lung cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Melanoma

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is melanoma.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is melanoma and the subject or the melanoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, BRAF, PRKDC and/or XRCC3 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is melanoma and the subject or the melanoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, BRAF and/or PRKDC gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is melanoma and the subject or the melanoma is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is melanoma and the subject or the melanoma is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the melanoma cell lines, particularly the subject or the melanoma is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Cervical cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is cervical cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is cervical cancer and the subject or the cervical cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRIP1, EGFR, REV3L and/or UIMC1 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is cervical cancer and the subject or the cervical cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from EGFR and/or REV3L gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is cervical cancer and the subject or the cervical cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is cervical cancer and the subject or the cervical cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the cervical cancer cell lines, particularly the subject or the cervical cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Breast cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is breast cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATR, BLM, BRCA1, BRCA2, ERBB2, FANCA, FANCE, FANCI, FBXO18, MLH3, MSH3, MYC, PRKDC, PTEN, RB1, SLX4, TP53 and/or Tmprss2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more

functional mutation(s) in one or more gene(s)/protein(s) selected from ATR, BRCA1, ERBB2, FANCA, MSH3, MYC, PRKDC, PTEN, RB1, TP53 and/or TMRSS2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, MSH3, MYC, PRKDC, RB1, TP53 and/or TMRSS2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from MSH3, PTEN, RB1 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is breast cancer and the subject or the breast cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the breast cancer cell lines, particularly the subject or the breast cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Pancreatic cancer

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is pancreatic cancer.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ARID1A, BRAF, DYRK1A, ERCC2, FBXW7, KRAS, MLH1, PALB2, PARP4, PRKDC and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRAF, DYRK1A, KRAS, PRKDC and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRAF, DYRK1A, and/or PRKDC gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRAF, PRKDC and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from DYRK1A, KRAS, PRKDC and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is pancreatic cancer and the subject or the pancreatic cancer is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the pancreatic cancer cell lines, particularly the subject or the pancreatic cancer is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Mantle cell lymphoma

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is mantle cell lymphoma.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or the mantle cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ATR, ATRX, BAP1, BRCA1, CHEK2, DCLRE1A, ERCC2, FANCM, KRAS, MLH3, MSH3, POLN, PRKDC, RB1, SLX4, TMPRSS2 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or the mantle cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, ATR, FANCM, KRAS, MLH3, MSH3, PRKDC, RB1, TMPRSS2 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or the mantle cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from FANCM, KRAS, MSH3 and/or TMPRSS2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or the mantle cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from FANCM, MSH3 and/or PRKDC gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or mantle cell lymphoma is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyperproliferative disease is mantle cell lymphoma and the subject or the mantle cell lymphoma is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the mantle cell lymphoma cell lines, particularly the subject or the mantle cell lymphoma is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Diffuse Large B-cell lymphoma (DLBCL)

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is diffuse large B-cell lymphoma.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from APC, ATM, BRAF, BRIP1, CDC7, ERCC2, FANCD, FEN1, PRKDC, MLH1, MYC, REV3L, TOP2A and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, BRAF, CDC7, PRKDC, MYC, TOP2A and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATM, BRAF, CDC7, MYC, PRKDC, TOP2A and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from MYC gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in at least two, particularly at least three gene(s)/protein(s) selected from APC, ATM, BRAF, BRIP1, CDC7, ERCC2, FANCD, FEN1, PRKDC, MLH1, MYC, REV3L, TOP2A and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell

lymphoma is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is diffuse large B-cell lymphoma and the subject or the diffuse large B-cell lymphoma is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the diffuse large B-cell lymphoma cell lines, particularly the subject or the diffuse large B-cell lymphoma is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Glioblastoma

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is glioblastoma.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or the glioblastoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATRX, CCNE1, ERBB2, FANCA, PRKDC, PTEN, RAD50, RAD54, TDP2 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or glioblastoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATRX, CCNE1, ERBB2, PRKDC, PTEN, TDP2 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or glioblastoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from CCNE1, ERBB2, PRKDC, PTEN, TDP2 and/or TP53 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or glioblastoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from ATRX and/or PTEN gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or the glioblastoma is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is glioblastoma and the subject or the glioblastoma is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the glioblastoma cell lines, particularly the subject or the glioblastoma is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Neuroblastoma

In another embodiment of the use/method/pharmaceutical compositions/kits of the present invention the hyper-proliferative disease is neuroblastoma.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is neuroblastoma and the subject or neuroblastoma is characterized by one or more functional mutation(s) in one or more gene(s)/protein(s) selected from CHEK2, MSH3 and/or PRKDC gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is neuroblastoma and the subject or the neuroblastoma is characterized by one or more functional mutation(s), particularly deleterious mutation(s), in ATM gene/protein and/or in BRCA2 gene/protein.

In another embodiment of the use/method/pharmaceutical composition/kit of the invention the hyper-proliferative disease is neuroblastoma and the subject or the neuroblastoma is characterized by one or more biomarker(s) described in the Experimental Section for one or more of the neuroblastoma cell lines, particularly the subject or the neuroblastoma is characterized by one or more functional mutation(s) in one or more genes which are described in Table 5.

Stratification Methods

Various stratification methods can be used in context of the present invention to identify one or more functional mutation(s) in one or more gene(s), an activation of the ALT pathway and/or microsatellite instability (MSI) in a sample.

Functional mutation(s)

The determination of functional mutations, particularly of deleterious and activating mutations, of gene(s)/protein(s) is known to the person skilled in the art. Deleterious mutations and activating mutations can be, for example, determined by one or more of the following stratification methods: Next generation sequencing (NGS) (Metzker ML, “Sequencing technologies—the next generation”, *Nat Rev Genet.* 2010;11:31–46); Sanger sequencing and other first generation sequencing methods (Lilian T. C. Franca, Emanuel Carrilho and Tarso B. L. Kist, A review of DNA sequencing techniques, *Quarterly Reviews of Biophysics* 35, 2 (2002), pp. 169–200); PCR, particularly multiplex PCR; Fluorescence in situ hybridization (FISH); array comparative genomic hybridization (array CGH); single nucleotide polymorphism microarray (SNP microarrays), in particular to determine copy number variants (CNVs); or immunohistochemistry (IHC), in particular to determine the loss or overexpression of the respective protein.

The term “NGS” does not denote a single technique; rather, it refers to a diverse collection of post-Sanger sequencing technologies developed in the last decade. These methods include sequencing-by-synthesis (Ronaghi M et al., “A sequencing method based on real-time pyrophosphate”, *Science.* 1998;281:363–365), sequencing-by-ligation (Shendure J et al., “Accurate multiplex polony sequencing of an evolved bacterial genome”, *Science.* 2005;309:1728–32.16), ion semiconductor sequencing (Rothberg JM et al., “An integrated semiconductor device enabling non-optical genome sequencing.”, *Nature.* 2011;475:348–52.17), and others.

Bioinformatics approaches are used for detecting and analyzing the sequence variants from NGS data (Teng S, “NGS for Sequence Variants.”, *Adv Exp Med Biol.* 2016;939:1-20). NGS variant detection consists of quality control (to remove potential artifacts and bias from data), sequence alignment (reads are mapped to positions on a reference genome), and variant calling (which is performed by comparing the aligned reads with known reference sequences to find which segments are different with the reference genomes).

The sequence variants detected from NGS can be classified to single nucleotide variants (SNVs), small insertions and deletions (INDELs), and large structural variants (SVs) based on their sequences in length.

SNVs, the most common type of sequence variants, are single DNA basepair differences in individuals. INDELs are defined as small DNA polymorphisms including both insertions and deletions ranging from 1 to 50 bp in length. SVs are large genomic alterations (>50 bp) including unbalanced variants (deletions, insertions, or duplications) and balanced changes (translocations and inversions). Copy number variants (CNVs), a large category of unbalanced SVs, are DNA alterations that result in the abnormal number of copies of particular DNA segments.

Variant analysis includes variant annotation which can be used to determine the effects of sequence variants on genes and proteins and filter the functional important variants from a background of neutral polymorphisms.

Variant association analyses connects the functional important variants with complex diseases or clinical traits. The disease-related casual variants can be identified by combining these approaches. Results of these variant analysis are stored in public databases, such as for example COSMIC (the Catalogue Of Somatic Mutations In Cancer, [www. cancer.sanger.ac.uk](http://www.cancer.sanger.ac.uk)), ClinVar (Landrum MJ, Lee JM, Riley GR, et al., “ClinVar: public archive of relationships among sequence variation and human phenotype.”, *Nucleic Acids Res.* 2014;42:D980–5), HGMD (Stenson PD, Mort M, Ball EV, et al., “The human gene mutation database: 2008 update.”, *Genome Med.* 2009;1:13) or “The Human Variome Project” (<http://www.humanvariomeproject.org/>), which has curated the gene-/disease- specific databases to collect the sequence variants and genes associated with diseases.

As described above, public data bases, relevant literatures and ongoing evidences associated with the recurrence and function of the gene are used to determine the reportable status of an alteration found from NGS data for the genes of interest. Functional mutations can be classified by any one of the following reportable status: deleterious mutation(s) and activation mutation(s).

Activation of the ALT pathway

In normal somatic cells, significant telomere shortening leads to p53-dependent senescence or apoptosis (Heaphy and Meeker, *J Cell Mol Med.* 15(6): 1227-1238 (2011)). Cancer cells rely on telomerase or the alternative lengthening of telomeres (ALT) pathway to overcome replicative mortality. Most neoplastic cells express telomerase to support immortalization and tumor progression. However, approximately 10% to 15% of cancers achieve immortalization via a telomerase-independent mechanism of telomere lengthening, the alternative lengthening of telomeres (ALT) (Cesare A.J., Reddel R.R. Alternative lengthening of telomeres: Models, mechanisms and implications. *Nat. Rev. Genet.* 2010;11:319–330.). ALT is a recombination-based mechanism of telomere maintenance characterized by heterogeneous, fluctuating telomere lengths, high levels of telomere sister chromatid exchanges (t-SCEs), abundant

extrachromosomal telomeric repeat DNA (ECTR), and a specialized telomeric DNA nuclear structure termed ALT-associated promyelocytic leukemia (PML) bodies (APBs) (Robert L. Dilley, Roger A. Greenberg, ALTernative Telomere Maintenance and Cancer, Trends Cancer. 2015 Oct 1; 1(2): 145–156). Specific mutational events including recurrent mutations of the Alpha Thalassemia/Mental Retardation Syndrome X-Linked (ATRX) or Death-Domain Associated Protein (DAXX) genes have been reported to influence ALT activation and maintenance (Amorim et al, The Role of ATRX in the Alternative Lengthening of Telomeres (ALT) Phenotype, Genes (Basel). 2016 Sep; 7(9): 66.). Recent studies have implicated the long-noncoding RNA telomeric repeat-containing RNA (TERRA), nuclear receptors, and RPA in the recombinogenic potential of ALT telomeres. ATR protein kinase, a critical regulator of recombination, recruited by the Replication Protein A might be involved in ALT regulation and become a viable target for treatment of ALT tumors (Flynn, R.L.; Cox, K.E.; Jeitany, M.; Wakimoto, H.; Bryll, A.R.; Ganem, N.J.; Bersani, F.; Pineda, J.R.; Suva, M.L.; Benes, C.H.; et al. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. Science 2015, 347, 273–277.)

Stratification methods known in the art can be used to identify subjects as having a cancer associated with activation of the ALT pathway (i.e., for identifying a cancer as associated with ALT activation, also referred to herein as an ALT cancer or an ALT+ cancer). For example, detection of maintenance of telomeres in the absence of telomerase activity (Bryan et al., EMBO J., 14:4240- 4248 (1995)); detection of a pattern of telomere lengths, e.g., by terminal restriction fragment Southern blots, ranging from very short to extremely long, and with a modal length approximately twice that in comparable telomerase-positive or normal cells (Bryan et al., EMBO J., 14:4240-4248 (1995); Gollahon et al., Oncogene, 17:709-717 (1998)), detection of rapid, unsynchronized changes in telomere length cause telomere length heterogeneity (Mumane et al., EMBO J., 13:4953-4962 (1994)), detection of ALT-associated PML bodies (APBs) (Yeager et al., Cancer Res., 59:4175-4179 (1999)), detection of copying of engineered telomeric tags from one telomere to another (Pickett et al. EMBO J., 28:799-809 (2009)), detection of tandem repeat instability at telomeres and the MS32 minisatellite (Jeyapalan et al., Hum. Mol. Genet., 14: 1785-1794 (2005)), detection of Telomere-sister chromatid exchange (T-SCE) (Fan et al. Nucleic Acids Res., 37:1740-1754 (2009)), detection of an increase in the level of telomeric t-circles (Cesare et al., Mol. Cell. Biol., 24:9948-9957 (2004)), detection of single stranded C-strand telomeric DNA (ss-C-strand) (Grudic et al., Nucleic Acids Res., 35:7267-7278 (2007)), detection of C circles (Henson et al., Nat. Biotechnol., 27:1181-1185 (2009)). See, e.g. Henson and Reddel, FEBS Lett. 584(17):3800-3811 (2010); and US20150247866. In some embodiments, a branched DNA assay in RNA in situ hybridization (RNA-ISH) is used, e.g., as described in WO2015/123565.

The activation of the ALT pathway is preferably determined by one of the stratification methods described above.

Microsatellite instability (MSI)

MSI analysis involves comparing allelic profiles of microsatellite markers generated by amplification of DNA from matching normal wildtype and test samples, which may be mismatch-repair (MMR) deficient. Alleles that are present in the test sample but not in corresponding normal wildtype samples indicate MSI. MSI can be for example analyzed by a MSI-PCR method which includes fluorescently labelled primers for co-amplification of microsatellite markers, by MSI- immunohistochemistry (IHC) staining of four MMR pathway proteins: MLH1, PMS2, MSH2, or MSH6, or by computational methods using next generation DNA sequencing (NGS) data detecting an abnormal number of microsatellite repeats:

The term “high microsatellite instability” (herein also called “MSI-high”) means that a significant number, particularly at least one, preferably at least two, of microsatellite markers were found:

MSI status can be determined by a MSI-PCR Analysis System (e.g. by Promega Corp, Madison, USA) which is based on the use of five nearly monomorphic mononucleotide microsatellite markers (BAT-25, BAT-26, NR- 21, NR-24, and MONO-27). In this system high microsatellite instability (“MSI-high”) is defined as the phenotype, in which at least 2 of the tested microsatellite markers (BAT-25, BAT-26, NR-21, NR-24, and MONO-27) are altered in the sample of the subject compared to a normal reference sample. Low microsatellite instability (“MSI-low”) is defined as the phenotype, in which only one of the tested microsatellite markers is altered in the sample of the subject compared to a normal reference sample. Microsatellite stable (MSS) is defined as the phenotype, in which none of the tested microsatellite markers is altered in the sample of the subject compared to a normal reference sample. MSI status can also be determined by a MSI-PCR method using the five microsatellite loci (BAT-25, BAT-26, D2S123, D5S346, and D17S250) recommended by the National Cancer Institute (NCI), which are amplified in a single multiplex PCR reaction. In this system high microsatellite instability (“MSI-high”) is defined as the phenotype, in which at least two of the tested mononucleotide microsatellite markers (BAT-25, BAT-26, D2S123, D5S346, and D17S250) are altered in the sample of the subject compared to a normal reference sample. Low microsatellite instability (“MSI-low”) is defined as the phenotype, in which only one of the tested mononucleotide microsatellite markers (BAT-25, BAT-26, D2S123, D5S346, and D17S250) is altered in the sample of the subject compared to a normal reference sample. Microsatellite stable (MSS) is defined as the phenotype, in which none of the tested microsatellite markers (BAT-25, BAT-26, D2S123, D5S346, and D17S250) is altered in the sample of the subject

compared to a normal reference sample (Boland CR, et al, A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 1998; 58(22):5248-5257).

In this context a “normal reference sample” used for MSI testing can be for example a genomic DNA template provided by the assay kit, or DNA isolated from blood or from another non-cancerous tissue from the subject to be tested.

MSI status can be assessed by computational methods to using next generation DNA sequencing (NGS) data generated from tumor or other tissues. These computational methods include but are not limited to mSINGS (Salipante, S.J. et al, “Microsatellite instability detection by next generation sequencing”, *Clin. Chem.* 60, 1192–1199, 2014), MSISensor (Niu B et al., “MSISensor: microsatellite instability detection using paired tumor-normal sequence data.”, *Bioinformatics.* 2014; 30(7):1015-1016.), MANTIS (Microsatellite Analysis for Normal Tumor InStability) (Kautto EA et al, “Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS”, *Oncotarget.* 2016 Dec 12), MOSAIC (Ronald J Hause et al, “Classification and characterization of microsatellite instability across 18 cancer types”, *Nature Medicine* 22, 1342–1350, 2016), or the Foundation Medicine NGS-MSI analysis using 114 intronic homopolymer repeat loci (10-20bp long in the human reference genome) (Michael J. Hall et al, *J Clin Oncol* 34, 2016 (suppl 4S; abstr 528).

In these computational methods MSI status can be defined by cutoff numbers for MSI-high, MSI-low or MSS based on an index score for each sample determined by using computer algorithm and validated by comparing to the other MSI detection methods.

MSI can also be detected by an immunohistochemistry (IHC) staining of four microsatellite marker proteins: MLH1, PMS2, MSH2, or MSH6. If any of these four proteins are found significantly reduced in quantity by IHC, particularly if at least one of the four proteins cannot be detected by IHC, the sample is labeled as MSI-high.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of one or more, preferably of two or more, microsatellite markers selected from BAT-25, BAT-26, NR-21, NR-24, MONO-27, D2S123, D5S346, D17S250 in the sample of the subject compared to a normal reference sample and/or microsatellite instability is characterized by the absence of one or more proteins selected from MLH1, PMS2, MSH2 and/or MSH 6.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of one or more, preferably of two or more, microsatellite markers selected from BAT-25, BAT-26, NR-21, NR-24, MONO-27, D2S123, D5S346 and/or D17S250 in the sample of the subject compared to a normal reference sample.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of one or more, preferably of two or more, microsatellite markers selected from BAT-25, BAT-26, NR-21, NR-24 and/or MONO-27 in the sample of the subject compared to a normal reference sample.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of at least one, preferably of at least two, microsatellite markers selected from BAT-25, BAT-26, NR- 21, NR-24, and MONO-27 in the sample of the subject compared to a normal reference sample.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of one or more microsatellite markers selected from BAT-25, BAT-26, D2S123, D5S346 and/or D17S250 in the sample of the subject compared to a normal reference sample.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the alteration of at least one, preferably of at least two, microsatellite markers selected from BAT-25, BAT-26, D2S123, D5S346 and/or D17S250 in the sample of the subject compared to a normal reference sample.

In another embodiment of the use/method/pharmaceutical composition/kit of the present invention the microsatellite instability is characterized by the absence of one or more proteins selected from MLH1, PMS2, MSH2 and/or MSH 6.

In another embodiment is determined by immunohistochemistry staining of MMR pathway proteins (MLH1, PMS2, MSH2, or MSH6). In this method the “MSI-high” phenotype is characterized by a significant reduction in quantity of one or more of the proteins selected from MLH1, PMS2, MSH2, and/or MSH6, particularly by a loss of expression of at least one of the proteins selected from MLH1, PMS2, MSH2 and/or MSH6. In this context the term “loss of expression” means no positive nuclear staining in a tumor cell, particularly in a tumor cell, by IHC.

The percentages in the tests and examples which follow are, unless indicated otherwise, percentages by weight; parts are parts by weight. Solvent ratios, dilution ratios and concentration data for liquid/liquid solutions are based in each case on volume.

Experimental Section

Preparation of Compound A

Compound A was prepared according to the procedure described in example 111 of International Patent Application WO2016020320.

Example 1

Treatment of different prostate cancer cell lines with Compound A

LAPC-4 human prostate cancer cells were obtained from the VTT Technical Research Center (Finland). They were plated in RPMI 1640 (RPMI = Roswell Park Memorial Institute) medium without phenol red + 10% charcoal-stripped FCS (FCS = Fetal Calf Serum) + 2 mM L-Glutamine at 4000 cells/well in a 96-well microtiter plate. After 1 day, the cells were treated with R1881 (1 nM) and Compound A (day 0). Cell number was determined by Alamar Blue staining (2 h) at day 7. Fluorescence was determined in a Victor X3 device (excitation 530 nm; emission 590 nm). The inhibition of cell growth was calculated by normalization with respect to the fluorescence reading (cell number) measured at the end of the experiment for cells treated with R1881 alone compared to the fluorescence reading (cell number) measured at the end of the experiment for DMSO-treated cells.

VCaP human prostate cancer cells were obtained from the VTT Technical Research Center (Finland). They were plated in DMEM medium (DMEM = Dulbecco's Modified Eagle Medium) with stable glutamine + 10% FCS at 16000 cells/well in a 96-well microtiter plate. Compound A was added at day 0. Cell number was determined by Alamar Blue staining (2 h) at day 0 and day 7. Fluorescence was determined in a Victor X3 device (excitation 530 nm; emission 590 nm). The inhibition of cell growth was calculated by normalization with respect to the fluorescence reading (cell number) measured at the end of the experiment for cells treated with R1881 alone compared to the fluorescence reading (cell number) measured at the start of the experiment for DMSO-treated cells.

LNCaP human prostate cancer cells were obtained from the DSMZ-German Collection of Microorganisms and Cell Cultures (Germany) (DSMZ ACC-256). They were plated in RPMI1640 medium without phenol red + 10% charcoal-stripped FCS at 600 cells/well in a 384-well white plate. R1881 (1 nM) and Compound A were added at day 0. Cell number was determined by CellTiter-Glow (Promega) at day 0 and day 6. Luminescence was determined in Victor X3. The inhibition of cell growth was calculated by normalization with respect to the fluorescence reading (cell number) measured at the

end of the experiment for cells treated with R1881 alone compared to the fluorescence reading (cell number) measured at the start of the experiment for DMSO-treated cells.

22RV1 human prostate cancer cells were obtained from the American Type Culture Collection (ATCC CRL-2505). They were plated in RPMI1640 medium supplemented with 10% FCS at 5000 cells/well in a 96-well microtiter plate. After 24 h, the cells from one microtiter plate were stained with crystal violet (==> 0 plate), whereas the cells in the test plates were exposed continuously for 4 days to test substances. Cell proliferation was determined by staining with crystal violet. The absorbance was determined photometrically at 595 nm using a Tecan Sunrise instrument. The percentage change of cell growth was calculated by normalization with respect to the absorbance reading (cell number) at the beginning of treatment of cells (0 plate) and the absorbance reading (cell number) of the untreated control group.

DU-145 human prostate cancer cells were obtained from the DSMZ-German Collection of Microorganisms and Cell Cultures (Germany) (DSMZ ACC-261). They were plated in DMEM/Ham's F12 medium at 5000 cells/well in a 96-well microtiter plate. After 24 h, the cells from one microtiter plate were stained with crystal violet (==> 0 plate), whereas the cells in the test plates were exposed continuously for 4 days to test substances. Cell proliferation was determined by staining with crystal violet. The absorbance was determined photometrically at 595 nm using a Tecan Sunrise instrument. The percentage change of cell growth was calculated by normalization with respect to the absorbance reading (cell number) at the beginning of treatment of cells (0 plate) and the absorbance reading (cell number) of the untreated control group.

PC-3 human prostate cancer cells were obtained from the DSMZ-German Collection of Microorganisms and Cell Cultures (Germany) (DSMZ ACC-465). They were plated in DMEM/Ham's F12 medium with stable Glutamine + 10% FCS at 5000 cells/well in a 96-well microtiter plate. After 24 h, the cells from one microtiter plate were stained with crystal violet (==> 0 plate), whereas the cells in the test plates were exposed continuously for 4 days to test substances. Cell proliferation was determined by staining with crystal violet. The absorbance was determined photometrically at 595 nm using a Tecan Sunrise instrument. The percentage change of cell growth was calculated by normalization with respect to the absorbance reading (cell number) at the beginning of treatment of cells (0 plate) and the absorbance reading (cell number) of the untreated control group.

Treatment of further cancer cell lines with Compound A

The cells (See Table 3: Test systems) were seeded in their appropriate medium supplemented with 10% FCS at 1,250 – 5,000 cells/well (depending on their proliferation rate) in 96-well microtiter plates. Cells were allowed to adhere for 24 h, and then the compound was added using a digital dispenser. The final

concentration of Compound A was between $1\text{E-}09$ mol/L and $3\text{E-}06$ mol/L, and the final concentration of the solvent DMSO was 0.03%. After 4 days of continuous incubation, the cells were fixed with glutaraldehyde, stained with crystal violet, and the absorbance was recorded at 595 nm. All measurements were done in quadruplicates. The values were normalized to the absorbance of solvent treated cells (=100%) and the absorbance of a reference plate which was fixed at the time point of compound application (=0%). Half-maximal growth inhibition (IC_{50}) was determined as compound concentration, which was required to achieve 50% inhibition of cellular growth using a 4-parameter fit.

Non-adherent growing GRANTA-519, Jeko-1, JVM-2, NCI-H929, Rec-1 and SU-DHL-8 cells were seeded in 150 μl of growth medium at 4000 cells/well (NCI-H929, 5000 cells/well) in 96-well microtiter plates and incubated for 24 h at 37°C . Compound A was added using a digital dispenser to the cells in the test plates and incubated continuously for 4 days at 37°C . To determine cell viability (corresponding to cell number) CTG solution (Promega Cell Titer Glo solution, # G755B and G756B) was added. After incubation for further 10 min luminescence was measured using Perkin Elmer Victor V equipment. All measurements were done in quadruplicates. The percentage change of cell viability was calculated by normalization with respect to the luminescence reading (cell number) at the beginning of treatment of cells (a reference plate was measured at the time point of compound application to the measurement plates) and the luminescence reading (cell number) of the untreated control group. Half-maximal growth inhibition (IC_{50}) was determined as compound concentration, which was required to achieve 50% inhibition of cellular growth using a 4-parameter fit.

Treatment of isogenic cancer cell lines with Compound A

The isogenic DLD-1 cell lines DLD-1 parental, DLD-1 BRCA2 (-/-) and DLD-1 ATM (-/-) (see Table 3: Test systems) were seeded in RPMI 1640 (RPMI = Roswell Park Memorial Institute) medium without phenol red + 10% charcoal-stripped FCS (FCS = Fetal Calf Serum) + 2 mM L-Glutamine + 25 mM Sodium Bicarbonate at 2,500 cells/well in a 96-well microtiter plate. Cells were allowed to adhere for 24 h, and then the compound was added using a digital dispenser. The final concentration of Compound A was between $7\text{E-}10$ mol/L and $5\text{E-}06$ mol/L, and the final concentration of the solvent DMSO was 0.03%. After 7 days of continuous incubation at 37°C cell viability was determined (corresponding to cell number) using CTG solution (Promega Cell Titer Glo solution, # G755B and G756B) was added. After incubation for further 10 min luminescence was measured using Perkin Elmer Victor V equipment. All measurements were done in quadruplicates. The percentage change of cell viability was calculated by normalization with respect to the luminescence reading (cell number) at the beginning of treatment of cells (a reference plate was measured at the time point of compound application to the measurement

plates) and the luminescence reading (cell number) of the untreated control group. Half-maximal growth inhibition (IC₅₀) was determined as compound concentration, which was required to achieve 50% inhibition of cellular growth using a 4-parameter fit.

Table 3: Test systems

Cell line	Tumor entity	Source
A2780	ovarian carcinoma	ECACC-93112519
AsPC1	pancreatic carcinoma	ATCC CRL-1682
BxPC3	pancreatic carcinoma	ATCC CRL-1687
Caco2	colorectal carcinoma	DSMZ ACC-169
GRANTA-519	mantle cell lymphoma	DSMZ ACC-342
DLD-1 (parental)	colorectal carcinoma	HD PAR-008
DLD-1 BRCA2 (-/-)	colorectal carcinoma	HD 105-007
DLD-1 ATM (-/-)	colorectal carcinoma	HD 105-061, clone 11517
HeLa	human cervical adenocarcinoma	ATCC CCL-2
HT-144	malignant melanoma	ATCC HTB-63
HT-29	colorectal carcinoma	DSMZ ACC-299
Jeko-1	mantle cell lymphoma	DSMZ ACC-553
LOVO	colorectal carcinoma	DSMZ ACC-350
MDA-MB-436	mammary carcinoma	CLS 300278
MDA-MB-468	mammary carcinoma	ATCC HTB-132
MIAPaca-2	pancreatic carcinoma	ATCC CRL-1420
NCI-H460	non-small cell lung carcinoma	ATCC HTB-177
NCI-H929	multiple myeloma	ATCC CRL-9068

Cell line	Tumor entity	Source
OVCAR-8	ovarian carcinoma	NCI-60 panel; Sample ID No. 25
REC-1	mantle cell lymphoma	ATCC CRL-3004
SK-OV-3	ovarian carcinoma	ATCC HTB-77
SU-DHL-8	germinal center B cell DLBCL	DSMZ ACC-573
JVM-2	mantle cell lymphoma	ATCC CRL-3002
TMD-8	activated B cell DLBCL	Charite, Berlin, Germany
C4-2B	prostate cancer	MD Anderson Cancer Center
HCT116	colorectal carcinoma	DSMZ ACC-581
IGR-OV-1	ovarian carcinoma	NCI-60 panel; Sample ID No. 26
NCI-H23	non-small cell lung carcinoma	ATCC CRL-5800
NCI-H1838	non-small cell lung carcinoma	ATCC CRL-5899
NCI-H1703	non-small cell lung carcinoma	ATCC CRL-5889
A549	non-small cell lung carcinoma	DSMZ ACC-107
NCI-H2030	non-small cell lung carcinoma	ATCC CRL-5914
HCC70	mammary carcinoma	ATCC CRL-2315
M059J	glioblastoma	ATCC CRL-2366
U-87MG	glioblastoma	ATCC HTB-14
SH-SY5Y	neuroblastoma	ATCC CRL-2266

ATCC = American Type Culture Collection; NCI = National Cancer Institute; CLS = Cell Line Service GmbH, Germany; DSMZ = Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany; MD Anderson Cancer Center, Houston, USA; HD = Horizon Discovery Ltd

Results:

The genetic mutations and DNA copy number alterations of the above-mentioned cancer cell lines were determined by targeted whole exome sequencing testing and/or acquired from the public databases of Cancer Cell Line Encyclopedia (CCLE, Barretina, Caponigro, Stransky et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012 28;483(7391):603-7.), Genentech Panel (Klijn C, Durinck S, Stawiski EW, Haverty PM, Jiang Z, et al, A comprehensive transcriptional portrait of human cancer cell lines. *Nat Biotechnol*. 2015 Mar;33(3):306-12.) and Sanger Cell Line Panel in COSMIC database (www.cancer.sanger.ac.uk; "COSMIC: exploring the world's knowledge of somatic mutations in human cancer", Forbes et al., *Nucleic Acids Res*. 2015, Jan; 43 (Database issue):D805-11. doi: 10.1093/nar/gku1075. Epub 2014 Oct 29).

Mutations in DNA damage (DDR) or mismatch repair (MMR) genes as well as mutations in genes inducing oncogenic replication stress or in TP53/tumor suppressor genes of these cancer cell lines (Table 3) are listed in Table 4, selected functional mutations of these cell lines are described in Table 5.

The activity of Compound A in these cell lines was tested. As shown in Table 4, Compound A inhibited the proliferation of the tumor cell lines tested. These results demonstrate that Compound A potently inhibits the proliferation of human tumor cell lines when tested as single agent and that the genetic background of the cells impacts their sensitivity to ATR inhibition.

The activity of Compound A was also tested in isogenic cell lines, in the parental DLD-1 cells and in two DLD-1 cell clones that are deficient for BRCA2 or ATM: DLD-1 BRCA2 (-/-) and DLD-1 ATM (-/-). As shown in Table 6, Compound A inhibited the proliferation of the parental DLD-1 to a lesser extent than the mutant cell lines. The strongest effect was detectable in ATM deficient DLD-1 cells. These results demonstrate that mutations in the genes BRCA2 and ATM sensitize tumor cells to treatment with Compound A.

Table 4 Inhibition of tumor cell proliferation by Compound A

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
Prostate	LNCaP	ARID1Afs, ATG5fs, ATMA1119V/K1572N, ATRXEE2264-2265E, BRCA2fs, CHEK2T430N, ERCC3A740T;R391W, ERCC5L1023I, FANCAE369D,Q652*, HDAC2A62V, MLH3I541V, MSH3PPA66-68-;fs, POLBfs, POLHD631G, PRKDCfs, RAD50fs, RAD54LL532M, RB1splice_acceptor, SLX4S605N, TDP2T308S, TP53BP1R639Q;Q111*, TRRAPR2665W;P3554L, WDR48G107*, XRCC3P87L, XRCC4L70M	APCR2714C, ATRK1379N, ERBB3K177N, MYCN45S, PTENfs, TOP2Afs, TOP2BG323*/V889A	18
Prostate	22Rv1	ATMK1101E, ARID1Afs, BARD1fs, BRCA2V1810I,fs, DCLRE1Cfs, FANCAfs, MSH3fs, NBNR43Q, PALB2V1123M, PARP4R970W, PRKDCfs, RAD18L314V, RAD50T532I,SLX4fs, TP53BP1fs, USP1fs, WRNfs, XRCC2fs	PIK3CAQ546R, ATRfs, BRAFL597R, ERBB3R683Q, TP53Q331R	36
Prostate	VCaP	MSH3PPA66-68-, MSH6fs	CCND1S219N, TMPRSS2-ERG, MYCamp, TP53R248W	51
Prostate	LapC4	ARID1A-2138-2139X, BAP1P723L, BRCA2fs, CDK12W1459X, ERCC2E313K, ERCC3R642Q/V443A/E259D, ERCC5G1080R, FANCD2N405S,P714L,P1081X,splice_donor, GEN1K42E, H2AFXN95S, LIG4T219A, MLH3K585X,	ATRR1951*stop/amp, CDC7A342V/D571G, EGFRV980D, TP53H178PX,R175H	55

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
		MSH2M300R,splice_donor, MSH3K381X, PALB2V398A, PARP1W589C, PARP3H441R, PARP4T1170I,V1065A,Q1059R,I1039T, V626D, POLHD140N, POLQS1797I,M587I, PRKDCQ4041H, RAD17R49X, RAD51C320Y, RAD54BI164T, REV3LS2862T,P2172L, TRRAPD394G,M1087I,P2026L,H3174Y, A3655V, WDR48S611P, WRNE3X,K1126X		
Prostate	DU-145	ATG5splice_donor, BRCA1E962K, BRCA2S2284L, BRIP1T132N, DYRK1AR226H, FANCBG702W, FANCIfs, GEN1Q554H, LIG4R32C, MLH1A586V,splice, MSH2L736I, MSH6S1067I, PMS2H189Y, POLA1A550S, POLLA285T, POLQV124M, PRKDCfs, RAD50N509K, RB1K715*, REV3LR2523C, RPA2E252D, TP53BP1splice_acceptor, TRRAPfs;A1389T, UIMC1A96D ,UBE2Nfs, USP1fs, XPAsfs, XRCC1splice_acceptor, XRCC2fs	KRASamp, DYRK1R226H, TMPRSS2splice_acc eptor, TP53V274F;P223L	110
Prostate	PC3	MSH3AAAAAAAAPP55-64A;PPA66-68, PRKDCfs	MYCamp, TP53fs	490
Prostate	C4-2B	MSI-H, ARID1A_c.854delG_p.G285fs*78, ATRX-E2265del, MSH2loss, TRRAP-Q1984*	PTENdel	50
Breast	HCC70	MSH3PPA66-68-, RB1DN479-480D	PTENfs, TP53R248Q	27

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
Breast	MDA-MB-436	BRCA1splice_donor, FANCIS812G, MLH3F92L, PRKDCL1824F;fs,	TMPRSS2V101F, MYCamp+, TP53fs	120
Breast	MDA-MB-468	BLMD554V, BRCA2M965I, FANCAQ869*,FANCEG245-, FBXO18G193A, SLX4E1784Q	ERBB2fs, ATRP? (2633+5A>G,Substitution - intronic) PTENsplice_donor, TP53R273H	130
Cervix	HeLa	BRIP1R855H, REV3LQ2891*, UIMC1R536W	EGFR1646L	150
Colorectal	LOVO	MSI-H, ARID1A-F2141fs*59, ATM-splice site 1236-2_1237delAGGC, CHEK2-T389fs*25, ERCC3Q711R, FANCA350W, FBXW7R505C,MSH2-G71del, MSH3L795H, NBNfs, POLHT477I, BLMfs, POLQfs, PRKDCfs, RAD50fs, XRCC2-L117fs*17	APCR1114*;R2816Q;fs, KRASG13D	71
Colorectal	HT29	FANCMfs, POLNR761*, POLQS1819*, PRKDCfs, WRNL1255V	APCE853*;fs, MYCamp, BRAFV600E,T119S, PIK3CAP449T, TP53R273H	160
Colorectal	Caco2	none	APCQ1367*, ERBB3D857N	240
Colorectal	HCT-116	ATMA1127V, ATRXTK1529-1530K, BRCA2fs, CDK12P250H, CHEK2L398P, ERCC5splice_donor, FANCAfs, FBXO18A1062V,GEN1R401Q, MLH1S252*,MSH3fs, MSH6fs, POLHA112T, POLQK2571N, PRKDCY2964C, RAD50fs, REV3Lfs, SLX4A1461fs*2, TRRAPH3023Y;T3663A, USP1R180*,	ERBB3Q261*, TOP2Afs, TOP2BR651H, KRAS G13D, PIK3CA H1047R	25

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
		WRNE480V		
Glioblastoma	U-87MG	RAD50D515G, RAD54LR691Q	ATRXN564S PTENsplice_donor,	64
Glioblastoma	M059J	FANCAR1409W, PRKDCfs, RAD54BP98L, TDP2R317*	CCNE1R95L, ERBB2W452S PTENfs, TP53E286K	80
Lung	NCI-H1838	ATMW1279*, BRCA1C328Y, CDK12R1473Q, RAD50L347P, MSH3AAAAAAAAPP55-64A;PPA66-68- , PRKDCfs	ATRP1991S, TP53R273L	24
Lung	NCI-H1703	ATMV1521L;G1998E, BRCA1G890V, FANCD2V97I, MSH3AAAAAAAAPP55- 64A;PPA66-68-, PRKDCfs	TP53splice_donor	46
Lung	NCI-H460	NBNG224A, REV3LQ1367L, MSH3AAAAAAAAPP55-64A, PRKDCfs	KRASQ61H, PIK3CAE545K, MYCamp	65
Lung	NCI-H2030	MSH3AAAAAAAAPP55-64A;PPA66-68- , PRKDCfs	KRASG12C, TP53G262V	160
Lung	NCI-H23	ATMQ1919P, BARD1A168T, BRIP1E1054A, FANCD1048Y, MSH3E251*, MDC1N1233S, NBNV153I, PMS2E491K, PRKDCfs,SLX4ED1148- 1149D	CCNE1D83N, KRASG12C/amp, TOP2BH977Y, MYCamp, NRASamp, TP53M246I	18
Lung	A549	ATRsplice_acceptor, MSH3PPA66-68-, PRKDCfs	KRASG12S, ATRsplice, CCNE1amp	29

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
Lymphoma, B cell	SU-DHL-8	ATMK1964E, FANCD2R1165Q, FEN1L190V, REV3LV1004E, PRKDCfs	MYCp72S,Q10H, BRAFT599TT, CDC7K42N, TOP2AG1197E, TP53R249G;Y234N	9
Lymphoma, B cell	TMD-8	BRIP1S59P, ERCC2H148R, MLH1V16L	APCK1170E, MYCF3L	179
Lymphoma, mantle cell	REC-1	ATMS707P/amp, PRKDCK3872R;fs, POLNN382S	KRASamp, TP53Q317*;G245D	10
Lymphoma, mantle cell	Jeko-1	ATMamp, ATRXR246C, BAP1S609G, BRCA1N742S, CHEK2V218A, DCLRE1AR1002C, ERCC2V231M, MLH3I397-, PRKDCfs, RB1R621S, SLX4G395C	ATRT1751A, TMPRSS2Y82D, TP53fs	18
Lymphoma, mantle cell	GRANT A-519	ATMR2832C	none	30
Lymphoma, mantle cell	JVM-2	FANCMQ1701*, MSH3AAAAAAAAPP55-64A, PRKDCfs	none	32
Melanoma	HT-144	ATMW2845*, PRKDCfs, XRCC3E278K	BRAFV600E	40
Neuroblastoma	SH-SY5Y	CHEK2fs, MSH3PPA66-68-, PRKDCfs	none	13
Ovarian	A2780	ARID1AQ1430*,R1721fs*4, ATMP604S, FANCMfs, PARP4G630E, POLHR356Q, PRKDCfs	ATRI123V,ERBB3V1082I, PTEN K128_R130del, TOP2BV530I, BRAFV226M, PIK3CAE365K	21
Ovarian	SK-OV-3	ARID1AQ586*, ATMsplice_acceptor,FANCMMA205V, FBXW7R505L, TDP1Y46C	APCfs, TOPBP1N295S, PIK3CAH1047R, KRASamp,	33

Indication	Cell line	DDR/MMR defects	Oncogenic replication stress or TP53/tumor suppressors	in vitro (IC50, nM)
			CDC7del, TP53fs	
Ovarian	IGROV-1	MSI-H, ARID1AD1850fs*4,G276fs*87, ATMR248Q, BRCA1K654fs*47, BRCA2P3150T, CHEK1fs, FANCA3prime_UTR, MLH1S505fs*3, MRE11AR525K, MSH3G539V;F780L;D943N,MSH6fs, PALB2T787I, POLQfs;L45I,POLNfs, PRKDCC1454Y, Y155C,RAD50fs, RAD52E130K, RB1fs, TDP1N179S, TRRAP2051F, USP1V636I, UIMC1A418T	ERBB3K742, PIK3CAR38C;*1069W, PTENfs, TOP2AH605Q, TOPBP1D395G, TP53Y126C	96
Ovarian	OVCAR 8	ATMV613L, MSH6T727S, REV3LL3040V	APCA1225S, ERBB2G776V, KRASP121H, MYCamp, TP53splice_acceptor	110
Pancreas	BxPC3	ERCC2R156Q, PRKDCfs	BRAFVTAPTP487-492A, TP53Y220C	44
Pancreas	AsPc-1	FBXW7R465C, PARP4M1110L, PRKDCfs	DYRK1A514C, KRASG12D, TP53fs	49
Pancreas	MIAPaC a2	ARID1AP1940L, MLH1T270I, PALB2S64L	KRASG12C, TP53R248W	380

Table 5: Functional mutations of genes of tested cell lines

Indication	Cell line	Functional Mutation		
		DDR/MMR deleterious	TP53/tumor suppressors	oncogenic replication stress
Prostate	LNCaP	ARID1Afs, ATGfs, ATRXEE2264-2265E, BRCA2fs, FANCAE369D,Q652*, MSH3PPA66-68-;fs, PRKDCfs, RAD50fs, RB1splice_acceptor, WDR48G107*	TP53BP1R639 Q;Q111*	ATRK1379N, ERBB3K177N, MYCN45S,TOP2Afs, TOP2BG323*/V889A
Prostate	22Rv1	ARID1Afs, BARD1fs, BRCA2V1810Lfs, DCLRE1Cfs, FANCAfs, MSH3fs, PRKDCfs, SLX4fs, USP1fs, WRNfs, XRCC2fs	TP53Q331R, TP53BP1fs	PIK3CAQ546R, ATRfs, BRAFL597R, ERBB3R683Q
Prostate	VCaP	MSH3PPA66-68-, MSH6fs	TP53R248W	CCND1S219N, TMPRSS2-ERG, MYCamp
Prostate	LapC4	ATM splice_donor, BRCA2fs, FANCD2 splice_donor, MSH2 splice_dono	TP53H178PX,R 175H	ATRR1951*stop/amp, CDC7A342V/D571G, EGFRV980D
Prostate	DU-145	ATG5splice_donor, FANCIfs, PRKDCfs, RB1K715*, TRRAPfs; UBE2Nfs, USP1fs, XPAfs, XRCC1splice_acceptor, XRCC2fs	TP53V274F;P2 23L, TP53BP1splice _acceptor	KRASamp, DYRK1R226H, TMPRSS2splice_acceptor
Prostate	PC3	MSH3AAAAAAAAPP55-64A;PPA66-68-, PRKDCfs	TP53fs	MYCamp

		Functional Mutation		
Indication	Cell line	DDR/MMR deleterious	TP53/tumor suppressors	oncogenic replication stress
Prostate	C4-2B	MSI-H, ARID1A_c.854delG_p.G285fs*78, MSH2loss, TRRAP-Q1984*, ATRX-E2265del	PTENdel	none
Breast	HCC70	MSH3PPA66-68-, RB1DN479-480D	PTENfs, TP53R248Q	none
Breast	MDA-MB-436	BRCA1splice_donor, PRKDCL1824F;fs	TP53fs	TMPRSS2V101F, MYCamp
Breast	MDA-MB-468	FANCAQ869*	PTENsplice_donor, TP53R273H	ERBB2fs, ATRP? (2633+5A>G, Substitution - intronic)
Cervix	HeLa	REV3LQ2891*	none	EGFR1646L
Colon	LOVO	MSI-H, ARID1A-F2141fs*59, ATM-splice site 1236-2_1237delAGGC, CHEK2-T389fs*25, MSH2-G71del, NBNfs, BLMfs, POLQfs, PRKDCfs, RAD50fs, XRCC2-L117fs*17	APCR1114*	KRASG13D
Colon	HT29	FANCMfs, POLNR761*, POLQS1819*, PRKDCfs	APCE853*,fs, TP53R273H	MYCamp, BRAFV600E, T119S, PIK3CAP449T
Colon	Caco2	none	APCQ1367*	ERBB3D857N

		Functional Mutation		
Indication	Cell line	DDR/MMR deleterious	TP53/tumor suppressors	oncogenic replication stress
Colon	HCT-116	ATMA1127V, ATRXTK1529-1530K, BRCA2fs, ERCC5splice_donor, FANCAfs, MLH1S252*,MSH3fs, MSH6fs, RAD50fs, REV3Lfs, SLX4A1461fs*2, USP1R180*	none	ERBB3Q261*, TOP2Afs, TOP2BR651H, KRAS G13D, PIK3CA H1047R
Glioblastoma	U-87MG	none	PTENsplice_donor	ATRXN564S
Glioblastoma	M059J	PRKDCfs, TDP2R317*	PTENfs, TP53E286K	CCNE1R95L, ERBB2W452S
Lung	NCI-H1838	ATMW1279*, MSH3AAAAAAAAPP55-64A;PPA66-68-, PRKDCfs	TP53R273L	ATRP1991S
Lung	NCI-H1703	ATMG1998E, MSH3AAAAAAAAPP55-64A;PPA66-68-, PRKDCfs	TP53splice_donor,	none
Lung	NCI-H460	MSH3AAAAAAAAPP55-64A, PRKDCfs	none	KRASQ61H, PIK3CAE545K, MYCamp
Lung	NCI-H2030	MSH3AAAAAAAAPP55-64A;PPA66-68-, PRKDCfs	TP53G262V	KRASG12C

		Functional Mutation		
Indication	Cell line	DDR/MMR deleterious	TP53/tumor suppressors	oncogenic replication stress
Lung	NCI-H23	ATMQ1919P, MSH3E251*, PRKDCfs,SLX4ED1148- 1149D	TP53M246I	CCNE1D83N, KRASG12C/amp, TOP2BH977Y, MYCamp, NRASamp
Lung	A549	MSH3PPA66-68-, PRKDCfs	none	KRASG12S, ATRsplice, CCNEamp
Lymphoma, B cell	SU-DHL-8	ATMK1964E, PRKDCfs,	TP53R249G;Y2 34N ,	MYCp72S,Q10H, BRAF599TT, CDC7K42N, TOP2AG1197E
Lymphoma, B cell	TMD-8	none	none	MYCF3L
Lymphoma, mantle cell	REC-1	PRKDCfs	TP53Q317*;G2 45D,	KRASamp
Lymphoma, mantle cell	Jeko-1	MLH3I397-, PRKDCfs, RB1R621S	TP53fs	ATRT1751A, TMPRSS2Y82D
Lymphoma, mantle cell	GRANTA- 519	ATMR2832C	none	none
Lymphoma, mantle cell	JVM-2	FANCMQ1701*, MSH3AAAAAAAAPP55- 64A, PRKDCfs	none	none
Melanoma	HT-144	ATMW2845*, PRKDCfs	none	BRAFFV600E
Neuroblastoma	SH-SY5Y	CHEK2fs, MSH3PPA66- 68-, PRKDCfs,	none	none
Ovarian	A2780	ARID1AQ1430*,R1721fs* 4, ATMP604S, FANCMfs, PRKDCfs,	PTEN K128_R130del	ATRI123V,ERBB3V1082 I, TOP2BV530I, BRAFFV226M, PIK3CAE365K

		Functional Mutation		
Indication	Cell line	DDR/MMR deleterious	TP53/tumor suppressors	oncogenic replication stress
Ovarian	SK-OV-3	ARID1AQ586*, ATMsplice_acceptor	APCfs, TP53fs	TOPBP1N295S, PIK3CAH1047R, KRASamp, CDC7del
Ovarian	IGROV-1	MSI-H, ARID1AD1850fs*4, G276fs*87, ATMR248Q, BRCA1K654fs*47, CHEK1fs, MLH1S505fs*3, MSH6fs, POLQfs; POLNfs, RAD50fs	PTENfs, TP53Y126C	ERBB3K742, PIK3CAR38C;*1069W, TOP2AH605Q, TOPBP1D395G
Ovarian	OVCAR8	ATMV613L	TP53splice_acc eptor	ERBB2G776V, KRASP121H, MYCamp
Pancreas	BxPC3	PRKDCfs	TP53Y220C	BRAFVTAPTP487-492A,
Pancreas	AsPc-1	PRKDCfs	TP53fs	DYRK1AS14C, KRASG12D
Pancreas	MIAPaCa2	none	TP53R248W	KRASG12C

Abbreviations used in Tables 4 and 5:DDR: DNA damage repair; MMR: Mismatch repair; fs: frame shift; del: deletion; *: stop codon; amp: gene amplification; MSI-H: Microsatellite Instability High
IC50: compound concentration required to achieve 50% inhibition of the maximal cell growth.

Table 6: Inhibition of isogenic tumor cell line proliferation by Compound A

Indication	Cell line	Defect	in vitro (IC ₅₀ , nM)
Colorectal	DLD-1 parental		50
Colorectal	DLD-1 BRCA2 (-/-)	BRCA2 deficiency (deleterious mutation of BRCA2)	27
Colorectal	DLD-1 ATM (-/-)	ATM deficiency (deleterious mutation of ATM)	1.5

Abbreviations used in Table 6:

IC₅₀: compound concentration required to achieve 50% inhibition of the maximal cell growth.

Example 2

In vivo xenotransplantation models

The anti-tumor activity of Compound A was examined in murine xenotransplantation models of human cancer. For this purpose, mice were implanted subcutaneously with tumor cells. At a mean tumor size of 20-30 mm² animals were randomized into treatment and control groups (n=10 animals/group) and treatment started with vehicle only or Compound A (formulation: 60% PEG400/10% Ethanol/30% Water; application route: p.o./per os, orally; dose/schedule: 50 mg/kg twice daily for 3 days on/ 4 days off). The oral application volume was 10 ml/kg. The time interval between two applications per day was 6-7h. The experiment was ended when the untreated control group had tumors of area ≤ 225 mm². The tumor size and the body weight were determined three times weekly. Changes in the body weight were a measure of treatment-related toxicity (> 10% = critical, stop of treatment until recovery, > 20% = toxic, termination). The tumor area was detected by means of an electronic caliper gauge [length (mm) x width (mm)]. In vivo anti-tumor efficacy is presented as T/C ratio (Treatment/Control) calculated with tumor areas at study end by the formula [(tumor area of treatment group at day x) - (tumor area of treatment group at day before first treatment)] / [(tumor area of control group at day x) - (tumor area of control group at day before first treatment)]. A compound having a T/C below 0.5 is defined as active (effective). Statistical analysis was assessed using SigmaStat software. A one-way analysis of variance was performed and differences to the

control were compared by a pair-wise comparison procedure (Dunn's method).

Results (Table 7):

Compound A showed potent anti-tumor efficacy in different xenograft models of human tumors upon monotherapy treatment inducing stable disease in ovarian (A2780), prostate (PC3), colorectal cancer (LOVO) and complete tumor remission in mantle cell lymphoma (REC-1) at good tolerability.

Table 7:

Anti-tumor activity of Compound A in different human cancer xenograft models in mice.

Xenograft Model	T/C^a	Max. weight loss^b (%)
REC-1	-0.13*	-10
PC3	-0.02*	-7
LOVO	0.13*	-8
A2780	0.13*	-6

* P < 0.05 (compared to vehicle treated control)

a) T/C = ratio of the tumor area of treatment versus [(tumor area of treatment group at day x) - (tumor area of treatment group at day before first treatment)] / [(tumor area of control group at day x) - (tumor area of control group at day before first treatment)].

b) Loss of body weight: Changes in body weight compared to the initial body weight at the start of treatment (> 10% = critical, stop of treatment until recovery, > 20% = toxic, termination).

The abbreviation 2QD means twice per day, po means peroral

Example 3

Treatment of isogenic DT40 chicken lymphoma cell lines with Compound A

DT40 cells from isogenic cell lines (see Table 8) were seeded in 40 µl of growth medium (RPMI 1640 medium containing stabilized glutamine (#FG1215, Merck/Biochrom), supplemented with 10% fetal calf

serum, 1% chicken serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 5E-05M β-mercaptoethanol) at 200 cells/well in 384-well white microtiter plates ((#6007680; Perkin Elmer Life Sciences) and incubated for 24 h at 37°C. Compound A was added using a digital dispenser (Tecan) to the cells in the test plates and incubated continuously for 3 days at 37°C. To determine cell viability (corresponding to cell number) 10 µl/well of CTG solution (Promega Cell Titer Glo solution, # G755B and G756B) was added. After incubation for further 10 min luminescence was measured using a PHERAstar FSX (BMG Labtech) equipment. All measurements were done in quadruplicates. The percentage change of cell viability was calculated by normalization with respect to the luminescence reading (cell number) at the beginning of treatment of cells (a reference plate was measured at the time point of compound application to the measurement plates) and the luminescence reading (cell number) of the untreated control group. Half-maximal growth inhibition (IC₅₀) was determined as compound concentration, which was required to achieve 50% inhibition of cellular growth using a 4-parameter fit.

To evaluate the relative cellular sensitivity of the isogenic DT40 cell lines towards Compound A the mean IC₅₀ of each mutant cell line was divided by the mean IC₅₀ of wild-type cells, and then the quotient was converted into logarithmic scale (base 2). Log₂ ratios of ≤ -1 or ≥ +1, corresponding to a 2-fold change in sensitivity relative to wild-type cells, were considered as particularly relevant.

Results

The activity of Compound A was tested in a panel of 46 isogenic cell lines derived from DT40 chicken lymphoma cells, which do not express TP53 (Takao et al., *Oncogene* 1999; 18: 7002-7009), covering inactivation of various genes involved in DNA damage signaling and DNA repair. Relative sensitivities against Compound A were calculated for the mutant cell lines versus the parental wild-type cell line (Table 9). The results indicate that cells deficient in the genes TP53BP1, RAD9A, RAD17, H2AFX, RAD52, BRCA1, BRCA2, UBE2N, PCNA, PARP1, TDP2, FANCD2, FANCG, POLL, POLL/POLB double mutated, REV3L, FEN1, XPA, ERCC5, or BLM are 2-fold or more than 2-fold more sensitive towards Compound A as compared to wild-type cells. Strong sensitizations (> 4-fold) were observed with RAD17, PARP1, FANCD2, UBE2N, RAD9A, REV3L, TP53BP1, ERCC5, and BLM deficient DT40 cells, whereas the strongest effects (> 8-fold) were detectable in PCNA, FEN1, H2AFX, BRCA1 deficient DT40 cells. These results demonstrate that deleterious mutations in the genes TP53BP1, RAD9A, RAD17, H2AFX, RAD52, BRCA1, BRCA2, UBE2N, PCNA, PARP1, TDP2, FANCD2, FANCG, POLL, POLL/POLB double mutated, REV3L, FEN1, XPA, ERCC5, or BLM sensitize tumor cells to treatment with Compound A.

Table 8: DT40 isogenic mutant cell lines. All cell lines were obtained from Kyoto University, Japan.

Cell line	Gene	Function of deleted (mutated) gene(s), and annotation	Ref.
KU70	XRCC6	Non-homologous end joining	1
LIGASE IV	LIG4	Non-homologous end joining	2
DNA-PKcs	PRKDC	Non-homologous end joining	3
RAP80	UIMC1	Functional interaction with Top2, Component of BRCA1-A complex, K63 poly-ubiquitin binding protein	4
53BP1	TP53BP1	Inhibition of homologous recombination (Homologous recombination)	5
ATM	ATM	Damage check point control	6
RAD9	RAD9A	Damage check point control	7
RAD17	RAD17	Damage check point control	7
H2AX	H2AFX	Homologous recombination	8
RAD52	RAD52	Homologous recombination, Rad51 like protein, Homologous recombination, single-strand DNA annealing	9
NBS1p70	NBN	Homologous recombination	10
BRCA1	BRCA1	Homologous recombination	11
BRCA2	BRCA2	Homologous recombination	12
UBC13	UBE2N	E2 ligase, post-replication repair, Homologous recombination	13
RAD18	RAD18	E3 ligase of PCNA, Post replication repair	14
PCNAK164R	PCNA	Post-replication repair	15
PARP1	PARP1	DNA damage sensing, poly(ADP-ribosyl)ation, SSB and DSB repair	16
TDP1	TDP1	Removal of Top1 cleavage complex (Top1cc)	17
TDP2	TDP2	Removal of Top2 cleavage complex (Top2cc)	18
TDP1/TDP2	TDP1/TDP2	(See above)	19
FANCC	FANCC	Interstrand crosslink repair, Homologous recombination	20
FANCD2	FANCD2	Interstrand crosslink repair, Homologous recombination	21
FANCG	FANCG	Interstrand crosslink repair, Homologous recombination	22
USP1	USP1	Interstrand crosslink repair, Homologous recombination	23
UAF1	WDR48	Interstrand crosslink repair, Homologous recombination, USP1 association factor	23
SNM1A/1B	DCLRE1A / DCLRE 1B	Interstrand crosslink repair	24
ARTEMIS	DCLRE1C	5'-3' exonuclease, non-homologous end joining	24
POLB	POLB	Base excision repair	25

POLL	POLL	DNA polymerase, Base excision repair	25
POLB/POLL	POLB/POLL	(See above)	25
POLN	POLN	Translesion synthesis DNA polymerase	26
POLQ	POLQ	Translesion synthesis DNA polymerase, Base excision repair, Helicase domain	26
POLN/POLQ	POLN-POLQ	(See above)	26
POLH	POLH	Translesion synthesis DNA polymerase	27
POLZ	REV3L	Translesion synthesis DNA polymerase	28
POLH/POLZ	POLH-REV3L	(See above)	29
FEN1	FEN1	5' flap endonuclease, base excision repair, Homologous recombination	30
XPA	XPA	Nuclear excision repair	31
XPG	ERCC5	Nuclear excision repair	32
FBH1	FBXO18	DNA helicase, Similar phenotype of BLM	33
BLM	BLM	RecQ helicase responsible for Bloom syndrome	34
WRN	WRN	RecQ helicase responsible for Werner syndrome	35
MSH3	MSH3	Mismatch repair	36
ATG5	ATG5	Autophagy related 5 homolog, autophagy, negative regulation of apoptosis	37

Table 9: Inhibition of proliferation of isogenic DT40 cells by Compound A and relative sensitivities (log₂ ratios).

Cell line	Gene	IC ₅₀ (M)	log ₂ (ratio)
Wild-type		1.3E-07	0.00
KU70	XRCC6	1.2E-07	-0.12
LIGASE IV	LIG4	1.0E-07	-0.38
DNA-PKcs	PRKDC	8.5E-08	-0.61
RAP80	UIMC1	1.0E-07	-0.38
53BP1	TP53BP1	3.7E-08	-1.81
ATM	ATM	1.1E-07	-0.24
RAD9	RAD9A	2.5E-08	-2.38
RAD17	RAD17	2.1E-08	-2.63
H2AX	H2AFX	1.2E-08	-3.44
RAD52	RAD52	5.4E-08	-1.27
NBS1p70	NBN	1.2E-07	-0.12
BRCA1	BRCA1	1.4E-08	-3.22
BRCA2	BRCA2	4.7E-08	-1.47
UBC13	UBE2N	2.4E-08	-2.44
RAD18	RAD18	6.7E-08	-0.96
PCNAK164R	PCNA	8.5E-09	-3.93
PARP1	PARP1	2.2E-08	-2.56
TDP1	TDP1	9.9E-08	-0.39
TDP2	TDP2	5.6E-08	-1.22
TDP1/TDP2	TDP1/TDP2	6.6E-08	-0.98
FANCC	FANCC	8.1E-08	-0.68
FANCD2	FANCD2	2.3E-08	-2.50
FANCG	FANCG	5.2E-08	-1.32
USP1	USP1	1.3E-07	0.00
UAF1	WDR48	7.0E-08	-0.89
SNM1A/1B	DCLRE1A / DCLRE 1B	9.9E-08	-0.39
ARTEMIS	DCLRE1C	2.0E-07	0.62
POLB	POLB	1.3E-07	0.00
POLL	POLL	4.7E-08	-1.47
POLB/POLL	POLB/POLL	5.6E-08	-1.22
POLN	POLN	1.3E-07	0.00
POLQ	POLQ	7.4E-08	-0.81
POLN/POLQ	POLN-POLQ	9.2E-08	-0.50
POLH	POLH	7.7E-08	-0.76

POLZ	REV3L	3.1E-08	-2.07
POLH/POLZ	POLH-REV3L	1.2E-07	-0.12
FEN1	FEN1	1.1E-08	-3.56
XPA	XPA	4.6E-08	-1.50
XPG	ERCC5	4.1E-08	-1.66
FBH1	FBXO18	1.5E-07	0.21
BLM	BLM	4.3E-08	-1.60
WRN	WRN	7.1E-08	-0.87
MSH3	MSH3	7.1E-08	-0.87
ATG5	ATG5	1.5E-07	0.21

References

1. Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, et al. Homologous recombination and non-homologous end-joining pathways of DNA doublestrand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *Embo J* 1998;17:5497-508.
2. Adachi N, Ishino T, Ishii Y, Takeda S, Koyama H. DNA ligase IV-deficient cells are more resistant to ionizing radiation in the absence of Ku70: Implications for DNA double-strand break repair. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98:12109-13.
3. Fukushima T, Takata M, Morrison C, Araki R, Fujimori A, Abe M, et al. Genetic analysis of the DNA-dependent protein kinase reveals an inhibitory role of Ku in late SG2 phase DNA double-strand break repair. *J Biol Chem* 2001;276:44413-8.
4. Iijima J, Zeng Z, Takeda S, Taniguchi Y. RAP80 Acts Independently of BRCA1 in Repair of Topoisomerase II Poison-Induced DNA Damage *Cancer Res.* 2010;70:8467–8474
5. Nakamura K, Sakai W, Kawamoto T, Bree RT, Lowndes NF, Takeda S, et al. Genetic dissection of vertebrate 53BP1: a major role in non-homologous end joining of DNA double strand breaks. *DNA Repair (Amst)* 2006;5:741-9.
6. Takao N, Kato H, Mori R, Morrison C, Sonada E, Sun X, et al. Disruption of ATM in p53-null cells causes multiple functional abnormalities in cellular response to ionizing radiation. *Oncogene* 1999;18:7002-9.
7. Kobayashi M, Hirano A, Kumano T, Xiang SL, Mihara K, Haseda Y, Matsui O, Shimizu H, Yamamoto K. Critical role for chicken Rad17 and Rad9 in the cellular response to DNA damage and stalled DNA replication. *Genes Cells* 2004; 9:291–303
8. Sonoda E, Zhao GY, Kohzaki M, Dhar PK, Kikuchi K, Redon C, et al. Collaborative roles of

- gammaH2AX and the Rad51 paralog Xrcc3 in homologous recombinational repair. *DNA Repair* (Amst) 2007;6:280-92.
9. Yamaguchi-Iwai Y, Sonoda E, Buerstedde JM, Bezzubova O, Morrison C, Takata M, et al. Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. *Mol Cell Biol* 1998;18:6430-5.
10. Nakahara M, Sonoda E, Nojima K, Sale JE, Takenaka K, Kikuchi K, et al. Genetic evidence for single-strand lesions initiating Nbs1-dependent homologous recombination in diversification of Ig v in chicken B lymphocytes. *PLoS genetics* 2009;5:e1000356.
11. Martin RW, Orelli BJ, Yamazoe M, Minn AJ, Takeda S, Bishop DK. RAD51 upregulation bypasses BRCA1 function and is a common feature of BRCA1-deficient breast tumors. *Cancer Res* 2007;67:9658-65.
12. Hatanaka A, Yamazoe M, Sale JE, Takata M, Yamamoto K, Kitao H, et al. Similar effects of Brca2 truncation and Rad51 paralog deficiency on immunoglobulin V gene diversification in DT40 cells support an early role for Rad51 paralogs in homologous recombination. *Molecular and cellular biology* 2005;25:1124-34.
13. Zhao GY, Sonoda E, Barber LJ, Oka H, Murakawa Y, Yamada K et al. A critical role for the ubiquitin-conjugating enzyme Ubc13 in initiating homologous recombination. *Mol Cell*. 2007;25:663-75.
14. Yamashita YM, Okada T, Matsusaka T, Sonoda E, Zhao GY, Araki K, et al. RAD18 and RAD54 cooperatively contribute to maintenance of genomic stability in vertebrate cells. *Embo J* 2002;21:5558-66.
15. Arakawa H, Moldovan GL, Saribasak H, Saribasak NN, Jentsch S, Buerstedde JM. A role for PCNA ubiquitination in immunoglobulin hypermutation. *PLoS biology* 2006;4:e366.
16. Hohegger H, Dejsuphong D, Fukushima T, Morrison C, Sonoda E, Schreiber V, et al. Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells. *Embo J*. 2006;25:1305-14.
17. Murai J, Huang SY, Das BB, Dexheimer TS, Takeda S, Pommier Y. Tyrosyl-DNA phosphodiesterase 1 (TDP1) repairs DNA damage induced by topoisomerases I and II and base alkylation in vertebrate cells. *J Biol Chem* 2012;287:12848-57.
18. Zeng Z, Cortes-Ledesma F, El Khamisy SF, Caldecott KW. TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J Biol Chem* 2011;286:403-9.
19. Zeng Z, Sharma A, Ju L, Murai J, Umans L, Vermeire L, et al. TDP2 promotes repair of topoisomerase I-mediated DNA damage in the absence of TDP1. *Nucleic Acids Res* 2012.
20. Hirano S, Yamamoto K, Ishiai M, Yamazoe M, Seki M, Matsushita N, et al. Functional

- relationships of FANCC to homologous recombination, translesion synthesis, and BLM. *Embo J* 2005;24:418-27.
21. Yamamoto K, Hirano S, Ishiai M, Morishima K, Kitao H, Namikoshi K, et al. Fanconi anemia protein FANCD2 promotes immunoglobulin gene conversion and DNA repair through a
5 mechanism related to homologous recombination. *Molecular and cellular biology* 2005;25:34-43.
22. Yamamoto K, Ishiai M, Matsushita N, Arakawa H, Lamerdin JE, Buerstedde JM, et al. Fanconi anemia FANCG protein in mitigating radiation- and enzyme-induced DNA double-strand breaks by homologous recombination in vertebrate cells. *Molecular and cellular biology* 2003;23:5421-30.
23. Murai J, Yang K, Dejsuphong D, Hirota K, Takeda S, D'Andrea AD. The USP1/UAF1 Complex
10 Promotes Double-Strand Break Repair through Homologous Recombination. *Mol Cell Biol* 2011;31:2462-9.
24. Ishiai M, Kimura M, Namikoshi K, Yamazoe M, Yamamoto K, Arakawa H, et al. DNA cross-link repair protein SNM1A interacts with PIAS1 in nuclear focus formation. *Molecular and cellular biology* 2004;24:10733-41.
- 15 25. Tano K, Nakamura J, Asagoshi K, Arakawa H, Sonoda E, Braithwaite EK, et al. Interplay between DNA polymerases beta and lambda in repair of oxidation DNA damage in chicken DT40 cells. *DNA Repair (Amst)* 2007;6:869-75.
26. Yoshimura M, Kohzaki M, Nakamura J, Asagoshi K, Sonoda E, Hou E, et al. Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. *Molecular cell*
20 2006;24:115-25.
27. Kawamoto T, Araki K, Sonoda E, Yamashita YM, Harada K, Kikuchi K, et al. Dual roles for DNA polymerase eta in homologous DNA recombination and translesion DNA synthesis. *Molecular cell* 2005;20:793-9.
28. Sonoda E, Okada T, Zhao GY, Tateishi S, Araki K, Yamaizumi M, et al. Multiple roles of Rev3, the catalytic subunit of polzeta in maintaining genome stability in vertebrates. *Embo J*
25 2003;22:3188-97.
29. Hirota K, Sonoda E, Kawamoto T, Motegi A, Masutani C, Hanaoka F, et al. Simultaneous disruption of two DNA polymerases, Poleta and Polzeta, in Avian DT40 cells unmasks the role of Poleta in cellular response to various DNA lesions. *PLoS genetics* 2010;6.
- 30 30. Matsuzaki Y, Adachi N, Koyama H. Vertebrate cells lacking FEN-1 endonuclease are viable but hypersensitive to methylating agents and H2O2. *Nucleic Acids Res* 2002;30:3273-7.
31. Okada T, Sonoda E, Yamashita YM, Koyoshi S, Tateishi S, Yamaizumi M, et al. Involvement of vertebrate polkappa in Rad18-independent postreplication repair of UV damage. *J Biol Chem* 2002;277:48690-5.
- 35 32. Kikuchi K, Taniguchi Y, Hatanaka A, Sonoda E, Hocheegger H, Adachi N, et al. Fen-1 facilitates

- homologous recombination by removing divergent sequences at DNA break ends. *Molecular and cellular biology* 2005;25:6948-55.
33. Kohzaki M, Hatanaka A, Sonoda E, Yamazoe M, Kikuchi K, Vu Trung N, et al. Cooperative roles of vertebrate Fbh1 and Blm DNA helicases in avoidance of crossovers during recombination initiated by replication fork collapse. *Mol Cell Biol* 2007;27:2812- 20.
34. Imamura O, Fujita K, Shimamoto A, Tanabe H, Takeda S, Furuichi Y, et al. Bloom helicase is involved in DNA surveillance in early S phase in vertebrate cells. *Oncogene* 2001;20:1143-51.
35. Imamura O, Fujita K, Itoh C, Takeda S, Furuichi Y, Matsumoto T. Werner and Bloom helicases are involved in DNA repair in a complementary fashion. *Oncogene* 2002;21:954-63.
36. Nojima K, Hochegger H, Saberi A, Fukushima T, Kikuchi K, Yoshimura M, Orelli BJ, Bishop DK, Hirano S, Ohzeki M, Ishiai M, Yamamoto K, Takata M, Arakawa H, Buerstedde JM, Yamazoe M, Kawamoto T, Araki K, Takahashi JA, Hashimoto N, Takeda S, Sonoda E. Multiple repair pathways mediate tolerance to chemotherapeutic cross-linking agents in vertebrate cells. *Cancer Res.* 2005; 65:11704-11711.
37. Maede Y, Shimizu H, Fukushima T, Kogame T et al. Differential and common DNA repair pathways for topoisomerase I- and II-targeted drugs in a genetic DT40 repair cell screen panel. *Mol Cancer Ther* 2013; 13; 214–20.

SEQUENCE LISTING

<110> Bayer Pharma Aktiengesellschaft

<120> Use of

2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine for treating prostate cancer

<130> BHC163077

<160> 222

<170> BiSSAP 1.3.6

<210> 1

<211> 11025

<212> DNA

<213> Homo sapiens

<220>

<223> >APC|ENSG00000134982|ENST00000457016|11025

<400> 1

acaagatggc ggagggcaag tagcaagggg gcggggtgtg gccgccggaa gcctagccgc	60
tgctcggggg ggacctgcgg gctcaggccc gggagctgcg gaccgaggtt ggctcgatgc	120
tgttcccagg tactgttggt ggctgttggt gaggaaggtg aagcactcag ttgccttctc	180
gggcctcggc gccccctatg tacgcctccc tgggctcggg tccggtcgcc cttttgccc	240
cttctgtacc accctcagtt ctcgggtcct ggagcaccgg cggcagcagg agctgcgtcc	300
ggcaggagac gaagagcccg ggcggcgctc gtacttctgg ccactgggcg agcgtctggc	360
aggtccaagg gtagccaagg atggctgcag cttcatatga tcagttgtta aagcaagttg	420
aggcactgaa gatggagaac tcaaactctc gacaagagct agaagataat tccaatcatc	480
ttacaaaact ggaaactgag gcatctaata tgaaggaagt acttaaaca ctacaaggaa	540
gtattgaaga tgaagctatg gcttcttctg gacagattga tttattagag cgtcttaaag	600
agcttaactt agatagcagt aatttccctg gagtaaaact gcggtcaaaa atgtccctcc	660
gttcttatgg aagccgggaa ggatctgtat caagccgttc tggagagtgc agtcctgttc	720
ctatgggttc atttccaaga agagggtttg taaatggaag cagagaaagt actggatatt	780
tagaagaact tgagaaagag aggtcattgc ttcttgctga tcttgacaaa gaagaaaagg	840
aaaaagactg gtattacgct caacttcaga atctcactaa aagaatagat agtcttcctt	900

taactgaaaa tttttcctta caaacagata tgaccagaag gcaattggaa tatgaagcaa	960
ggcaaatcag agttgcatg gaagaacaac taggtacctg ccaggatatg gaaaaacgag	1020
cacagcgaag aatagccaga attcagcaaa tcgaaaagga catacttcgt atacgacagc	1080
ttttacagtc ccaagcaaca gaagcagaga ggtcatctca gaacaagcat gaaaccggct	1140
cacatgatgc tgagcggcag aatgaaggtc aaggagtggg agaaatcaac atggcaactt	1200
ctggtaatgg tcagggttca actacacgaa tggaccatga aacagccagt gttttgagtt	1260
ctagtagcac acactctgca cctcgaaggc tgacaagtca tctgggaacc aaggtggaaa	1320
tggtgtattc attgttgtca atgcttggtc ctcagtataa ggatgatatg tcgcgaactt	1380
tgctagctat gtctagctcc caagacagct gtatatccat gcgacagtct ggatgtcttc	1440
ctctcctcat ccagctttta catggcaatg acaaagactc tgtattgttg ggaaattccc	1500
ggggcagtaa agaggctcgg gccagggcca gtgcagcact ccacaacatc attcactcac	1560
agcctgatga caagagaggc aggcgtgaaa tccgagtcct tcctcttttg gaacagatac	1620
gcgcttactg tgaaacctgt tgggagtggc aggaagctca tgaaccaggc atggaccagg	1680
acaaaaatcc aatgccagct cctgttgaac atcagatctg tcctgctgtg tgtgttctaa	1740
tgaaactttc atttgatgaa gagcatagac atgcaatgaa tgaactaggg ggactacagg	1800
ccattgcaga attattgcaa gtggactgtg aaatgtatgg gcttactaat gaccactaca	1860
gtattacact aagacgatat gctggaatgg ctttgacaaa cttgactttt ggagatgtag	1920
ccaacaaggc tacgctatgc tctatgaaag gctgcatgag agcacttgtg gcccaactaa	1980
aatctgaaaag tgaagactta cagcaggtta ttgcgagtgt tttgaggaat ttgtcttggc	2040
gagcagatgt aaatagtaaa aagacgttgc gagaagttgg aagtgtgaaa gcattgatgg	2100
aatgtgcttt agaagttaaa aaggaatcaa ccctcaaaag cgtattgagt gccttatgga	2160
atttgtcagc acattgcact gagaataaag ctgatatatg tgctgtagat ggtgcacttg	2220
catttttggt tggcactctt acttaccgga gccagacaaa cacttttagcc attattgaaa	2280
gtggaggtgg gatattacgg aatgtgtcca gcttgatagc tacaaatgag gaccacaggc	2340
aaatcctaag agagaacaac tgtctacaaa ctttattaca acacttaaaa tctcatagtt	2400
tgacaatagt cagtaatgca tgtggaactt tgtggaatct ctgagcaaga aatcctaaag	2460
accaggaagc attatgggac atgggggcag ttagcatgct caagaacctc attcattcaa	2520

agcacaaaat gattgctatg ggaagtgctg cagctttaag gaatctcatg gcaaataggc	2580
ctgcgaagta caaggatgcc aatattatgt ctctggctc aagcttgcca tctcttcatg	2640
ttaggaaaca aaaagcccta gaagcagaat tagatgctca gcacttatca gaaacttttg	2700
acaatataga caatttaagt cccaaggcat ctcatcgtag taagcagaga cacaagcaaa	2760
gtctctatgg tgattatggt tttgacacca atcgacatga tgataatagg tcagacaatt	2820
ttaatactgg caacatgact gtcctttcac catatttgaa tactacagtg ttaccagct	2880
cctcttcatc aagaggaagc ttagatagtt ctcgttctga aaaagataga agtttgaga	2940
gagaacgcgg aattggtcta ggcaactacc atccagcaac agaaaatcca ggaacttctt	3000
caaagcgagg tttgcagatc tccaccactg cagcccagat tgccaaagtc atggaagaag	3060
tgtagccat tcatacctct caggaagaca gaagttctgg gtctaccact gaattacatt	3120
gtgtgacaga tgagagaaat gcacttagaa gaagctctgc tgcccataca cattcaaaca	3180
cttacaattt cactaagtcg gaaaattcaa ataggacatg ttctatgcct tatgccaaat	3240
tagaatacaa gagatcttca aatgatagtt taaatagtgt cagtagtagt gatggttatg	3300
gtaaaagagg tcaaatgaaa ccctcgattg aatcctattc tgaagatgat gaaagtaagt	3360
tttgcagtta tggtaatac ccagccgacc tagcccataa aatacatagt gcaaatcata	3420
tggatgataa tgatggagaa ctagatacac caataaatta tagtcttaaa tattcagatg	3480
agcagttgaa ctctggaagg caaagtcctt cacagaatga aagatgggca agacccaaac	3540
acataataga agatgaaata aaacaaagtg agcaaagaca atcaaggaat caaagtacaa	3600
cttatcctgt ttatactgag agcactgatg ataaacacct caagttccaa ccacattttg	3660
gacagcagga atgtgtttct ccatacaggt cacggggagc caatggttca gaaacaaatc	3720
gagtgggttc taatcatgga attaatacaa atgtaagcca gtctttgtgt caagaagatg	3780
actatgaaga tgataagcct accaattata gtgaacgtta ctctgaagaa gaacagcatg	3840
aagaagaaga gagaccaaca aattatagca taaaatataa tgaagagaaa cgtcatgtgg	3900
atcagcctat tgattatagt ttaaaatatg ccacagatat tccttcatca cagaaacagt	3960
cattttcatt ctcaaagagt tcatctggac aaagcagtaa aaccgaacat atgtcttcaa	4020
gcagtgagaa tacgtccaca cttcatcta atgccaagag gcagaatcag ctccatccaa	4080
gttctgcaca gagtagaagt ggtcagcctc aaaaggctgc cacttgcaaa gtttcttcta	4140

ttaaccaaga aacaatacag acttattgtg tagaagatac tccaatatgt ttttcaagat	4200
gtagttcatt atcatctttg tcatcagctg aagatgaaat aggatgtaat cagacgacac	4260
aggaagcaga ttctgctaata accctgcaaa tagcagaaat aaaagaaaag attggaacta	4320
ggtcagctga agatcctgtg agcgaagttc cagcagtgtc acagcacctt agaaccaaat	4380
ccagcagact gcagggttct agtttatctt cagaatcagc caggcacaaa gctgttgaat	4440
tttcttcagg agcgaaatct ccctccaaaa gtgggtgtca gacacccaaa agtccacctg	4500
aacactatgt tcaggagacc ccactcatgt ttagcagatg tacttctgtc agttcacttg	4560
atagttttga gagtcgttcg attgccagct ccgttcagag tgaacctatgc agtggaatgg	4620
taagtggcat tataagcccc agtgatcttc cagatagccc tggacaaacc atgccaccaa	4680
gcagaagtaa aacacctcca ccacctctc aaacagctca aaccaagcga gaagtaccta	4740
aaaataaagc acctactgct gaaaagagag agagtggacc taagcaagct gcagtaaattg	4800
ctgcagttca gaggggtccag gttcttccag atgctgatac tttattacat tttgccacgg	4860
aaagtactcc agatggattt tcttgttcat ccagcctgag tgctctgagc ctcgatgagc	4920
catttataca gaaagatgtg gaattaagaa taatgcctcc agttcaggaa aatgacaatg	4980
ggaatgaaac agaatacagag cagcctaaag aatcaaatga aaaccaagag aaagaggcag	5040
aaaaaactat tgattctgaa aaggacctat tagatgattc agatgatgat gatattgaaa	5100
tactagaaga atgtattatt tctgccatgc caacaaagtc atcacgtaaa gcaaaaaagc	5160
cagcccagac tgcttcaaaa ttacctccac ctgtggcaag gaaaccaagt cagctgcctg	5220
tgtacaaact tctaccatca caaaacaggt tgcaaccca aaagcatgtt agttttacac	5280
cgggggatga tatgccacgg gtgtattgtg ttgaaggagc acctataaac ttttccacag	5340
ctacatctct aagtgatcta acaatcgaat cccctccaaa tgagttagct gctggagaag	5400
gagttagagg aggggcacag tcaggtgaat ttgaaaaacg agataccatt cctacagaag	5460
gcagaagtac agatgaggct caaggaggaa aaacctcatc tgtaaccata cctgaattgg	5520
atgacaataa agcagaggaa ggtgatattc ttgcagaatg cattaattct gctatgccca	5580
aagggaaaag tcacaagcct ttccgtgtga aaaagataat ggaccaggtc cagcaagcat	5640
ctgcgtcttc ttctgcaccc aacaaaaatc agttagatgg taagaaaaag aaaccaactt	5700
caccagtaaa acctatacca caaaatactg aatataggac acgtgtaaga aaaaatgcag	5760

actcaaaaaa taattttaa	gctgagagag ttttctcaga	caacaaagat tcaaagaaac	5820
agaatttgaa aaataattcc	aagggtcttca atgataagct	cccaaataat gaagatagag	5880
tcagaggaag ttttgctttt	gattcacctc atcattacac	gcctattgaa ggaactcctt	5940
actgtttttc acgaaatgat	tctttgagtt ctctagattt	tgatgatgat gatgttgacc	6000
tttccagggg aaaggctgaa	ttaagaaagg caaaagaaaa	taaggaatca gaggctaaag	6060
ttaccagcca cacagaacta	acctccaacc aacaatcagc	taataagaca caagctattg	6120
caaagcagcc aataaatcga	ggtcagccta aaccatact	tcagaaacaa tccacttttc	6180
cccagtcatc caaagacata	ccagacagag gggcagcaac	tgatgaaaag ttacagaatt	6240
ttgctattga aaatactccg	gtttgctttt ctcataattc	ctctctgagt tctctcagtg	6300
acattgacca agaaaacaac	aataaagaaa atgaacctat	caaagagact gagccccctg	6360
actcacaggg agaaccaagt	aaacctcaag catcaggcta	tgctcctaaa tcatttcatg	6420
ttgaagatac cccagtttgt	ttctcaagaa acagtttctt	cagtttcttt agtattgact	6480
ctgaagatga cctgttgag	gaatgtataa gctccgcaat	gccaaaaaag aaaaagcctt	6540
caagactcaa gggtgataat	gaaaaacata gtcccagaaa	tatgggtggc atattaggtg	6600
aagatctgac acttgatttg	aaagatatac agagaccaga	ttcagaacat ggtctatccc	6660
ctgattcaga aaattttgat	tggaaagcta ttcaggaagg	tgcaaattcc atagtaagta	6720
gtttacatca agctgctgct	gctgcatgtt tatctagaca	agcttcgtct gattcagatt	6780
ccatcctttc cctgaaatca	ggaatctctc tgggatcacc	atttcatctt acacctgac	6840
aagaagaaaa accctttaca	agtaataaag gccacgaat	tctaaaacca ggggagaaaa	6900
gtacattgga aactaaaaag	atagaatctg aaagtaaagg	aatcaaagga ggaaaaaaag	6960
tttataaaaag tttgattact	ggaaaagttc gatctaattc	agaaatttca ggccaaatga	7020
aacagcccct tcaagcaaac	atgccttcaa tctctcgagg	caggacaatg attcatattc	7080
caggagtctg aaatagctcc	tcaagtacaa gtcctgtttc	taaaaaaggc ccacccctta	7140
agactccagc ctccaaaagc	cctagtgaag gtcaaacagc	caccacttct cctagaggag	7200
ccaagccatc tgtgaaatca	gaattaagcc ctgttgccag	gcagacatcc caaataggtg	7260
ggatcaagtaa agcaccttct	agatcaggat ctagagattc	gaccccttca agacctgccc	7320
agcaaccatt aagtagacct	atacagtctc ctggccgaaa	ctcaatttcc cctggtagaa	7380

atggaataag tcctcctaac aaattatctc aacttccaag gacatcatcc cctagtactg	7440
cttcaactaa gtcctcaggt tctggaaaaa tgtcatatac atctccaggt agacagatga	7500
gccaacagaa ccttaccaaa caaacagggt tatccaagaa tgccagtagt attccaagaa	7560
gtgagtctgc ctccaaagga ctaaatacaga tgaataatgg taatggagcc aataaaaagg	7620
tagaactttc tagaatgtct tcaactaaat caagtggaag tgaatctgat agatcagaaa	7680
gacctgtatt agtacgccag tcaactttca tcaaagaagc tccaagccca accttaagaa	7740
gaaaattgga ggaatctgct tcatttgaat ctctttctcc atcatctaga ccagcttctc	7800
ccactagggtc ccaggcacia actccagttt taagtccttc ctttcctgat atgtctctat	7860
ccacacattc gtctgttcag gctgggtggat ggcgaaaact ccacacctaat ctcagtccca	7920
ctatagagta taatgatgga agaccagcaa agcgccatga tattgcacgg tctcattctg	7980
aaagtccttc tagacttcca atcaataggt caggaacctg gaaacgtgag cacagcaaac	8040
attcatcatc ctttcctcga gtaagcactt ggagaagaac tggaagttca tcttcaattc	8100
tttctgcttc atcagaatcc agtgaaaaag caaaaagtga ggatgaaaaa catgtgaact	8160
ctatttcagg aaccaaaaca agtaaagaaa accaagtatc cgcaaaaagga acatggagaa	8220
aaataaaaga aaatgaattt tctcccacia atagtacttc tcagaccgtt tcctcaggtg	8280
ctacaaatgg tgctgaatca aagactctaa tttatcaaat ggacacctgct gtttctaaaa	8340
cagaggatgt ttgggtgaga attgaggact gtcccattaa caatcctaga tctggaagat	8400
ctcccacagg taatactccc ccggtgattg acagtgtttc agaaaaggca aatccaaaca	8460
ttaaagattc aaaagataat caggcaaaac aaaatgtggg taatggcagt gttcccatgc	8520
gtaccgtggg tttggaaaat cgcctgaact cttttattca ggtggatgcc cctgacaaaa	8580
aaggaaactga gataaaacca ggacaaaata atcctgtccc tgtatcagag actaatgaaa	8640
gttctatagt ggaacgtacc ccattcagtt ctagcagctc aagcaaacac agttcaccta	8700
gtgggactgt tgctgccaga gtgactcctt ttaattacaa cccaagccct aggaaaagca	8760
gcgcagatag cacttcagct cggccatctc agatcccaac tccagtgaat aacaacacaa	8820
agaagcgaga ttccaaaact gacagcacag aatccagtgg aacccaaagt cctaagcgcc	8880
attctgggtc ttaccttggtg acatctgttt aaaagagagg aagaatgaaa ctaagaaaat	8940
tctatgttaa ttacaactgc tatatagaca ttttgtttca aatgaaactt taaaagactg	9000

aaaaattttg taaatagggtt tgattcttgt tagagggttt ttgttctgga agccatattt	9060
gatagtatac tttgtcttca ctggctttat tttgggaggc actcttgatg gttaggaaaa	9120
aaatagtaaa gccaaagtatg tttgtacagt atgtttttaca tgtattttaa gtagcatccc	9180
atcccaactt cctttaatta ttgcttgtct taaaataatg aacactacag atagaaaata	9240
tgatatattg ctgttatcaa tcatttctag attataaact gactaaactt acatcaggga	9300
aaaattggta tttatgcaaa aaaaaatgtt tttgtccttg tgagtccatc taacatcata	9360
attaatcatg tggctgtgaa attcacagta atatggttcc cgatgaacaa gtttaccag	9420
cctgctttgc tttactgcat gaatgaaact gatggttcaa tttcagaagt aatgattaac	9480
agttatgtgg tcacatgatg tgcatagaga tagctacagt gtaataattt acactatttt	9540
gtgctccaaa caaaacaaaa atctgtgtaa ctgtaaaaca ttgaatgaaa ctattttacc	9600
tgaactagat tttatctgaa agtaggtaga atttttgcta tgctgtaatt tgttgtatat	9660
tctggatatt gaggtgagat ggctgctctt ttattaatga gacatgaatt gtgtctcaac	9720
agaaactaaa tgaacatttc agaataaatt attgctgtat gtaaactgtt actgaaattg	9780
gtattttgtt gaagggtctt gtttcacatt tgtattaata attgtttaaa atgcctcttt	9840
taaaagctta tataaatttt tttcttcagc ttctatgcat taagagtaaa attcctctta	9900
ctgtaataaa aacaattgaa gaagactgtt gccacttaac cattccatgc gttggcactt	9960
atctattcct gaaatttctt ttatgtgatt agctcatctt gatttttaat atttttccac	10020
ttaaactttt ttttcttact ccactggagc tcagtaaaag taaattcatg taatagcaat	10080
gcaagcagcc tagcacagac taagcattga gcataatagg cccacataat ttcctctttc	10140
ttaatattat agaattctgt acttgaaatt gattcttaga cattgcagtc tcttcgaggc	10200
tttacagtgt aaactgtctt gccccttcat cttcttgttg caactgggtc tgacatgaac	10260
actttttatc accctgatg ttagggcaag atctcagcag tgaagtataa tcagcacttt	10320
gccatgctca gaaaattcaa atcacatgga acttttagagg tagatttaat acgattaaga	10380
tattcagaag tatatttttag aatccctgcc tgttaaggaa actttatttg tggtaggtag	10440
agttctgggg tacatgttaa gtgtcccctt atacagtgga gggaagtctt ccttcctgaa	10500
ggaaaataaa ctgacactta ttaactaaga taatttactt aatatatctt ccctgatttg	10560
ttttaaaga tcagagggtg actgatgata catgcataca tatttgttga ataaatgaaa	10620

atattattttt agtgataaga ttcatatact ctgtatttgg ggagggaana cttttttaag	10680
catggtgggg cactcagata ggagtgaata cacctacctg gtgccttgaa aatcacatca	10740
agtagttaat tatctacccc ttacctgtgt ttataacttc caggtaatga gaatgatttt	10800
ttttaagct aaaatgccag taaataaaag tgctatgact tgagctaaga tatttgactc	10860
caatgcctgt actgtgtcta ctgcaccact ttgtaaacac ttcaatttac tatctttgaa	10920
atgattgacc tttaaatttt tgccaaatgt tatctgaaat tgtctatgaa taccatctac	10980
ttctgttggt ttcccaggct tccataaaca atggagatac atgca	11025

<210> 2
 <211> 8577
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >ARID1A|ENSG00000117713|ENST00000324856|8577

<400> 2	
gaaagcggag agtcacagcg gggccaggcc ctggggagcg gagcctccac cgccccctc	60
attcccaggc aagggcttgg ggggaatgag ccgggagagc cgggtcccga gcctacagag	120
ccgggagcag ctgagccgcc ggcgcctcgg ccgccgccgc cgctctctcc tcctccgccg	180
ccgccagccc ggagcctgag ccggcggggc gggggggaga ggagcgagcg cagcgcagca	240
gcggagcccc gcgaggcccc cccgggcggg tggggagggc agcccggggg actgggcccc	300
ggggcggggg gggagggggg gagaagacga agacagggcc gggtctctcc gcggacgaga	360
cagcggggat catggccgcg caggtcggcc ccgccgccgc cagcagcctg ggcaaccgcg	420
cgccgccgcc gccctcggag ctgaagaaag ccgagcagca gcagcgggag gaggcggggg	480
gcgaggcggc ggccggcgga gcggccgagc gcgggggaaat gaaggcagcc gccgggcagg	540
aaagcgaggg ccccgccgtg gggccgccgc agccgctggg aaaggagctg caggacgggg	600
ccgagagcaa tgggggtggc ggcggcgccg gagccggcag cggcggcggg cccggcgccg	660
agccggacct gaagaactcg aacgggaacg cgggccctag gcccgcctg aacaataacc	720
tcacggagcc gcccgcgccg ggcgggtggc gcagcagcga tgggggtgggg gcgcctctc	780
actcagccgc ggccgccttg ccgccccag cctacggctt cgggcaacc tacggccgga	840
gcccgtctgc cgtcgccgc gccgcggccg ccgtcttcca ccaacaacat ggcggacaac	900

aaagccctgg cctggcagcg ctgcagagcg gcggcggcgg gggcctggag ccctacgcgg	960
ggccccagca gaactctcac gaccacggct tccccaacca ccagtacaac tcctactacc	1020
ccaaccgcag cgcctacccc ccgcccggcc cggcctacgc gctgagctcc ccgagaggtg	1080
gcactccggg ctccggcgcg gcggcggctg ccggctccaa gccgcctccc tcctccagcg	1140
cctccgcctc ctcgctgtct tcgtccttcg ctacgcagcg cttcggggcc atggggggag	1200
gcggccccctc cgcggccggc gggggaactc cccagcccac cgccaccccc accctcaacc	1260
aactgctcac gtcgcccagc tcggcccggg gctaccaggg ctaccccggg ggcgactaca	1320
gtggcgggccc ccaggacggg ggcgccggca agggcccggc ggacatggcc tcgcagtgtt	1380
ggggggctgc ggcgccggca gctgcggcgg cggccgcctc gggagggggc caacaaagga	1440
gccaccacgc gcccatgagc cccgggagca gcggcggcgg ggggcagccg ctgcccggga	1500
cccctcagcc atccagtcca atggatcaga tgggcaagat gagacctcag ccatatggcg	1560
ggactaacc atactcgag caacaggac ctccgtcagg accgcagcaa ggacatgggt	1620
accaggggca gccatacggg tcccagacc cgcagcggta cccgatgacc atgcagggcc	1680
gggcgcagag tgccatgggc ggcctctctt atacacagca gattcctcct tatggacaac	1740
aaggccccag cgggtatggt caacagggcc agactccata ttacaaccag caaagtcctc	1800
accctcagca gcagcagcca ccctactccc agcaaccacc gtcccagacc cctcatgccc	1860
aaccttcgta tcagcagcag ccacagtctc aaccaccaca gctccagtcc tctcagcctc	1920
catactccca gcagccatcc cagcctccac atcagcagtc cccggctcca taccctccc	1980
agcagtcgac gacacagcag cccccccaga gccagcccc ctactcacag ccacaggctc	2040
agtctcctta ccagcagcag caacctcagc agccagcacc ctcgacgctc tcccagcagg	2100
ctgcgtatcc tcagccccag tctcagcagt cccagcaaac tgcctattcc cagcagcgct	2160
tccctccacc gcaggagcta tctcaagatt catttgggtc tcaggcatcc tcagccccct	2220
caatgacctc cagtaaggga gggcaagaag atatgaacct gagccttcag tcaagacctc	2280
ccagcttgcc tgatctatct ggttcaatag atgacctccc catggggaca gaaggagctc	2340
tgagtcctgg agtgagcaca tcagggattt ccagcagcca aggagagcag agtaatccag	2400
ctcagctctc tttctctcct catacctccc ctcacctgcc tggcatccga ggcccttccc	2460
cgtccccctgt tggctctccc gccagtgttg ctcagtctcg ctcaggacca ctctcgcctg	2520

ctgcagtgcc aggcaaccag atgccacctc ggccacccag tggccagtcg gacagcatca	2580
tgcattccttc catgaaccaa tcaagcattg cccaagatcg aggttatatg cagaggaacc	2640
cccagatgcc ccagtacagt tccccccagc ccggctcagc cttatctccg cgtcagcctt	2700
ccggaggaca gatacacaca ggcatgggct cctaccagca gaactccatg gggagctatg	2760
gtccccaggg gggtcagtat ggcccacaag gtggctaccc caggcagcca aactataatg	2820
ccttgcccaa tgccaactac cccagtgcag gcatggctgg aggcataaac cccatgggtg	2880
ccggagggtca aatgcatgga cagcctggca tcccacctta tggcacactc cctccaggga	2940
ggatgagtca cgcctccatg ggcaaccggc cttatggccc taacatggcc aatatgccac	3000
ctcaggttgg gtcagggatg tgtccccac cagggggcat gaaccgaaa acccaagaaa	3060
ctgctgtcgc catgcatgtt gctgccaact ctatccaaaa caggccgcca ggctacccca	3120
atatgaatca agggggcatg atgggaactg gacctcctta tggacaaggg attaatagta	3180
tggctggcat gatcaaccct cagggacccc catattccat ggggtggaacc atggccaaca	3240
attctgcagg gatggcagcc agcccagaga tgatgggcct tggggatgta aagttaactc	3300
cagccaccaa aatgaacaac aaggcagatg ggacacccaa gacagaatcc aaatccaaga	3360
aatccagttc ttctactaca accaatgaga agatcaccaa gttgtatgag ctgggtggtg	3420
agcctgagag gaagatgtgg gtggaccgtt atctggcctt cactgaggag aaggccatgg	3480
gcatgacaaa tctgcctgct gtgggtagga aacctctgga cctctatcgc ctctatgtgt	3540
ctgtgaagga gattggtgga ttgactcagg tcaacaagaa caaaaaatgg cgggaacttg	3600
caaccaacct caatgtgggc acatcaagca gtgctgccag ctccttgaaa aagcagtata	3660
tccagtgtct ctatgccttt gaatgcaaga ttgaacgggg agaagaccct ccccagaca	3720
tctttgcagc tgctgattcc aagaagtccc agcccaagat ccagcctccc tctcctgcgg	3780
gatcaggatc tatgcagggg ccccagactc cccagtcaac cagcagttcc atggcagaag	3840
gaggagactt aaagccacca actccagcat ccacaccaca cagtcagatc cccccattgc	3900
caggcatgag caggagcaat tcagttggga tccaggatgc ctttaatgat ggaagtgact	3960
ccacattcca gaagcggaat tccatgactc caaacctgg gtatcagccc agtatgaata	4020
cctctgacat gatggggcgc atgtcctatg agccaaataa ggatccttat ggcagcatga	4080
ggaaagctcc agggagtgat cccttcatgt cctcagggca gggccccaac ggcgggatgg	4140

gtgaccccta cagtcgtgct gccggccctg ggctagga aa tgtggcgatg ggaccacgac	4200
agcactatcc ctatggaggt ccttatgaca gagtgaggac ggagcctgga atagggcctg	4260
agggaacat gagcactggg gccccacagc cgaatctcat gccttccaac ccagactcgg	4320
ggatgtattc tcctagccgc tccccccgc agcagcagca gcagcagcag caacgacatg	4380
attcctatgg caatcagttc tccaccaag gcaccccttc tggcagcccc ttccccagcc	4440
agcagactac aatgtatcaa cagcaacagc agaattacaa gcggccaatg gatggcacat	4500
atggccctcc tgccaagcgg cacgaagggg agatgtacag cgtgccatac agcactgggc	4560
aggggcagcc tcagcagcag cagttgcccc cagcccagcc ccagcctgcc agccagcaac	4620
aagctgcca gccttcccct cagcaagatg tataacaacca gtatggcaat gcctatcctg	4680
ccactgccac agctgctact gagcgccgac cagcaggcgg cccccagaac caatttccat	4740
tccagtttgg ccgagaccgt gtctctgcac cccctggcac caatgccag caaaacatgc	4800
caccacaaat gatgggcggc cccatacagg catcagctga ggttgctcag caaggcacca	4860
tgtggcaggg gcgtaatgac atgacctata attatgcaa caggcagagc acgggctctg	4920
ccccccaggg cccgcctat catggcgtga accgaacaga tgaaatgctg cacacagatc	4980
agagggccaa ccacgaaggc tcgtggcctt cccatggcac acgccagccc ccatatggtc	5040
cctctgcccc tgtgcccccc atgacaaggc cccctccatc taactaccag cccccacaa	5100
gcatgcagaa tcacattcct caggtatcca gccctgctcc cctgccccgg ccaatggaga	5160
accgcacctc tcctagcaag tctccattcc tgcactctgg gatgaaaatg cagaaggcag	5220
gtcccccagt acctgcctcg cacatagcac ctgcccctgt gcagcccccc atgattcggc	5280
gggatatcac cttcccacct ggctctgttg aagccacaca gcctgtgttg aagcagagga	5340
ggcggctcac aatgaaagac attggaaccc cggaggcatg gcgggtaatg atgtccctca	5400
agtctggtct cctggcagag agcacatggg cattagatac catcaacatc ctgctgtatg	5460
atgacaacag catcatgacc ttcaacctca gtcagctccc agggttgcta gagctccttg	5520
tagaatattt ccgacgatgc ctgattgaga tctttggcat tttaaaggag tatgaggtgg	5580
gtgaccagg acagagaacg ctactggatc ctgggaggtt cagcaaggtg tctagtccag	5640
ctcccatgga ggggtggggaa gaagaagaag aacttctagg tcctaaacta gaagaggaag	5700
aagaagagga agtagttgaa aatgatgagg agatagcctt ttcaggcaag gacaagccag	5760

cttcagagaa tagtgaggag aagctgatca gtaagtttga caagcttcca gtaaagatcg	5820
tacagaagaa tgatccattt gtggtggact gctcagataa gcttgggcgt gtgcaggagt	5880
ttgacagtgg cctgctgcac tggcggattg gtggggggga caccactgag catatccaga	5940
cccacttcga gagcaagaca gagctgctgc cttcccggcc tcacgcaccc tgcccaccag	6000
cccctcggaa gcatgtgaca acagcagagg gtacaccagg gacaacagac caggaggggc	6060
ccccaccta tggacctcca gaaaaacgga tcacagccac tatggatgac atgttgtcta	6120
ctcggcttag caccttgacc gaggatggag ctaagagttc agaggccatc aaggagagca	6180
gcaagtttcc atttggcatt agcccagcac agagccaccg gaacatcaag atcctagagg	6240
acgaaccca cagtaaggat gagacccac tgtgtaccct tctggactgg caggattctc	6300
ttgccaagcg ctgctgtgt gtgtccaata ccattcgaag cctgtcattt gtgccaggca	6360
atgactttga gatgtccaaa caccagggc tgctgctcat cctgggcaag ctgatcctgc	6420
tgcaccacaa gcaccagaa cggaagcagg caccactaac ttatgaaaag gaggaggaac	6480
aggaccaagg ggtgagctgc acaaagtgg agtgggtgtg ggactgcttg gagatgctcc	6540
gggaaaacac cttggttaca ctgcacca tctcggggca gttggacct tctccatacc	6600
ccgagagcat ttgcctgcct gtcctggacg gactcctaca ctgggcagtt tgcccttcag	6660
ctgaagcca ggacccctt tccaccctgg gcccgaatgc cgtcctttcc ccgcagagac	6720
tggctcttga aaccctcagc aaactcagca tccaggacaa caatgtggac ctgattctgg	6780
ccacaccccc cttcagccgc ctggagaagt tgtatagcac tatggtgcgc ttcctcagtg	6840
accgaaagaa cccggtgtgc cgggagatgg ctgtggtact gctggccaac ctggctcagg	6900
gggacagcct ggcagctcgt gccattgcag tgcagaaggg cagtatcggc aacctcctgg	6960
gcttcctaga ggacagcctt gccgccacac agttccagca gagccaggcc agcctcctcc	7020
acatgcagaa cccaccctt gagccaacta gtgtggacat gatgcggcgg gctgcccgcg	7080
cgctgcttgc cttggccaag gtggacgaga accactcaga gtttactctg tacgaatcac	7140
ggctgttga catctcggt taccgttga tgaactcatt ggtttcacaa gtcatttgtg	7200
atgtactgtt tttgattggc cagtcatgac agccgtggga cacctcccc cccgtgtgt	7260
gtgtgctgt gtggagaact tagaaactga ctgttgccct ttatttatgc aaaaccacct	7320
cagaatccag tttaccctgt gctgtccagc ttctcccttg ggaaaaagtc tctcctgttt	7380

ctctctctc cttccacctc ccctccctcc atcacctcac gcctttctgt tccttgcct	7440
caccttactc ccctcaggac cctacccac cctctttgaa aagacaaagc tctgcctaca	7500
tagaagactt tttttatttt aaccaaagtt actgttgttt acagtgagtt tggggaaaaa	7560
aaataaaaata aaaatggctt tcccagtcct tgcacacac ggatgccaca tttcataact	7620
gtttttaatg gtaaaaaaaaa aaaaaaaaaa taaaaaaaaa aattctgaag gacaaaaaag	7680
gtgactgctg aactgtgtgt ggtttattgt tgtacattca caatcttgca ggagccaaga	7740
agttcgagtt tgtgaacaga ccctgttcac tggagaggcc tgtgcagtag agtgtagacc	7800
ctttcatgta ctgtactgta cacctgatac tgtaaacata ctgtaataat aatgtctcac	7860
atggaaacag aaaacgctgg gtcagcagca agctgtagtt tttaaaaatg tttttagtta	7920
aacgttgagg agaaaaaaaa aaaaggcttt tcccccaaag tatcatgtgt gaacctacaa	7980
caccctgacc tctttctctc ctcttgatt gtatgaataa ccctgagatc acctcttaga	8040
actggtttta acctttagct gcagcggcta cgctgccacg tgtgtatata tatgacgttg	8100
tacattgcac atacccttgg atccccacag tttggctctc ctcccagcta cccctttata	8160
gtatgacgag ttaacaagtt ggtgacctgc acaaagcgag acacagctat ttaatctctt	8220
gccagatatc gcccctcttg gtgcgatgct gtacaggtct ctgtaaaaag tccttgctgt	8280
ctcagcagcc aatcaactta tagtttattt ttttctgggt ttttgttttg ttttgttttc	8340
tttctaatac aggtgtgaaa aagttctagg ttcagttgaa gttctgatga agaaacacaa	8400
ttgagatttt ttcagtata aaatctgcat atttgtattt caacaatgta gctaaaactt	8460
gatgtaaatt cctccttttt ttcctttttt ggcttaatga atatcattta ttcagtatga	8520
aatctttata ctatatgttc cacgtgttaa gaataaatgt acattaaatc ttggtaa	8577

<210> 3
 <211> 3216
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >ATG5|ENSG0000057663|ENST00000369076|3216

ctggacttgt ggtgcgctgc cagggtccg cagcgttgcc ggttgatttc gctggatacc	60
--	----

agagggcgga agtgcagcag ggttcagctc cgacctccgc gccggtgctt tttgcggctg	120
cgcgggcctt ctggagtcct gctaccgcgt ccccgagga cagtgtgtca ggcgggcagc	180
ttgccccgcc gccccaccgg agcgcggaat ctgggcgtcc ccaccagtgc ggggagccgg	240
aaggaggagc catagcttgg agtaggtttg gctttggttg aaataagaat ttagcctgta	300
tgtactgctt taactcctgg aagaatgaca gatgacaaag atgtgcttcg agatgtgtgg	360
tttggacgaa ttccaacttg tttcacgcta tatcaggatg agataactga aagggaagca	420
gaaccatact atttgctttt gccaaagagta agttatttga cgttggtaac tgacaaagtg	480
aaaaagcact ttcagaaggt tatgagacaa gaagacatta gtgagatatg gtttgaatat	540
gaaggcacac cactgaaatg gcattatcca attggtttgc tatttgatct tcttgcata	600
agttcagctc ttccttggaa catcacagta cattttaaga gttttccaga aaaagacctt	660
ctgcactgtc catctaagga tgcaattgaa gctcatttta tgtcatgtat gaaagaagct	720
gatgctttaa aacataaaaag tcaagtaatc aatgaaatgc agaaaaaaga tcacaagcaa	780
ctctggatgg gattgcaaaa tgacagattt gaccagtttt gggccatcaa tcggaaactc	840
atggaatatc ctgcagaaga aaatggattt cgttatatcc cctttagaat atatcagaca	900
acgactgaaa gacctttcat tcagaagctg tttcgtcctg tggctgcaga tggacagttg	960
cacacactag gagatctcct caaagaagtt tgtccttctg ctattgatcc tgaagatggg	1020
gaaaaaaaga atcaagtgat gattcatgga attgagccaa tgttggaaac acctctgcag	1080
tggctgagtg aacatctgag ctaccggat aattttcttc atattagtat catcccacag	1140
ccaacagatt gaaggatcaa ctatttgcct gaacagaatc atccttaa at gggatttatc	1200
agagcatgtc acccttttgc ttcaatcagg tttggtggag gcaacctgac cagaaacact	1260
tcgctgctgc aagccagaca ggaaaaagat tccatgtcag ataaggcaac tgggctggtc	1320
ttactttgca tcacctctgc tttcctccac tgccatcatt aaacctcagc tgtgacatga	1380
aagacttacc ggaccactga aggtcttctg taaaatataa tgaagctgaa acctttggcc	1440
taagaagaaa atggaagtat gtgccactcg atttgtattt ctgattaaca aataaacagg	1500
ggtatttcct aaggtgacca tggttgaact ttagctcatg aaagtggaaa cattggttta	1560
attttcaaga gaattaagaa agtaaaagag aaattctgtt atcaataact tgcaagtaat	1620
tttttgtaaa agattgaatt acagtaaacc catctttcct taacgaaaat ttcctatgtt	1680

tacagtctgt ctattggtat gcaatcttgt aactttgata atgaacagtg agagatTTTT	1740
aaataaagcc tctaaatatg ttttgtcatt taataacata cagttttgtc acttttcaag	1800
tactttctga ctcacataca gtagatcact ttttactctg tgttaccatt ttgactggtc	1860
gtcattggca tgggggtggat atagggcata ggattacttg tctcagaagc tgtcatagaa	1920
tttcttgctg ccaattaaaa aacctgtgtt ctttacacac tacacgtata aatattgtaa	1980
ctgttcatct ttgttgtttt atcactgtaa gcctgtcaaa tcatagtatc ctaagcatct	2040
gtaaatgcta attttgcatt tttggaaaaa cccattcctt ccaagctagt gtttttcatt	2100
ggctccaggt ctaatttttc actgtgggtcc ctggcagcca gtcttttgaa gtttaaagat	2160
tacctgtctc ttgactgcag taccttttct ttaattttta ccaaaaatat ccagaggtta	2220
ctggagtctt tattcaatat aaggaaagt ttgctgcactt tattaccaag cctctgggat	2280
tttaccagtc aaacatattt gtgcattaca tttcatttct tgtgagctag ctggctgtcc	2340
atattgaatg ttgaccatt tgagtacgct aaaaggctta cagtatcaga cacgatcatg	2400
gttttagatc ccataataaa aatgaatgtt tttcttataa aaaattatac aaatgctgaa	2460
gtgagattct actattgttc attgcttcct tttctttttc cttttgcat tttcactgat	2520
taatagcaca tttcttcaca aaattagata aagttgggtca aagaccagat attctggaat	2580
ggaaattgta aagcttaatc aaaaagaata gccagtacag catacaatct cagaaactta	2640
gaagcaagta gaaaataatt ggttgatgta aacgaaagt ccattttagt aaaggcagga	2700
aaaaaatagc aatatttgag ttatgtaagg ataaaaaat cactgacttg tatttttgca	2760
caagaggctg gtctgaatat gattgttcac attaagagt tttattcgtc ggttcatttt	2820
ggggattttc ccccttgatg ttttgacaga ttgaagtgag ctttagtgag caaaaggatc	2880
agaatgcagg gaacactaag ctgtgatgaa gaaagtgtgg taaaaagcca gagtagtttt	2940
atacagacaa aaccagtgtc aggcctttgc agtaggcttg agtgaacttc tgatctagat	3000
ttgaaagtaa attttatgaa gacattgccc atttttactt cctcattcat tattgtacca	3060
gcatcatagc tttattactc taatcccagg taagtcaagc ctacaatgcc ctagaggaag	3120
agtaaaacca gaaattcatg ctggcttaaa taatctattt ttgtttcttt tcatttgaat	3180
atttaaattt tatggtttat taaaaatta aataaa	3216

<211> 13147
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >ATM|ENSG00000149311|ENST00000278616|13147

<400> 4
 ccggagcccc agccgaaggg cgagccgcaa acgctaagtc gctggccatt ggtggacatg 60
 gcgcaggcgc gtttgctccg acgggccgaa tgttttgggg cagtgttttg agcgcggaga 120
 ccgcgtgata ctggatgcgc atgggcatac cgtgctctgc ggctgcttgg cgttgcttct 180
 tcctccagaa gtgggcgctg ggagtcacg cagggtttga accggaagcg ggagtaggta 240
 gctgcgtggc taacggagaa aagaagccgt ggccgcggga ggaggcgaga ggagtcggga 300
 tctgcgctgc agccaccgcc gcggttgata ctactttgac cttccgagtg cagtgcagtg 360
 gatgtgtgtt ctgaaattgt gaacatgag tctagtactt aatgatctgc ttatctgctg 420
 ccgtcaacta gaacatgata gagctacaga acgaaagaaa gaagttgaga aatttaagcg 480
 cctgattcga gatcctgaaa caattaaaca tctagatcgg cattcagatt ccaaacaagg 540
 aaaatatattg aattgggatg ctgttttttag atttttacag aaatatattc agaaagaaac 600
 agaatgtctg agaatagcaa aaccaaattgt atcagcctca acacaagcct ccaggcagaa 660
 aaagatgcag gaaatcagta gtttggtcaa atacttcac aaatgtgcaa acagaagagc 720
 acctaggcta aaatgtcaag aactcttaaa ttatatcatg gatacagtga aagattcatc 780
 taatggtgct atttacggag ctgattgtag caacatacta ctcaaagaca ttctttctgt 840
 gagaaaatac tgggtgtgaaa tatctcagca acagtggta gaattgttct ctgtgtactt 900
 caggctctat ctgaaacctt cacaagatgt tcatagagtt ttagtggcta gaataattca 960
 tgctgttacc aaaggatgct gttctcagac tgacggatta aattccaaat ttttggactt 1020
 tttttccaag gctattcagt gtgcgagaca agaaaagagc tcttcaggtc taaatcatat 1080
 cttagcagct ctactatct tcctcaagac tttggctgtc aactttcgaa ttcgagtgtg 1140
 tgaattagga gatgaaattc ttcccacttt gctttatatt tggactcaac ataggcttaa 1200
 tgattcttta aaagaagtca ttattgaatt atttcaactg caaatattata tccatcatcc 1260
 gaaaggagcc aaaacccaag aaaaaggtgc ttatgaatca acaaaatgga gaagtatttt 1320
 atacaactta tatgatctgc tagtgaatga gataagtcac ataggaagta gaggaaagta 1380

ttcttcagga tttcgtataa ttgccgtcaa agaaaatttg attgaattga tggcagatat	1440
ctgtcaccag gtttttaatg aagataccag atccttggag atttctcaat cttacactac	1500
tacacaaaga gaatctagt attacagtgt cccttgcaaa aggaagaaaa tagaactagg	1560
ctgggaagta ataaaagatc accttcagaa gtcacagaat gattttgatc ttgtgccttg	1620
gctacagatt gcaaccaat taatatcaaa gtatcctgca agtttaccta actgtgagct	1680
gtctccatta ctgatgatac tatctcagct tctaccccaa cagcgacatg gggaacgtac	1740
accatatgtg ttacgatgcc ttacggaagt tgcatttgtt caagacaaga ggtcaaacct	1800
agaaagctca caaaagtcag atttattaaa actctggaat aaaatttggt gtattacctt	1860
tcgtgggtata agttctgagc aaatacaagc tgaaaacttt ggcttacttg gagccataat	1920
tcagggtagt ttagttgagg ttgacagaga attctggaag ttatttactg ggtcagcctg	1980
cagaccttca tgcctgcag tatgctgttt gactttggca ctgaccacca gtatagttcc	2040
aggaacggta aaaatgggaa tagagcaaaa tatgtgtgaa gtaaatagaa gcttttcttt	2100
aaaggaatca ataatgaaat ggctcttatt ctatcagtta gagggtgact tagaaaatag	2160
cacagaagtg cctccaattc ttcacagtaa ttttcctcat cttgtactgg agaaaattct	2220
tgtgagtctc actatgaaaa actgtaaagc tgcaatgaat tttttccaaa gcgtgccaga	2280
atgtgaacac caccaaaaag ataaagaaga actttcattc tcagaagtag aagaactatt	2340
tcttcagaca acttttgaca agatggactt ttttaaccatt gtgagagaat gtggtataga	2400
aaagcaccag tccagtattg gcttctctgt ccaccagaat ctcaaggaat cactggatcg	2460
ctgtcttctg ggattatcag aacagcttct gaataattac tcacttgaga ttacaaattc	2520
agaaactctt gtccggtgtt cacgtctttt ggtgggtgtc cttggctgct actgttacat	2580
gggtgtaata gctgaagagg aagcatataa gtcagaatta ttccagaaag ccaagtctct	2640
aatgcaatgt gcaggagaaa gtatcactct gtttaaaaat aagacaaatg aggaattcag	2700
aattggttcc ttgagaaata tgatgcagct atgtacacgt tgcttgagca actgtaccaa	2760
gaagagtcca aataagattg catctggctt tttcctgcga ttgttaacat caaagctaata	2820
gaatgacatt gcagatatct gtaaaagttt agcatccttc atcaaaaagc catttgaccg	2880
tggagaagta gaatcaatgg aagatgatac taatggaaat ctaatggagg tggaggatca	2940
gtcatccatg aatctattta acgattacc tgaatagtagt gttagtgatg caaacgaacc	3000

tggagagagc	caaagtacca	taggtgccat	taatccttta	gctgaagaat	atctgtcaaa	3060
gcaagatcta	cttttcttag	acatgctcaa	gttcttgtgt	ttgtgtgtaa	ctactgctca	3120
gaccaatact	gtgtccttta	gggcagctga	tattcggagg	aaattgttaa	tgттаattga	3180
ttctagcacg	ctagaaccta	ccaaatccct	ccacctgcat	atgtatctaa	tgctttttaa	3240
ggagcttcct	ggagaagagt	accccttgcc	aatggaagat	gttcttgaac	ttctgaaacc	3300
actatccaat	gtgtgttctt	tgtatcgctg	tgaccaagat	gtttgtaaaa	ctatttttaa	3360
ccatgtcctt	catgtagtga	aaaacctagg	tcaaagcaat	atggactctg	agaacacaag	3420
ggatgctcaa	ggacagtttc	ttacagtaat	tggagcattt	tggcatctaa	caaaggagag	3480
gaaatatata	ttctctgtaa	gaatggccct	agtaaattgc	cttaaaactt	tgcttgaggc	3540
tgatccttat	tcaaaatggg	ccattcttaa	tgtaatggga	aaagactttc	ctgtaaataa	3600
agtatttaca	caatttcttg	ctgacaatca	tcaccaagtt	cgcatgttgg	ctgcagagtc	3660
aatcaataga	ttgttccagg	acacgaaggg	agattcttcc	aggttactga	aagcacttcc	3720
tttgaagctt	cagcaaacag	cttttgaaaa	tgcatacttg	aaagctcagg	aaggaatgag	3780
agaaatgtcc	catagtgctg	agaaccctga	aactttggat	gaaatttata	atagaaaatc	3840
tgttttactg	acgttgatag	ctgtgggttt	atcctgtagc	cctatctgcg	aaaaacaggc	3900
tttgtttgcc	ctgtgtaaat	ctgtgaaaga	gaatggatta	gaacctcacc	ttgtgaaaaa	3960
ggtttttagag	aaagtttctg	aaacttttgg	atatagacgt	ttagaagact	ttatggcatc	4020
tcatttagat	tatctgggtt	tggaatggct	aaatcttcaa	gatactgaat	acaacttatc	4080
ttcttttcct	tttattttat	taaactacac	aaatattgag	gattttctata	gatcttgtta	4140
taaggttttg	attccacatc	tggtgattag	aagtcatttt	gatgagggtga	agtcatttgc	4200
taatcagatt	caagaggact	ggaaaagtct	tctaacagac	tgctttccaa	agattcttgt	4260
aaatattctt	ccttattttg	cctatgaggg	taccagagac	agtgggatgg	cacagcaaag	4320
agagactgct	accaaggctt	atgatatgct	taaaagtga	aacttatttg	gaaaacagat	4380
tgatcactta	ttcattagta	atttaccaga	gattgtgggtg	gagttattga	tgacgttaca	4440
tgagccagca	aattctagt	ccagtcagag	cactgacctc	tgtgactttt	caggggattt	4500
ggatcctgct	cctaattccac	ctcattttcc	atcgcatgtg	attaaagcaa	catttgccta	4560
tatcagcaat	tgtcataaaa	ccaagttaaa	aagcatttta	gaaattcttt	ccaaaagccc	4620

tgattcctat cagaaaattc ttcttgccat atgtgagcaa gcagctgaaa caaataatgt	4680
ttataagaag cacagaattc ttaaaatata tcacctgttt gttagtttat tactgaaaga	4740
tataaaaagt ggcttaggag gagcttgggc ctttgttctt cgagacgtta ttataacttt	4800
gattcactat atcaaccaa ggccttcttg tatcatggat gtgtcattac gtagcttctc	4860
cctttgttgt gacttattaa gtcaggtttg ccagacagcc gtgacttact gtaaggatgc	4920
tctagaaaac catcttcattg ttattgttgg tacacttata ccccttgtgt atgagcaggt	4980
ggaggttcag aaacaggtat tggacttggt gaaatactta gtgatagata acaaggataa	5040
tgaaaacctc tatatcacga ttaagctttt agatcctttt cctgaccatg ttgtttttaa	5100
ggatttgcgt attactcagc aaaaaatcaa atacagtaga ggaccctttt cactcttgga	5160
ggaaattaac cattttctct cagtaagtgt ttatgatgca cttccattga caagacttga	5220
aggactaaag gatcttcgaa gacaactgga actacataaa gatcagatgg tggacattat	5280
gagagcttct caggataatc cgcaagatgg gattatggtg aaactagttg tcaatttggt	5340
gcagttatcc aagatggcaa taaaccacac tggtgaaaaa gaagttctag aggctgttgg	5400
aagctgcttg ggagaagtgg gtcctataga tttctctacc atagctatac aacatagtaa	5460
agatgcatct tataccaagg cccttaagtt atttgaagat aaagaacttc agtggacctt	5520
cataatgctg acctacctga ataacacact ggtagaagat tgtgtcaaag ttcgatcagc	5580
agctgttacc tgtttgaaaa acatttttagc cacaaagact ggacatagtt tctgggagat	5640
ttataagatg acaacagatc caatgctggc ctatctacag ccttttagaa catcaagaaa	5700
aaagttttta gaagtacca gatttgacaa agaaaacct tttgaaggcc tggatgatat	5760
aatctgtgg attcctctaa gtgaaaatca tgacatttgg ataaagacac tgacttgtgc	5820
ttttttggac agtggaggca caaatgtga aattcttcaa ttattaaagc caatgtgtga	5880
agtgaaaact gacttttgtc agactgtact tccatacttg attcatgata ttttactcca	5940
agatacaaat gaatcatgga gaaatctgct ttctacacat gttcagggat ttttcaccag	6000
ctgtcttcga cacttctcgc aaacgagccg atccacaacc cctgcaaact tggattcaga	6060
gtcagagcac tttttccgat gctgtttgga taaaaaatca caaagaacaa tgcttgctgt	6120
tgtggactac atgagaagac aaaagagacc ttcttcagga acaattttta atgatgcttt	6180
ctggctggat ttaaattatc tagaagttgc caaggtagct cagtcttgtg ctgctcactt	6240

tacagcttta ctctatgcag aaatctatgc agataagaaa agtatggatg atcaagagaa	6300
aagaagtctt gcatttgaag aaggaagcca gaatacaact atttctagct tgagtgaana	6360
aagtaaagaa gaaactggaa taagtttaca ggatcttctc ttagaaatct acagaagtat	6420
aggggagcca gatagtttgt atggctgtgg tggagggag atgttacaac ccattactag	6480
actacgaaca tatgaacacg aagcaatgtg gggcaaagcc ctagtaacat atgacctcga	6540
aacagcaatc ccctcatcaa cacgccaggc aggaatcatt caggccttgc agaatttggg	6600
actctgccat attctttccg tctattttaa aggattggat tatgaaaata aagactggtg	6660
tcctgaacta gaagaacttc attaccaagc agcatggagg aatatgcagt gggaccattg	6720
cacttccgtc agcaaagaag tagaaggaa cagttaccat gaatcattgt acaatgctct	6780
acaatctcta agagacagag aattctctac attttatgaa agtctcaa atgccagagt	6840
aaaagaagtg gaagagatgt gtaagcgcag ccttgagtct gtgtattcgc tctatcccac	6900
acttagcagg ttgcaggcca ttggagagct ggaaagcatt ggggagcttt tctcaagatc	6960
agtcacacat agacaactct ctgaagtata tattaagtgg cagaaacact cccagcttct	7020
caaggacagt gattttagtt ttcaggagcc tatcatggct ctacgcacag tcattttgga	7080
gatcctgatg gaaaaggaaa tggacaactc acaaagagaa tgtattaagg acattctcac	7140
caaacacctt gtagaactct ctatactggc cagaactttc aagaacactc agctccctga	7200
aagggcaata tttcaaatta aacagtacaa ttcagttagc tgtggagtct ctgagtggca	7260
gctggaagaa gcacaagtat tctgggcaaa aaaggagcag agtcttgccc tgagtattct	7320
caagcaaatg atcaagaagt tggatgccag ctgtgcagcg aacaatcca gcctaaaact	7380
tacatacaca gaatgtctga gggtttgtgg caactggtta gcagaaacgt gcttagaaaa	7440
tcctgcggtc atcatgcaga cctatctaga aaaggcagta gaagttgctg gaaattatga	7500
tggagaaagt agtgatgagc taagaaatgg aaaaatgaag gcatttctct cattagcccg	7560
gttttcagat actcaatacc aaagaattga aaactacatg aaatcatcgg aatttgaaaa	7620
caagcaagct ctctgaaaa gagccaaaga ggaagtaggt ctccttaggg aacataaaat	7680
tcagacaaac agatacacag taaaggttca gcgagagctg gagttggatg aattagccct	7740
gcgtgcactg aaagaggatc gtaaagcgtt cttatgtaaa gcagttgaaa attatatcaa	7800
ctgcttatta agtggagaag aacatgatat gtgggtattc cgactttgtt ccctctggct	7860

tgaaaattct ggagtttctg aagtcaatgg catgatgaag agagacggaa tgaagattcc	7920
aacatataaa tttttgcctc ttatgtacca attggctgct agaatgggga ccaagatgat	7980
gggaggccta ggatttcattg aagtcctcaa taatctaatt tctagaattt caatggatca	8040
cccccatcac actttgttta ttatactggc cttagcaaat gcaaacagag atgaatttct	8100
gactaaacca gaggtagcca gaagaagcag aataactaaa aatgtgccta aacaaagctc	8160
tcagcttgat gaggatcgaa cagaggctgc aaatagaata atatgtacta tcagaagtag	8220
gagacctcag atggtcagaa gtgttgaggc actttgtgat gcttatatta tattagcaaa	8280
cttagatgcc actcagtgga agactcagag aaaaggcata aatattccag cagaccagcc	8340
aattactaaa cttagaatt tagaagatgt tgttgtccct actatggaaa ttaaggtgga	8400
ccacacagga gaatatggaa atctggtgac tatacagtca tttaaagcag aatttcgctt	8460
agcaggaggt gtaaatttac caaaaataat agattgtgta gggtccgatg gcaaggagag	8520
gagacagctt gttaagggcc gtgatgacct gagacaagat gctgtcatgc aacaggtctt	8580
ccagatgtgt aatacattac tgcagagaaa cacggaaact aggaagagga aattaactat	8640
ctgtacttat aagggtggtc ccctctctca gcgaagtggg gttcttgaat ggtgcacagg	8700
aactgtcccc attggtgaat ttcttgtaa caatgaagat ggtgctcata aaagatacag	8760
gccaaatgat ttcagtgcct ttcagtgcc aaagaaaatg atggaggtgc aaaaaagtc	8820
ttttgaagag aaatatgaag tcttcatgga tgtttgcaa aattttcaac cagttttccg	8880
ttacttctgc atggaaaaat tcttgatcc agctatttgg tttgagaagc gattggctta	8940
tacgcgcagt gtagctactt cttctattgt tggttacata cttggacttg gtgatagaca	9000
tgtacagaat atcttgataa atgagcagtc agcagaactt gtacatatag atctaggtgt	9060
tgcttttgaa cagggcacaa tccttcctac tcctgagaca gttcctttta gactcaccag	9120
agatattgtg gatggcatgg gcattacggg tgttgaaggt gtcttcagaa gatgctgtga	9180
gaaaaccatg gaagtgatga gaaactctca ggaaactctg ttaaccattg tagaggtcct	9240
tctatatgat ccactctttg actggaccat gaatcctttg aaagctttgt atttacagca	9300
gaggccggaa gatgaaactg agcttcaccc tactctgaat gcagatgacc aagaatgcaa	9360
acgaaatctc agtgatattg accagagttt caacaaagta gctgaacgtg tcttaatgag	9420
actacaagag aaactgaaag gagtggaaga aggcactgtg ctcatgtttg gtggacaagt	9480

gaatttgctc atacagcagg ccatagaccc caaaaatctc agccgacttt tcccaggatg	9540
gaaagcttgg gtgtgatctt cagtatatga attacccttt cattcagcct ttagaaatta	9600
tatttttagcc tttattttta acctgccaac atactttaag tagggattaa tattttaagt	9660
aactattgtg ggtttttttg aatgttggtt ttaataactg atttaatcac cactcaaaaa	9720
tgttttgatg gtcttaagga acatctctgc tttcactctt tagaaataat ggtcattcgg	9780
gctgggcgca gcggctcacg cctgtaatcc cagcactttg ggaggccgag gtgagcggat	9840
cacaaggta ggagttcgag accagcctgg ccaagagacc agcctggcca gtatggtgaa	9900
accctgtctc tactaaaaat acaaaaatta gccgagcatg gtggcgggca cctgtaatcc	9960
cagctactcg agaggctgag gcaggagaat ctcttgaacc tgggaggtga aggttgctgt	10020
gggccaaaaat catgccattg cactccagcc tgggtgacaa gagcgaaact ccatctcaaa	10080
aaaaaaaaaa aaaaaacaga aacgtatttg gatctttcct agtaagatca ctcagtgtta	10140
ctaaataatg aagttgttat ggagaacaaa tttcaaagac acagttagtg tagttactat	10200
ttttttaagt gtgtattaaa acttctcatt ctattctctt tatcttttaa gcccttctgt	10260
actgtccatg tatgttatct ttctgtgata acttcataga ttgccttcta gttcatgaat	10320
tctcttgta gatgtatata atctctttta ccctatccat tgggcttctt ctttcagaaa	10380
ttgtttttca tttctaatta tgcatactt ttcagatctc tgtttcttga tgtcattttt	10440
aatgtttttt taatgttttt tatgtcacta attattttta atgtctgtac ttgatagaca	10500
ctgtaatagt tctattaaat ttagttcctg ctgtttatat ctgttgattt ttgtatttga	10560
taggctgttc atccagtttt gtctttttga aaagtgaagt tattttcagc aaggctttat	10620
ctatgggaat cttgagtgtc tgtttatgtc atattcccag ggctgttgct gcacacaagc	10680
ccattcttat ttttaatttct tggcttttagg gtttccatac ctgaagtgtg gcataaatac	10740
tgataggaga tttcccaggc caaggcaaac acatttcctc ctcatctcct tgtgctagt	10800
ggcagaatat ttgattgatg cttttttcac tgagagtata agcttccatg tgtccacct	10860
ttatggcagg ggtggaagga ggtacattta attcccactg cctgcctttg gcaagccctg	10920
ggttctttgc tccccatata gatgtctaag ctaaaagccg tgggttaatg agactggcaa	10980
attgttccag gacagctaca gcatcagctc acatattcac ctctctggtt tttcattccc	11040
ctcatttttt tctgagacag agtcttgctc tgtcacccag gctggagtgc agtggcatga	11100

tctcagctca ctgaaacctc tgcctcctgg gttcaagcaa ttctcctgcc tcagcctccc	11160
gagtagctgg gactacaggc gtgtgccaac acgcccggct aatTTTTTgt atTTTTatta	11220
gagacggagt ttcaccgtgt tagccaggat ggtctcgatc gcttgacctc gtgatccacc	11280
ctcctcggcc tcccaaagtg ctgggattac aggtgtgagc caccgcgcc ggcctcattc	11340
ccctcatttt tgaccgtaag gatttcccct ttcttgtaag ttctgctatg tatttaaaag	11400
aatgttttct acattttatc cagcatttct ctgtgttctg ttggaaggga agggcttagg	11460
tatctagttt gatacatagg tagaagtgga acatttctct gtccccagc tgtcatcata	11520
taagataaac atcagataaa aagccacctg aaagtaaaac tactgactcg tgtattagtg	11580
agtataatct cttctccatc cttaggaaaa tgttcatccc agctgcggag attaacaat	11640
gggtgattga gtttctcct cgtatttgga ccttgaaggc tatataaatt ttttcttat	11700
gaagagttgg catttctttt tattgccaat ggcaggcact cattcatatt tgatctcctc	11760
accttcccct cccctaaaac caatctccag aactttttgg actataaatt tcttggttg	11820
acttctggag aactgttcag aatattactt tgcatttcaa attacaaact taccttggtg	11880
tatctttttc ttacaagctg cctaaatgaa tatttggtat atattggtag ttttattact	11940
atagtaaadc aaggaaatgc agtaaactta aaatgtcttt aagaaagccc tgaaatcttc	12000
atgggtgaaa ttagaaatta tcaactagat aatagtatag ataaatgaat ttgtagctaa	12060
ttcttgctag ttgttgcatc cagagagctt tgaataacat cattaatcta ctctttagcc	12120
ttgcatggta tgctatgagg ctctgttct gttcaagtat tctaataat ggctttgaaa	12180
agtttatcaa atttacatac agatcacaag cctaggagaa ataactaatt cacagatgac	12240
agaattaaga ttataaaaga tttttttttt gtaatttttag tagagacagg gttgccattg	12300
tattccagcc ttggcgacag agcaagactc tgcctcaaaa aaaaaaaaaa aaaggttttg	12360
gcaagctgga actctttctg caaatgacta agatagaaaa ctgccaagga caaatgagga	12420
gtagttagat ttgaaaata ttaatcatag aatagttgtt gtatgctaag tcaactgacc	12480
atattatgta cagcatttct gatctttact ttgcaagatt agtgatacta tcccaatata	12540
ctgctggaga aatcagaatt tggagaaata agttgtccaa ggcaagaaga tagtaaat	12600
taagtacaag tgtaatatgg acagtatcta acttgaaaag atttcaggcg aaaagaatct	12660
ggggtttgcc agtcagttgc tcaaaaggc aatgaaaacc aaatagtga gctatcagag	12720

aagctaataa attatagact gcttgaacag ttgtgtccag attaaggag ataatagctt	12780
tcccacccta ctttgtgcag gtcatacctc cccaaagtgt ttacctaadc agtaggttca	12840
caaactcttg gtcattatag tatatgccta aaatgtatgc acttaggaat gctaaaaatt	12900
taaatatggt ctaaagcaaa taaaagcaaa gaggaaaaac tttggacagc gtaaagacta	12960
gaatagtctt ttaaaaagaa agccagtata ttggtttgaa atatagagat gtgtcccaat	13020
ttcaagtatt ttaattgcac cttaatgaaa ttatctatct tctatagatt ttagtactat	13080
tgaatgtatt actttactgt tacctgaatt tattataaag tgtttttgaa taaataattc	13140
taaaagc	13147

<210> 5
 <211> 8249
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >ATR|ENSG00000175054|ENST00000350721|8249

<400> 5	
ttccgggagg agttttggcc tccacacggc tccgtcgggc gccgcgctct tccggcagcg	60
gtagcttttg agacgccggg aacccgcgtt ggcgtgggtg actagtgcct cgcagcctca	120
gcatggggga acatggcctg gagctggctt ccatgatccc cgccctgcgg gagctgggca	180
gtgccacacc agaggaatat aatacagttg tacagaagcc aagacaaatt ctgtgtcaat	240
tcattgaccg gatacttaca gatgtaaatg ttgttgctgt agaacttgta aagaaaactg	300
actctcagcc aacctccgtg atgttgcttg atttcatcca gcatatcatg aaatcctccc	360
cacttatgtt tgtaaagtgt agtggaagcc atgaggccaa aggcagttgt attgaattca	420
gtaattggat cataacgaga cttctgcgga ttgcagcaac tccctcctgt catttgttac	480
acaagaaaat ctgtgaagtc atctgttcat tattatttct ttttaaaagc aagagtcctg	540
ctatTTTTGG ggtactcaca aaagaattat tacaactttt tgaagacttg gtttacctcc	600
atagaagaaa tgtgatgggt catgctgtgg aatggccagt ggtcatgagc cgatttttaa	660
gtcaattaga tgaacacatg ggatatttac aatcagctcc tttgcagttg atgagtatgc	720
aaaatttaga atttattgaa gtcactttat taatggttct tactcgtatt attgcaattg	780

tgtttttttag aaggcaagaa ctcttacttt ggcagatagg ttgtgttctg ctagagtatg	840
gtagtccaaa aattaaatcc ctagcaatta gctttttaac agaacttttt cagcttggag	900
gactaccagc acaaccagct agcacttttt tcagctcatt tttggaatta ttaaaacacc	960
ttgtagaaat ggatactgac caattgaaac tctatgaaga gccattatca aagctgataa	1020
agacactatt tccctttgaa gcagaagctt atagaaatat tgaacctgtc tatttaaata	1080
tgctgctgga aaaactctgt gtcattgtttg aagacggtgt gctcatgcgg cttaaagtctg	1140
atttgctaaa agcagctttg tgccatttac tgcagtattt ccttaaattt gtgccagctg	1200
ggatatgaatc tgctttacaa gtcaggaagg tctatgtgag aaatatttgt aaagctcttt	1260
tggatgtgct tggaattgag gtagatgcag agtacttggt gggcccactt tatgcagctt	1320
tgaaaatgga aagtatggaa atcattgagg agattcaatg ccaaactcaa caggaaaacc	1380
tcagcagtaa tagtgatgga atatcaccca aaaggcgtcg tctcagctcg tctctaaacc	1440
cttctaaaag agcaccaaaa cagactgagg aaattaaaca tgtggacatg aaccaaaga	1500
gcatattatg gtagtgcactg aaacagaaag ctgaatccct tcagatttcc cttgaataca	1560
gtggcctaaa gaatcctggt attgagatgt tagaaggaat tgctgttgtc ttacaactga	1620
ctgctctgtg tactgttcat tgttctcatc aaaacatgaa ctgccgtact ttcaaggact	1680
gtcaacataa atccaagaag aaaccttctg tagtgataac ttggatgtca ttggattttt	1740
acacaaaagt gcttaagagc tgtagaagtt tgttagaatc tgttcagaaa ctggacctgg	1800
aggcaaccat tgataagggtg gtgaaaattt atgatgcttt gatttatatg caagtaaaca	1860
gttcatttga agatcatatc ctggaagatt tatgtggtat gctctcactt ccatggattt	1920
attcccattc tgatgatggc tgtttaaagt tgaccacatt tgccgctaatt cttctaakat	1980
taagctgtag gatttcagat agctattcac cacaggcaca atcacgatgt gtgtttcttc	2040
tgactctggt tccaagaaga atattccttg agtggagaac agcagtttac aactgggccc	2100
tgacagagctc ccatgaagta atccgggcta gttgtgttag tggatttttt atcttattgc	2160
agcagcagaa ttcttgtaac agagttccca agattcttat agataaagtc aaagatgatt	2220
ctgacattgt caagaaagaa tttgcttcta tacttgggtca acttgtctgt actcttcacg	2280
gcatgtttta tctgacaagt tctttaacag aacctttctc tgaacacgga catgtggacc	2340
tcttctgtag gaacttgaaa gccacttctc aacatgaatg ttcattcttct caactaaaag	2400

cttctgtctg caagccattc cttttcctac tgaaaaaaaa aatacctagt ccagtaaaac	2460
ttgctttcat agataatcta catcatcttt gtaagcatct tgattttaga gaagatgaaa	2520
cagatgtaaa agcagttctt ggaactttat taaatttaat ggaagatcca gacaaagatg	2580
ttagagtggc ttttagtgga aatatcaagc acatattgga atccttggac tctgaagatg	2640
gatttataaa ggagcttttt gtcttaagaa tgaaggaagc atatacacat gcccaaatat	2700
caagaaataa tgagctgaag gataccttga ttcttacaac aggggatatt ggaagggccg	2760
caaaaggaga tttggtacca tttgcactct tacacttatt gcattgtttg ttatccaagt	2820
cagcatctgt ctctggagca gcatacacag aaattagagc tctggttgca gctaaaagtg	2880
ttaaactgca aagttttttc agccagtata agaaacccat ctgtcagttt ttggtagaat	2940
cccttcactc tagtcagatg acagcacttc cgaatactcc atgccagaat gctgacgtgc	3000
gaaaacaaga tgtggctcac cagagagaaa tggctttaaa tacgttgtct gaaattgcca	3060
acgttttcga ctttcctgat cttaatcggt ttcttactag gacattacaa gttctactac	3120
ctgatcttgc tgccaaagca agccctgcag cttctgctct cattcgaact ttaggaaaac	3180
aattaaatgt caatcgtaga gagattttta taaacaactt caaatatatt ttttctcatt	3240
tggtctgttc ttgttccaaa gatgaattag aacgtgccct tcattatctg aagaatgaaa	3300
cagaaattga actggggagc ctgttgagac aagatttcca aggattgcat aatgaattat	3360
tgctgcgtat tggagaacac tatcaacagg tttttaatgg tttgtcaata cttgcctcat	3420
ttgcatccag tgatgatcca tatcagggcc cgagagatat catatcacct gaactgatgg	3480
ctgattattt acaacccaaa ttgttgggca ttttggcttt ttttaacatg cagttactga	3540
gctctagtgt tggcattgaa gataagaaaa tggccttgaa cagtttgatg tctttgatga	3600
agttaatggg acccaaacat gtcagttctg tgagggtgaa gatgatgacc aactgagaa	3660
ctggccttcg attcaaggat gattttcctg aattgtgttg cagagcttgg gactgctttg	3720
ttcgctgcct ggatcatgct tgtctgggct cccttctcag tcatgtaata gtagctttgt	3780
tacctttat acacatccag cctaaagaaa ctgcagctat cttccactac ctcataattg	3840
aaaacaggga tgctgtgcaa gattttcttc atgaaatata ttttttacct gatcatccag	3900
aattaaaaaa gataaaagcc gttctccagg aatacagaaa ggagacctct gagagcactg	3960
atcttcagac aactcttcag ctctctatga aggccattca acatgaaaat gtcgatgttc	4020

gtattcatgc tcttacaagc ttgaaggaaa ccttgtataa aaatcaggaa aaactgataa	4080
agtatgcaac agacagtgaac acagtagaac ctattatctc acagttgggtg acagtgcctt	4140
tgaaagggtg ccaagatgca aactctcaag ctcggttgct ctgtggggaa tgtttagggg	4200
aattgggggc gatagatcca ggtcgattag atttctcaac aactgaaact caaggaaaag	4260
attttacatt tgtgactgga gtagaagatt caagctttgc ctatggatta ttgatggagc	4320
taacaagagc ttaccttgcg tatgctgata atagccgagc tcaagattca gctgcctatg	4380
ccattcagga gttgctttct atttatgact gtagagagat ggagaccaac ggcccaggtc	4440
accaattgtg gaggagattt cctgagcatg ttcgggaaat actagaacct catctaaata	4500
ccagatacaa gaggttctcag aagtcaaccg attgggtctgg agtaaagaag ccaatttact	4560
taagtaaatt gggtagtaac tttgcagaat ggtcagcatc ttgggcaggt tatcttatta	4620
caaaggttcg acatgatctt gccagtaaaa ttttcacctg ctgtagcatt atgatgaagc	4680
atgatttcaa agtgaccatc tatcttcttc cacatattct ggtgtatgtc ttactgggtt	4740
gtaatcaaga agatcagcag gaggtttatg cagaaattat ggagtttcta aagcatgacg	4800
atcagcatac cataaatacc caagacattg catctgatct gtgtcaactc agtacacaga	4860
ctgtgttctc catgcttgac catctcacac agtgggcaag gcacaaattt caggcactga	4920
aagctgagaa atgtccacac agcaaatcaa acagaaataa ggtagactca atggtatcta	4980
ctgtggatta tgaagactat cagagtgtaa cccgttttct agacctcata cccagagata	5040
ctctggcagt agcttccttt cgctccaaag catacacacg agctgtaatg cactttgaat	5100
catttattac agaaaagaag caaaatattc aggaacatct tggattttta cagaaattgt	5160
atgctgctat gcatgaacct gatggagtgg ccggagtcag tgcaattaga aaggcagaac	5220
catctctaaa agaacagatc cttgaacatg aaagccttgg cttgctgagg gatgccactg	5280
cttgttatga cagggctatt cagctagaac cagaccagat cattcattat catggtgtag	5340
taaagtccat gttaggtctt ggtcagctgt ctactgttat cactcaggtg aatggagtgc	5400
atgctaacag gtccgagtgg acagatgaat taaacacgta cagagtggaa gcagcttgga	5460
aattgtcaca gtgggatttg gtggaaaact atttggcagc agatggaaaa tctacaacat	5520
ggagtgtcag actgggacag ctattattat cagccaaaaa aagagatatc acagcttttt	5580
atgactcact gaaactagtg agagcagaac aaattgtacc tctttcagct gcaagctttg	5640

aaagaggctc ctaccaacga ggatatgaat atattgtgag attgcacatg ttatgtgagt	5700
tggagcatag catcaaacca cttttccagc attctccagg tgacagttct caagaagatt	5760
ctctaaactg ggtagctcga ctagaaatga cccagaattc ctacagagcc aaggagccta	5820
tcctggctct ccggagggct ttactaagcc tcaacaaaag accagattac aatgaaatgg	5880
ttggagaatg ctggctgcag agtgccaggg tagctagaaa ggctggtcac caccagacag	5940
cctacaatgc tctccttaat gcaggggaat cacgactcgc tgaactgtac gtggaaaggg	6000
caaagtggtc ctgggtccaag ggtgatgttc accaggcact aattgttctt caaaaaggtg	6060
ttgaattatg ttttcctgaa aatgaaaccc cacctgaggg taagaacatg ttaatccatg	6120
gtcgagctat gctactagt ggccgattta tggaagaaac agctaacttt gaaagcaatg	6180
caattatgaa aaaatataag gatgtgaccg cgtgcctgcc agaatgggag gatgggcatt	6240
tttaccttgc caagtactat gacaaattga tgcccatggt cacagacaac aaaatggaaa	6300
agcaaggtga tctcatccgg tatatagttc ttcattttgg cagatctcta caatatggaa	6360
atcagttcat atatcagtca atgccacgaa tgtaaactct atggcttgat tatggtacaa	6420
aggcatatga atgggaaaaa gctggccgct ccgatcgtgt acaaagagg aatgatttgg	6480
gtaaaataaa caaggttatc acagagcata caaactattt agctccatat caatTTTTga	6540
ctgctttttc acaattgatc tctcgaattt gtcatttcta cgatgaagtt tttgttgtct	6600
tgatggaaat aatagccaaa gtatttctag cctatcctca acaagcaatg tggatgatga	6660
cagctgtgtc aaagtcactt tatcccatgc gtgtgaacag atgcaaggaa atcctcaata	6720
aagctattca tatgaaaaaa tccttagaga agtttgttgg agatgcaact cgcctaacag	6780
ataagcttct agaattgtgc aataaaccgg ttgatggaag tagttccaca ttaagcatga	6840
gcactcattt taaaatgctt aaaaagctgg tagaagaagc aacatttagt gaaatcctca	6900
ttcctctaca atcagtcag atacctacac ttccatcaat tctgggtacc catgctaacc	6960
atgctagcca tgaaccattt cctggacatt gggcctatat tgcagggttt gatgatatgg	7020
tggaatttct tgcttctctt cagaaaccaa agaagatttc tttaaaaggc tcagatggaa	7080
agttctacat catgatgtgt aagccaaaag atgacctgag aaaggattgt agactaatgg	7140
aattcaattc cttgattaat aagtgtctaa gaaaagatgc agagtctcgt agaagagaac	7200
ttcatattcg aacatatgca gttattccac taaatgatga atgtgggatt attgaatggg	7260

tgaacaacac tgctggtttg agacctattc tgaccaaact atataaagaa aagggagtgt	7320
atatgacagg aaaagaactt cgccagtgtg tgctaccaa gtcagcagct ttatctgaaa	7380
aactcaaagt attccgagaa tttctcctgc ccaggcatcc tcctatTTTT catgagtgg	7440
ttctgagaac attccctgat cctacatcat ggtacagtag tagatcagct tactgccgtt	7500
ccactgcagt aatgtcaatg gttgggtata ttctggggct tggagaccgt catggtgaaa	7560
atatttcttt tgattctttg actggtgaat gcgtacatgt agatttcaat tgtcttttca	7620
ataagggaga aacctttgaa gttccagaaa ttgtgccatt tcgcctgact cataatatgg	7680
ttaatggaat gggtcctatg ggaacagagg gtctttttcg aagagcatgt gaagttacaa	7740
tgaggctgat gcgtgatcag cgagagcctt taatgagtgt cttaaagact tttctacatg	7800
atcctcttgt ggaatggagt aaaccagtga aagggcattc caaagcgcca ctgaatgaaa	7860
ctggagaagt tgtcaatgaa aaggccaaga cccatgttct tgacattgag cagcgactac	7920
aaggtgtaat caagactcga aatagagtga caggactgcc gttatctatt gaaggacatg	7980
tgcattacct tatacaggaa gctactgatg aaaacttact atgccagatg tatcttgggt	8040
ggactccata tatgtgaaat gaaattatgt aaaagaatat gttaataatc taaaagtaat	8100
gcatttggtg tgaatctgtg gttgtatctg ttcaattcta aagtacaaca taaatttacg	8160
ttctcagcaa ctgttatttc tctctgatca ttaattatat gtaaaataat atacattcag	8220
ttattaagaa ataaactgct ttcttaata	8249

<210> 6
 <211> 11167
 <212> DNA
 <213> Homo sapiens

<220>
 <223> ATRX|ENSG00000085224|ENST00000373344|11167

<400> 6	
ctcggcccaa caaatggcg gcggcagcgg tgtcgctttg tttccgcggc tcctgcggcg	60
gtggcagtgg tagcggcctt tgagctgtgg ggaggttcca gcagcagcta cagtgcagac	120
taagactcca gtgcatttct atcgtaaccg ggcgcggggg agcgcagatc ggcgccagc	180
aatcacagaa gccgacaagg cgttcaagcg aaaacatgac cgctgagccc atgagtgaaa	240
gcaagttgaa tacattggtg cagaagcttc atgacttcct tgcacactca tcagaagaat	300

ctgaagaaac aagttctcct ccacgacttg caatgaatca aaacacagat aaaatcagtg	360
gttctggaag taactctgat atgatggaaa acagcaagga agagggaact agctcttcag	420
aaaaatccaa gtcttcagga tcgtcacgat caaagaggaa accttcaatt gtaacaaagt	480
atgtagaatc agatgatgaa aaaccttttg atgatgaaac tgtaaatgaa gatgcgtcta	540
atgaaaattc agaaaatgat attactatgc agagcttgcc aaaaggtaca gtgattgtac	600
agccagagcc agtgctgaat gaagacaaaag atgattttta agggcctgaa tttagaagca	660
gaagtaaaat gaaaactgaa aatctcaaaa aacgcggaga agatgggctt catgggattg	720
tgagctgcac tgcttgtgga caacaggtca atcattttca aaaagattcc atttatagac	780
acccttcatt gcaagttctt atttgtaaga attgctttta gtattacatg agtgatgata	840
ttagccgtga ctcatgatga atggatgaac aatgtaggtg gtgtgcggaa ggtggaaact	900
tgatttgttg tgacttttgc cataatgctt tctgcaagaa atgcattcta cgcaaccttg	960
gtcgaaagga gttgtccaca ataatggatg aaaacaacca atggtattgc tacatttgtc	1020
accagagcc tttgttggac ttggtcactg catgtaacag cgtatttgag aatttagaac	1080
agttgttgca gcaaaataag aagaagataa aagttgacag tgaaaagagt aataaagtat	1140
atgaacatac atccagattt tctccaaaga agactagttc aaattgtaat ggagaagaaa	1200
agaaattaga tgattcctgt tctggctctg taacctactc ttattccgca ctaattgtgc	1260
ccaaagagat gattaagaag gcaaaaaaac tgattgagac cacagccaac atgaactcca	1320
gttatgttaa atttttaag caggcaacag ataattcaga aatcagttct gctacaaaat	1380
tacgtcagct taaggctttt aagtctgtgt tggctgatat taagaaggct catcttgcac	1440
tggaagaaga cttaaattcc gagtttcgag cgatggatgc tgtaaacaaa gagaaaaata	1500
ccaaagagca taaagtcata gatgctaagt ttgaaacaaa agcacgaaaa ggagaaaaac	1560
cttgtgcttt ggaaaagaag gatattttcaa agtcagaagc taaactttca agaaaacagg	1620
tagatagtga gcacatgcat cagaatgttc caacagagga acaaagaaca aataaaagta	1680
ccggtggtga acataagaaa tctgatagaa aagaagaacc tcaatatgaa cctgccaaca	1740
cttctgaaga tttagacatg gatattgtgt ctgttccttc ctgattcca gaagacattt	1800
ttgagaatct tgagactgct atggaagttc agagttcagt tgatcatcaa ggggatggca	1860
gcagtggaaac tgaacaagaa gtggagagtt catctgtaaa attaaatatt tcttcaaaag	1920

acaacagagg aggtattaaa tcaaaaacta cagctaaagt aacaaaagaa ttatatgtta	1980
aactcactcc tgtttccctt tctaattccc caattaaagg tgctgattgt caggaagttc	2040
cacaagataa agatggctat aaaagttgtg gtctgaaccc caagttagag aaatgtggac	2100
ttggacagga aaacagtgat aatgagcatt tggttgaaaa tgaagtttca ttacttttag	2160
aggaatctga tcttcgaaga tccccacgtg taaagactac acccttgagg cgaccgacag	2220
aaactaacc tgaacatct aattcagatg aagaatgtaa tgaaacagtt aaggagaaac	2280
aaaaactatc agttccagtg agaaaaaagg ataagcgtaa ttcttctgac agtgctatag	2340
ataatcctaa gcctaataaa ttgccaaaat ctaagcaatc agagactgtg gatcaaaatt	2400
cagattctga tgaaatgcta gcaatcctca aagaggtgag caggatgagt cacagttctt	2460
cttcagatac tgatattaat gaaattcata caaaccataa gactttgtat gatttaaaga	2520
ctcaggcggg gaaagatgat aaaggaaaaa ggaaacgaaa aagttctaca tctggctcag	2580
attttgatac taaaaagggc aaatcagcta agagctctat aatttctaaa aagaaacgac	2640
aaaccagtc tgagtcttct aattatgact cagaattaga aaaagagata aagagcatga	2700
gtaaaattgg tgctgccaga accaccaaaa aaagaattcc aaatacaaaa gattttgact	2760
cttctgaaga tgagaaacac agcaaaaaag gaatggataa tcaagggcac aaaaatttga	2820
agacctcaca agaaggatca tctgatgatg ctgaaagaaa acaagagaga gagactttct	2880
cttcagcaga aggcacagtt gataaagaca cgaccatcat ggaattaaga gatcgacttc	2940
ctaagaagca gcaagcaagt gcttcactg atggtgtcga taagctttct gggaaagagc	3000
agagttttac ttctttggaa gttagaaaag ttgctgaaac taaagaaaag agcaagcatc	3060
tcaaaaccaa aacatgtaaa aaagtacagg atggcttattc tgatattgca gagaaattcc	3120
taaagaaaga ccagagcgat gaaacttctg aagatgataa aaagcagagc aaaaagggaa	3180
ctgaagaaaa aaagaaacct tcagacttta agaaaaaagt aattaaaatg gaacaacagt	3240
atgaatcttc atctgatggc actgaaaagt tacctgagcg agaagaaatt tgtcattttc	3300
ctaagggcat aaaacaaatt aagaatggaa caactgatgg agaaaagaaa agtaaaaaaa	3360
taagagataa aacttctaaa aagaaggatg aattatctga ttatgctgag aagtcaacag	3420
ggaaaggaga tagttgtgac tcttcagagg ataaaaagag taagaatgga gcatatggta	3480
gagagaagaa aaggtgcaag ttgcttggaa agagttcaag gaagagacaa gattgttcat	3540

catctgatac tgagaaatat tccatgaaag aagatggttg taactcttct gataagagac	3600
tgaaaagaat agaattgagg gaaagaagaa atttaagttc aaagagaaat actaaggaaa	3660
tacaaagtgg ctcatcatca tctgatgctg aggaaagttc tgaagataat aaaaagaaga	3720
agcaaagaac ttcattctaaa aagaaggcag tcattgtcaa ggagaaaaag agaaactccc	3780
taagaacaag cactaaaagg aagcaagctg acattacatc ctcatcttct tctgatatag	3840
aagatgatga tcagaattct ataggtgagg gaagcagcga tgaacagaaa attaagcctg	3900
tgactgaaaa tttagtgcctg tcttcacata ctggattttg ccaatcttca ggagatgaag	3960
ccttatctaa atcagtcgct gtcacagtgg atgatgatga tgacgacaat gatcctgaga	4020
atagaattgc caagaagatg cttttagaag aaattaaagc caatctttcc tctgatgagg	4080
atggatcttc agatgatgag ccagaagaag ggaaaaaaag aactggaaaa caaatgaag	4140
aaaaccagg agatgaggaa gcaaaaaatc aagtcaattc tgaatcagat tcagattctg	4200
aagaatctaa gaagccaaga tacagacata ggcttttgcg gcacaaattg actgtgagt	4260
acggagaatc tggagaagaa aaaaagacaa agcctaaaga gcataaagaa gtcaaaggca	4320
gaaacagaag aaaggtgagc agtgaagatt cagaagattc tgattttcag gaatcaggag	4380
ttagtgaaga agttagtgaa tccgaagatg aacagcggcc cagaacaagg tctgcaaaga	4440
aagcagagtt ggaagaaaat cagcggagct ataaacagaa aaagaaaagg cgacgtatta	4500
aggttcaaga agattcatcc agtgaaaaca agagtaattc tgaggaagaa gaggaggaaa	4560
aagaagagga ggaggaagag gaggaggagg aggaagagga ggaggaagat gaaaatgatg	4620
attccaagtc tcctggaaaa ggcagaaaga aaattcgga gattcttaaa gatgataaac	4680
tgagaacaga aacacaaaat gctcttaagg aagaggaaga gagacgaaaa cgtattgctg	4740
agagggagcg tgagcgagaa aaattgagag aggtgataga aattgaagat gcttcaccca	4800
ccaagtgtcc aataacaacc aagttggttt tagatgaaga tgaagaaacc aaagaacctt	4860
tagtgcaggt tcatagaaat atggttatca aattgaaacc ccatcaagta gatggtgttc	4920
agtttatgtg ggattgctgc tgtgagtctg tgaaaaaac aaagaaatct ccaggttcag	4980
gatgcattct tgcccactgt atgggccttg gtaagacttt acaggtggta agttttcttc	5040
atacagttct tttgtgtgac aaactggatt tcagcacggc gttagtggtt tgtcctctta	5100
atactgcttt gaattggatg aatgaatttg agaagtggca agagggatta aaagatgatg	5160

agaagcttga ggtttctgaa ttagcaactg tgaaacgtcc tcaggagaga agctacatgc	5220
tgcagagggtg gcaagaagat ggtgggtgtta tgatcatagg ctatgagatg tatagaaatc	5280
ttgctcaagg aaggaatgtg aagagtcgga aacttaaaga aatatttaac aaagctttgg	5340
ttgatccagg ccctgatttt gttgtttgtg atgaaggcca tattctaaaa aatgaagcat	5400
ctgctgtttc taaagctatg aattctatac gatcaaggag gaggattatt ttaacaggaa	5460
caccacttca aaataaccta attgagtatc attgtatggt taattttatc aaggaaaatt	5520
tacttggatc cattaaggag ttcaggaata gatttataaa tccaattcaa aatggtcagt	5580
gtgcagattc taccatggta gatgtcagag tgatgaaaaa acgtgctcac attctctatg	5640
agatgttagc tggatgtgtt cagaggaaag attatacagc attaacaaaa ttcttgcctc	5700
caaaacacga atatgtgtta gctgtgagaa tgacttctat tcagtgcaag ctctatcagt	5760
actacttaga tcacttaaca ggtgtgggca ataatagtga aggtggaaga ggaaaggcag	5820
gtgcaaagct tttccaagat tttcagatgt taagtagaat atggactcat ccttgggtgtt	5880
tgcagctaga ctacattagc aaagaaaata agggttatct tgatgaagac agtatggatg	5940
aatttatagc ctgagattct gatgaaacct ccatgagttt aagctccgat gattatacaa	6000
aaaagaagaa aaaagggaaa aaggggaaaa aagatagtag ctcaagtgga agtggcagtg	6060
acaatgatgt tgaagtgatt aaggtctgga attcaagatc tcggggagggt ggtgaaggaa	6120
atgtggatga aacaggaaac aatccttctg tttctttaaa actggaagaa agtaaagcta	6180
cttcttcttc taatccaagc agcccagctc cagactggta caaagatttt gttacagatg	6240
ctgatgctga ggttttagag cattctggga aaatgggtact tctctttgaa attcttcgaa	6300
tggcagagga aattggggat aaagtccttg ttttcagcca gtccctcata tctctggact	6360
tgattgaaga ttttcttgaa ttagctagta gggagaagac agaagataaa gataaacccc	6420
ttatttataa aggtgagggg aagtggcttc gaaacattga ctattaccgt ttagatggtt	6480
ccactactgc acagtcaagg aagaagtggg ctgaagaatt taatgatgaa actaatgtga	6540
gaggacgatt atttatcatt tctactaaag caggatctct aggaattaat ctggtagctg	6600
ctaactcagat aattatattc gacgcttctt ggaatccatc ttatgacatc cagagtatat	6660
tcagagttta tcgctttgga caaactaagc ctgtttatgt atataggttc ttagctcagg	6720
gaaccatgga agataagatt tatgatcggc aagtaactaa gcagtcactg tcttttcgag	6780

ttgttgatca gcagcaggtg gagcgtcatt ttactatgaa tgagcttact gaactttata	6840
cttttgagcc agacttatta gatgacccta attcagaaaa gaagaagaag agggatactc	6900
ccatgctgcc aaaggatacc atacttgagc agctccttca gatacataaa gaacacattg	6960
taggatacca tgaacatgat tctcttttgg accacaaaga agaagaagag ttgactgaag	7020
aagaaagaaa agcagcttgg gctgagtatg aagcagagaa gaagggactg accatgcgtt	7080
tcaacatacc aactgggacc aattttacccc ctgtcagttt caactctcaa actccttata	7140
ttcctttcaa tttgggagcc ctgtcagcaa tgagtaatca acagctggag gacctcatta	7200
atcaaggaag agaaaaagt gtagaagcaa caaacagtgt gacagcagtg aggattcaac	7260
ctcttgagga tataatttca gctgtatgga aggagaacat gaatctctca gaggcccaag	7320
tacaggcggt agcattaagt agacaagcca gccaggagct tgatgttaaa cgaagagaag	7380
caatctacaa tgatgtattg acaaaacaac agatgttaat cagctgtgtt cagcgaatac	7440
ttatgaacag aaggctccag cagcagtaca atcagcagca acagcaacaa atgacttatac	7500
aacaagcaac actgggtcac ctcatgatgc caaagcccc aaatttgatc atgaatcctt	7560
ctaactacca gcagattgat atgagaggaa tgtatcagcc agtggctggt ggtatgcagc	7620
caccaccatt acagcgtgca ccacccccaa tgagaagcaa aaatccagga ccttcccaag	7680
ggaaatcaat gtgattttgc actaaaagct taatggattg ttaaaatcat agaaagatct	7740
tttatTTTTT taggaatcaa tgacttaaca gaactcaact gtataaatag tttggtcccc	7800
ttaaagcca atcttcata ttagttttac tttttttttt tttaaatagg gcataccatt	7860
tcttcctgac atttgtcagt gatgttgctt agaatcttct tacacacgct gagtacagaa	7920
gatatttcaa attgttttca gtgaaaacaa gtccttccat aatagtaaca actccacaga	7980
tttcctctct aaatttttat gcctgctttt agcaaccata aaattgtcat aaaattaata	8040
aatttaggaa agaataaaga tttatatatt cattctttac atataaaaac acacagctga	8100
gttcttagag ttgattcctc aagttatgaa atacttttgt acttaatcca tttcttgatt	8160
aaagtgattg aaatggtttt aatgttcttt tgactgaagt ctgaaactgg gctcctgctt	8220
tattgtctct gtgactgaaa gttagaaact gagggttatc tttgacacag aattgtgtgc	8280
aatattctta aatactactg ctctaaaagt tggagaagtc ttgcagttat cttagcattg	8340
tataaacagc cttaagtata gcctaagaag agaattcctt tttcttcttt agtccttctg	8400

ccatttttta	ttttcagtta	tatgtgctga	aataattact	ggtaaaat	cagggttg	8460
gattatcttc	cacacatgaa	ttttctctct	cctggcacga	atataaagca	catctcttaa	8520
ctgcatgggtg	ccagtgcata	tgcttcatcc	tggtgctggc	agtgggatgt	ggacttagaa	8580
aatcaagttc	tagcatttta	gtaggttaac	actgaagttg	tggttgttag	gttcacaccc	8640
tgttttataa	acaacatcaa	aatggcagaa	ccattgctga	ctttaggttc	acatgaggaa	8700
tgtactttta	acaattccca	gtactatcag	tattgtgaaa	taattcctct	gaaagataag	8760
aatcactggc	ttctatgcgc	ttcttttctc	tcatcatcat	gttcttttac	cccagtttcc	8820
ttacattttt	ttaaattggt	tcagagtttg	tttttttttt	agtttagatt	gtgaggcaat	8880
tattaaatca	aaattaattc	atccaatacc	cctttactag	aagttttact	agaaaatgta	8940
ttacatttta	ttttttctta	atccagttct	gcaaaaatga	cctataaatt	tattcatgta	9000
caatttttgt	tacttgaatt	gttaaagaaa	acattgtttt	tgactatggg	agtcaactca	9060
acatggcaga	accatttttg	agatgatgat	acaacaggta	gtgaaacagc	ttaagaattc	9120
caaaaaaaaa	aaaaaaaaaa	aaaaaaagaa	aactgggttt	gggctttgct	ttaggtatca	9180
ctggattaga	atgagtttaa	cattagctaa	aactgctttg	agttgtttgg	atgattaaga	9240
gattgccatt	tttatcttgg	aagaactagt	ggtaaaacat	ccaagagcac	taggattgtg	9300
atacagaatt	tgtgaggttt	ggtggatcca	cgcccctctc	ccccactttc	ccatgatgaa	9360
atatcactaa	taaatcctgt	atatttagat	attatgctag	ccatgtaatc	agattttatt	9420
aattgggtgg	ggcaggtgtg	tatttacttt	agaaaaaatg	aaaaagacaa	gatttatgag	9480
aaatatttga	aggcagtaca	ctctggccaa	ctgttaccag	ttggtatttc	tacaagttca	9540
gaatatttta	aacctgattt	actagacctg	ggaattttca	acatggtcta	attattttact	9600
caaagacata	gatgtgaaaa	ttttaggcaa	ccttctaaat	ctttttcacc	atggatgaaa	9660
ctataactta	aagaataata	cttagaaggg	ttaattggaa	atcagagttt	gaaataaaac	9720
ttggaccact	ttgtatacac	tcttctcact	tgacatttta	gctatataat	atgtactttg	9780
agtataacat	caagctttta	caaatattta	aagacaaaaa	aatcacgtca	gtaaaatact	9840
aaaaggctca	tttttatatt	tgtttttagat	gttttaaata	gttgcaatgg	attaaaaatg	9900
atgatttaaa	atgttgcttg	taatacagtt	ttgcctgcta	aattctccac	attttgaac	9960
ctgttttatt	tctttgggtg	taaagcgttt	ttgcttagta	ttgtgatatt	gtatatgttt	10020

tgtcccagtt gtatagtaat gtttcagtcc atcatccagc tttggctgct gaaatcatac	10080
agctgtgaag acttgccttt gtttctgtta gactgctttt cagttctgta ttgagtatct	10140
taagtactgt agaaaagatg tcacttcttc ctttaaggct gttttgtaat atatataagg	10200
actggaattg tgtttttaaa gaaaagcatt caagtatgac aatatactat ctgtgttttc	10260
accattcaaa gtgctgttta gtagttgaaa cttaaactat ttaatgtcat ttaataaagt	10320
gaccaaaatg tgttgtgctc tttattgtat tttcacagct ttgaaaatct gtgcacatac	10380
tgtttcatag aaaatgtata gcttttgttg tcctatataa tgggtggttct tttgcacatt	10440
tagttattta atattgagag gtcacgaagt ttggttattg aatctgttat atactaaatt	10500
ctgtaaaggg agatctctca tctcaaaaag aatttacata ccaggaagtc catgtgtgtt	10560
tgtgttagtt ttggatgtct ttgtgtaatc cagccccatt tcctgtttcc caacagctgt	10620
aacactcatt ttaagtcaag cagggtacc aaccacact tgatagaaaa gctgcttacc	10680
attcagaagc ttccttatta cctggcctcc aaatgagctg aatattttgt agccttcct	10740
tagctatgtt cattttccct ccattatcat aaaatcagat cgatatttat gtgccccaaa	10800
caaaacttta agagcagtta cattctgtcc cagtagccct tgtttccttt gagagtagca	10860
tgttgtgagg ctatagagac ttattctacc agtaaaacag gtcaatcctt ttacatgttt	10920
attatactaa aaattatgtt cagggtattt actactttat ttcaccagac tcagtctcaa	10980
gtgacttggc tatctccaaa tcagatctac ccttagagaa taaacatttt tctaccgtta	11040
ttttttttca agtctataat ctgagccagt cccaaaggag tgatcaagtt tcagaaatgc	11100
tttcatcttc acaacatttt atatatacta ttatatgggg tgaataaagt tttaaatccg	11160
aaatata	11167

<210> 7

<211> 2590

<212> DNA

<213> Homo sapiens

<220>

<223> >ATRIP|ENSG00000164053|ENST00000320211|2590

<400> 7

gggcactcgc ggcggaggca agcggcggcg cgcggacggt tgggtccagtt ctccggcctg	60
--	----

gcggcaggca agtctagctc ggcgctgtcg gatacttggg gtgagcggaa agcatggcgg	120
ggacctccgc gccaggcagc aagaggcggg gcgagccccc ggcgccctcg cccggcccg	180
cgccggggcac cgggcacccc ccgagcaagc gggcccgggg cttctccgca gccgtgccc	240
cggaccctga cgaccggtt ggcgcgcatg gggacttcac tgccgacgac ctggaggagc	300
ttgacaccct cgcgtcacag gccctgagcc aatgtccggc cgcggtcgg gacgtgtcca	360
gtgatcataa ggtccacaga ttattagatg gcatgtcaaa aaatccttca gggaaaaaca	420
gagaaactgt tccaattaaa gataatttcg aattagaggt acttcaggca caatacaaag	480
aacttaaaga aaagatgaaa gtaatggaag aagaagttct cattaagaat ggagaaatta	540
aaattttgcg agactcacta catcagacgg aatccgttct agaggaaacag agaagatcac	600
attttcttct tgagcaagag aaaaccaag cactcagtga caaggaaaag gaattctcca	660
aaaagctcca atcattgcag tctgaactcc agtttaaaga tgcagagatg aatgaattaa	720
ggacaaaagct ccagaccagt gaacgagcaa ataaactggc tgctccctct gtttcccatg	780
tcagtcctag gaaaaaccct tctgtgggta taaagccaga agcatgttct ccacaatttg	840
gaaaaacatc ttttcctaca aaggagtctt ttagtgctaa catgtccctt cccacccct	900
gccagacgga gtcaggatac aagcctctgg tgggcagaga ggatagtaag cccacagtc	960
tgagaggtga ctccataaaa caagaagagg ccagaaaag ctttgttgac agctggagac	1020
agagatcaaa cactcaaggt tccattttga taaacctgct cctgaagcag cttttgatcc	1080
cagggtcac cctaagcctt tgccacctcc tgagtagtag ttctgagtct cctgctggca	1140
ccccctgca gccaccaggg tttggcagta cttggctgg aatgtcaggc ctcaggacca	1200
caggttctta tgatgggtca ttttccctct cagccctgag agaagcacag aacctggcat	1260
tcactggact gaatctgggt gcccggaatg agtgctcacg tgatggagac ccagcagagg	1320
gaggcagaag ggccttcca ctctgccagc ttcctggagc cgtgcatttc ctcccccttg	1380
tacagttctt catcggtta cactgccagg ccctgcagga cttggcagct gctaagagaa	1440
gcggagcacc tggggactca ccgacacatt cctcctgcgt gagctctggg gtagagacca	1500
accctgagga ctcagtgtgc atcctggaag gcttctctgt gactgcactt agcattcttc	1560
agcacctggg gtgccacagc ggagcagtcg tctccctatt actgtcagga gtgggggcag	1620
attctgctgc tggggaagga aacaggagcc tggttcacag gcttagtgat ggagatatga	1680

cctcagccct aaggggggtt gctgatgacc aaggacagca cccactgttg aagatgcttc	1740
ttcacctgtt ggcttttctt tctgcagcaa caggtcacct tcaagccagt gtcctgaccc	1800
agtgccttaa ggttttgggt aaattagccg aaaacacttc ctgtgatttc ttgccaggt	1860
tccagtgtgt gttccaagt ctgccaaagt gcctcagccc agagacaccc ctgcctagcg	1920
tgctgctggc tgttgagctc ctctccctgc tggcggacca cgaccagctg gcacctcagc	1980
tctgttccca ctcagaaggc tgcctcctgc tgctgctgta catgtacatc acatcacggc	2040
ctgacagagt ggccttggag acacaatggc tccagctgga acaagaggtg gtgtggctcc	2100
tggctaagct tgggtgtcag agccccctgc cccagtcac tggctccaac tgccagtgt	2160
atgtggaggt ggtcagagcg ctcacggtga tgttgcacag acagtggctg acagtgcgga	2220
gggcaggggg accccaagg accgaccagc agaggcggac agtgcgctgt ctgcgggaca	2280
cgggtgtgct gctgcacggc ctatcgaga aggacaagct cttcatgatg cactgcgtgg	2340
aggtcctgca tcagtttgac caggtgatgc cgggggtcag catgctcatc cgagggttc	2400
ctgatgtgac ggactgtgaa gaggcagccc tggatgacct ctgtccgcg gaaaccgatg	2460
tggaagaccc cgaggtggag tgtggctgag gccctgagt tccagccaca tgggtggcacc	2520
agcaccactc ctttccttac cacatcaact gattaaagca gtgaccagca ggaactgccc	2580
agagaactgg	2590

<210> 8

<211> 3937

<212> DNA

<213> Homo sapiens

<220>

<223> >BAP1|ENSG00000163930|ENST00000460680|3937

<400> 8

cgctccgcc ctcccctgc agcaccggg ctagtactg cccgtccgc ccctccttc	60
gagcctcagc gtcagcatc gcccggacc cctcttccct tcgcccgcct cgtcccgacc	120
ctccccttcg ccccgctcc gcccgcgcc tccccttcgc ccccgctccg tccgccccg	180
cccctcccct tcgccccgt ccctcccctt cgcgccgct cctccgcgcg tgcgcgttcg	240
ccttcgagcg catgcccga tctgctgtcc gacaggcgg agacagccc agaggcggag	300
cagggccgtc gcgccttggg gacgtctgcc gccggcgcgg gcgggtgac cgactgggcc	360

cgttgctctgt gtgtgggact gaggggcccc gggggcggtg ggggctcccg gtgggggcag	420
cgggtggggag ggagggcctg gacatggcgc tgaggggccc ccccgcgga agatgaataa	480
gggctggctg gagctggaga gcgaccagg cctcttcacc ctgctcgtgg aagatttcgg	540
tgtcaagggg gtgcaagtgg aggagatcta cgaccttcag agcaaattgc agggccctgt	600
atatggattt atcttcctgt tcaaattgat cgaagagcgc cgggtcccggc gaaaggtctc	660
taccttggtg gatgatacgt ccgtgattga tgatgatatt gtgaataaca tgttctttgc	720
ccaccagctg atacccaact cttgtgcaac tcatgccttg ctgagcgtgc tcctgaactg	780
cagcagcgtg gacctgggac ccaccctgag tcgcatgaag gacttcacca agggtttcag	840
ccctgagagc aaaggatatg cgattggcaa tgccccggag ttggccaagg ccataatag	900
ccatgccagg cccgagccac gccacctccc tgagaagcag aatggcctta gtgcagtgcg	960
gaccatggag gcgttcact ttgtcagcta tgtgcctatc acaggccggc tctttgagct	1020
ggatgggctg aaggtctacc ccattgacca tgggccctgg ggggaggacg aggagtggac	1080
agacaaggcc cggcgggtca tcatggagcg tatcggcctc gccactgcag gggagcccta	1140
ccacgacatc cgcttcaacc tgatggcagt ggtgcccgc cgaggatca agtatgaggc	1200
caggctgcat gtgctgaagg tgaaccgtca gacagtacta gaggctctgc agcagctgat	1260
aagagtaaca cagccagagc tgattcagac ccacaagtct caagagtcac agctgcctga	1320
ggagtccaag tcagccagca acaagtcccc gctggtgctg gaagcaaaca gggcccctgc	1380
agcctctgag ggcaaccaca cagatggtgc agaggaggcg gctggttcat gcgcacaagc	1440
cccatccac agccctcca acaaaccac gctagtgggtg aagcctccag gcagcagcct	1500
caatgggggt caccacaacc ccactcccat tgtccagcgg ctgccggcct ttctagacaa	1560
tcacaattat gccaaagtc ccatgcagga ggaagaagac ctggcggcag gtgtgggccg	1620
cagccgagtt ccagtccgc caccacagca gtactcagat gatgaggatg actatgagga	1680
tgacgaggag gatgacgtgc agaacaccaa ctctgccctt aggtataagg ggaagggaac	1740
agggaagcca ggggcattga gcggttctgc tgatgggcaa ctgtcagtg tgcagcccaa	1800
cacatcaac gtcttggtg agaagctcaa agagtcccag aaggacctct caattcctct	1860
gtccatcaag actagcagcg gggctgggag tccggctgtg gcagtgccca cacactcgca	1920
gccctcacc accccagca atgagagtac agacacggcc tctgagatcg gcagtgtttt	1980

caactcgcca ctgcgctcgc ctatccgctc agccaacccg acgcggccct ccagccctgt	2040
cacctccac atctccaagg tgctttttgg agaggatgac agcctgctgc gtgttgactg	2100
catacgctac aaccgtgctg tccgtgatct gggtcctgtc atcagcacag gcctgctgca	2160
cctggctgag gatggggtgc tgagtcccct ggcgctgaca gaggggtggga agggttcctc	2220
gccctccatc agaccaatcc aaggcagcca ggggtccagc agcccagtgg agaaggaggt	2280
cgtggaagcc acggacagca gagagaagac ggggatgggtg aggcctggcg agcccttgag	2340
tggggagaaa tactcaccca aggagctgct ggcactgctg aagtgtgtgg aggctgagat	2400
tgcaaaactat gaggcgtgcc tcaaggagga ggtagagaag aggaagaagt tcaagattga	2460
tgaccagaga aggaccaca actacgatga gttcatctgc acctttatct ccatgctggc	2520
tcaggaaggc atgctggcca acctagtggg gcagaacatc tccgtgcggc ggcgccaagg	2580
ggtcagcatc ggccggctcc acaagcagcg gaagcctgac cggcggaaac gctctcgccc	2640
ctacaaggcc aagcgccagt gaggactgct ggccctgact ctgcagccca ctcttgccgt	2700
gtggccctca ccagggtcct tccctgcccc acttcccctt ttcccagtat tactgaatag	2760
tcccagctgg agagtccagg ccctgggaat gggaggaacc aggccacatt ccttccatcg	2820
tgccctgagg cctgacacgg cagatcagcc ccatagtgtc caggaggcag catctggagt	2880
tggggcacag cgaggctactg cagcttcctc cacagccggc tgtggagcag caggacctgg	2940
cccttctgcc tgggcagcag aatatatatt ttacctatca gagacatcta tttttctggg	3000
ctccaacca acatgccacc atgttgacat aagttcctac ctgactatgc tttctctcct	3060
aggagctgtc ctgggtgggc caggtccttg tatcatgcca cgggtccaac tacagggtcc	3120
tagctggggg cctgggtggg ccctgggctc tgggccctgc tgctctagcc ccagccacca	3180
gcctgtccct gttgtaagga agccaggtct tctctcttca ttcctcttag gagagtgcca	3240
aactcaggga cccagcactg ggctggggtg ggagtagggt gtcccagtgg gggtggggtg	3300
agcaggctgc tgggatccca tggcctgagc agagcatgtg ggaactgttc agtggcctgt	3360
gaactgtctt cttgtttcta gccaggctgt tcaagactgc tctccatagc aaggttctag	3420
ggctcttcgc cttcagtgtt gtggccctag ctatgggcct aaattgggct ctaggtctct	3480
gtccctggcg cttgaggctc agaagagcct ctgtccagcc cctcagtatt accatgtctc	3540
cctctcaggg gtagcagaga cagggttgct tataggaagc tggcaccact cagctcttcc	3600

tgctactcca gtttcctcag cctctgcaag gcactcaggg tgggggacag caggatcaag	3660
acaacccgtt ggagcccctg tgttccagag gacctgatgc caaggggtaa tgggccagc	3720
agtgcctctg gagcccaggc cccaacacag ccccatggcc tctgccagat ggctttgaaa	3780
aaggtgatcc aagcaggccc ctttatctgt acatagtac tgagtggggg gtgctggcaa	3840
gtgtggcagc tgcctctggg ctgagcacag cttgaccct ctagccctg taaatactgg	3900
atcaatgaat gaataaaact ctcctaagaa tctcctg	3937

<210> 9
 <211> 5499
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >BARD1|ENSG00000138376|ENST00000260947|5499

cctctggcgg cccgccgtcc cagacgcggg aagagcttgg ccggtttcga gtcgctggcc	60
tgcagcttcc ctgtggtttc ccgaggcttc cttgcttccc gctctgcgag gaggctttca	120
tccgaaggcg ggacgatgcc ggataatcgg cagccgagga accggcagcc gaggatccgc	180
tccgggaacg agcctcgttc cgcgcccgc atggaaccgg atggtcgcgg tgcctgggcc	240
cacagtcgcg ccgcgctcga ccgcctggag aagctgctgc gctgctcgcg ttgtactaac	300
attctgagag agcctgtgtg tttaggagga tgtgagcaca tcttctgtag taattgtgta	360
agtgactgca ttggaactgg atgtccagtg tgttacaccc cggcctggat acaagacttg	420
aagataaata gacaactgga cagcatgatt caactttgta gtaagcttcg aaatttgcta	480
catgacaatg agctgtcaga tttgaaagaa gataaaccta ggaaaagttt gtttaatgat	540
gcaggaaaca agaagaattc aattaaaatg tggtttagcc ctcgaagtaa gaaagtcaga	600
tatgttgatg gtaaagcttc agtgcaaacc cagcctgcaa taaaaaaga tgcaagtgt	660
cagcaagact catatgaatt tgtttcccca agtcctcctg cagatgtttc tgagagggct	720
aaaaaggctt ctgcaagatc tggaaaaaag caaaaaaaga aaactttagc tgaaatcaac	780
caaaaatgga atttagaggc agaaaaagaa gatggtgaat ttgactcaa agaggaatct	840
aagcaaaagc tggtatcctt ctgtagccaa ccatctgtta tctccagtcc tcagataaat	900

ggtgaaatag acttactagc aagtggctcc ttgacagaat ctgaatgttt tggaagttta	960
actgaagtct ctttaccatt ggctgagcaa atagagtctc cagacactaa gagcaggaat	1020
gaagtagtga ctcctgagaa ggtctgcaaa aattatctta catctaagaa atctttgcc	1080
ttagaaaata atggaaaacg tggccatcac aatagacttt ccagtcccat ttctaagaga	1140
tgtagaacca gcattctgag caccagtgga gattttgtta agcaaacggt gccctcagaa	1200
aatataccat tgcctgaatg ttcttcacca cttcatgca aacgtaaagt tgggtgtaca	1260
tcaggaggga aaaacagtaa catgtccgat gaattcatta gtctttcacc aggtacacca	1320
ccttctacat taagtagttc aagttacagg cgagtgatgt ctagtccctc agcaatgaag	1380
ctgttgccca atatggctgt gaaaagaaat catagaggag agactttgct ccatattgct	1440
tctattaagg gcgacatacc ttctgttgaa taccttttac aaaatggaag tgatccaaat	1500
gttaaagacc atgctggatg gacaccattg catgaagctt gcaatcatgg gcacctgaag	1560
gtagtggaat tattgctcca gcataaggca ttggtgaaca ccaccgggta tcaaatgac	1620
tcaccacttc acgatgcagc caagaatggg catgtggata tagtcaagct gttactttcc	1680
tatggagcct ccagaaatgc tgtaaatata tttggctctgc ggctgtcga ttatacagat	1740
gatgaaagta tgaaatcgct attgctgcta ccagagaaga atgaatcatc ctcagctagc	1800
cactgctcag taatgaacac tgggcagcgt agggatggac ctcttgact tataggcagt	1860
gggctgtctt cagaacaaca gaaaatgctc agtgagcttg cagtaattct taaggctaaa	1920
aaatatactg agtttgacag tacagtaact catgttggtg ttcttggtga tgcagttcaa	1980
agtaccttga agtgtatgct tgggattctc aatggatgct ggattctaaa atttgaatgg	2040
gtaaaagcat gtctacgaag aaaagtatgt gaacaggaag aaaagtatga aattcctgaa	2100
ggtccacgca gaagcaggct caacagagaa cagctgttgc caaagctggt tgatggatgc	2160
tacttctatt tgtggggaac cttcaaacac catccaaagg acaaccttat taagctcgtc	2220
actgcaggtg ggggccagat cctcagtaga aagccaagc cagacagtga cgtgactcag	2280
accatcaata cagtcgcata ccatgcgaga cccgattctg atcagcgctt ctgcacacag	2340
tatatcatct atgaagattt gtgtaattat caccagaga gggttcggca gggcaaagtc	2400
tggaaggctc cttcgagctg gtttatagac tgtgtgatgt cctttgagtt gcttcctctt	2460
gacagctgaa tattatacca gatgaacatt tcaaattgaa tttgcacggt ttgtgagagc	2520

ccagtcattg tactgttttt aatgttcaca tttttacaaa taggtagagt cattcatatt	2580
tgtctttgaa tcaaaaaaaaa aaaaaaagtc taatgccaga ttaggaattc atgttatgtt	2640
taccatttag aagctgggat tgctttttaa ggtttttctt tttaaaattg gcatgttttt	2700
gatttatcat gtctttctat tcagattatt gggatatcaa gattaatgag gacaccagaa	2760
tcttggttaa atagacaagt ggtatcatta ctgtttgagt cttttaatat tctccatacc	2820
tgccaccagt gaaaaaactt gccttttttt tttttttttt tttagtaaac agaataattat	2880
caaacaattt attttggctt tattgaaaaa agagtatttg gtctaaatgt gccaccatag	2940
gtgttaaatt ctcctatctg caattgtctt tatcctatat tgtgttcatt tcttttctta	3000
ataatttact ttgttgtgtg tttctacact ttcatccctg ttttttatct tgtatatcat	3060
caggaaattg tgatttaatc attaacattg gtttttttgt gtgtgtggta aaaatcaaca	3120
ctaggctcat ggtacatatt tttattctgt acatttgctt gtaactatca atttgtaact	3180
ctgtttatct actacatgtg tatatatact tagagcattt tctctaacac attttaatgt	3240
tagtattttt taaaaggctt gaccagtcta gcaaattgtc agtccaacgt cattacttta	3300
aattaagaag cagtcttctt ctggtaaacc ttgttgggtat ttgtaaaata attttgaagg	3360
tcttaatttc ttcctttgta aaaggaaaag gtttttttta aagtttttag gttggcatgg	3420
aggcagaagt tgggtgattac ttgatttaca acagattttt tccagatcat acaaaaggcc	3480
atacagtaag tatagaagta ggtatgggga gggcttacta atatcaaata ggcaaggcct	3540
tagtgagtgg gcaggatacc acctgagagt ggccagatgt ggggagggtta ctctgctctg	3600
ggtgctctca ttcatgaatc gacaaggata cattagatta ttttgaaaca tttttttaag	3660
aagcagaatt ctttaataat tccttcctag acattgaata tacttataaa attaaagact	3720
tggggaagga gacactgaga gacttgccag tttggttcct catgaacaaa agaggacagt	3780
ttgataacta ccagaataga atatccctag ttttaaaata gtgagaatct ctgaagttca	3840
tcaacatctt aagatgcact tacttgaaag tttgagattc tgtttatcat ttgaaaacac	3900
attttgcttt aattctttct ttgacatgtt gttttttcat atcaagaaat atatgaacaa	3960
aataataacc ttttgaccct gacctgctg ggtgaattag ctctgaaaca ctctctacaa	4020
ccagtaatgc atttgtcca catttcattc tgatagaaaa tgaacaccat agcaccaaac	4080
aaaaatccga ggcgttagat aatgtctgga ttaaataatt taagactctc taggattttg	4140

gttgtcattt tttatttata acagacttta agtcactttc tgttgccctca taggtcacat	4200
tttagacagg tttgtgtctg ttccttgcat ctgaattcct gattgtaaag acacctatga	4260
ggtctcttag tttttgtcat tcattttctt ggtttatcac ccctcccttc tttttgttgt	4320
ttttccctga ctgttaagca gtttcatctt tgcttttggt aaatatttga cagcagttag	4380
tttgtgttaa gctcttgaaa cttgtgattg tactttctgt gtagatatac atgtaattat	4440
tttttatttt tcaatcatag attcaagctt ctttcttttt taccacaaat cattaaagtt	4500
atttgtgttt ccatatacct gtgtcttgta taaaattggc ttattctgtg ctgttgaatg	4560
aggctcaaca tgacttggtg aggaagtcta ttaactaaca aaagcttatc ttttttaaca	4620
taatgctttt taattaattt tgaataaaaa tatttctaaa gtgtactaga tactttatta	4680
ccttagatta ttccgaatac agtataactt tgatagtttg gaatagtcac taagaaacaa	4740
ttacacactg attgctttgt gtctctaaaa gtgagaggct ggtagctttt ccacattctc	4800
atggctattt tctagttcta cttgaattta taactgtttc cttttttcct tgacagctgc	4860
cactttgtag ctatttttct gtctctgcta atactttacc atatctatct caattgtttt	4920
ttcttttgac ttgctgaaaa atagaaacca gatgggaagt atattagcat tatgattgaa	4980
ataagggtaa atgagcaatg tgtgaagggt ttcactgact tcacctaaaa gatagtttag	5040
ctacttgaat tttagtaaag agaatttttc ctttatttca tcggtcccc cactttttt	5100
tttttttgca cctgccttgt aaatttaata gttaagtgc ctctgcctag aggatgatat	5160
ttggggagggt ttgatgtttc ctgtgggaat aagacgattc acaggtgaga gtggggccac	5220
attagctgtt attgtttcca tgggtcagtg tggaaaatgc attaatacata ttctaaacgt	5280
tcatgggcct cattacagtc acaattgtct attctgtttc ctaccctgaa cacattaaaa	5340
tggtaggaac taatgcttgt cttatttaat tactaaaagc caccattttc tttgatagat	5400
tgagctacag attgtaaact tcatgtattt ctttataagt caaccctttt caaagatacg	5460
cacatcaaac tgaatgaata aataaatatt gagaagttg	5499

<210> 10

<211> 4555

<212> DNA

<213> Homo sapiens

<220>

<223> >BLM|ENSG00000197299|ENST00000355112|4555

<400> 10

aatcgggaata ggcaagcttc cggcgggaag tgagccaggg cttggcgcg cggccgtggt	60
tgcggcgcgg gaagtttgga tcctggttcc gtccgctagg agtctgcgtg cgaggattat	120
ggctgctgtt cctcaaaata atctacagga gcaactagaa cgtcactcag ccagaacact	180
taataataaa ttaagtcttt caaaacaaa attttcaggt ttcactttta aaaagaaaac	240
atcttcagat aacaatgtat ctgtaactaa tgtgtcagta gcaaaaacac ctgtattaag	300
aaataaagat gttaatgtta ccgaagactt ttccttcagt gaacctctac ccaacaccac	360
aaatcagcaa aggggtcaagg acttctttaa aaatgctcca gcaggacagg aaacacagag	420
aggtggatca aaatcattat tgccagattt cttgcagact ccgaaggaag ttgtatgcac	480
tacccaaaac acaccaactg taaagaaatc ccgggatact gctctcaaga aattagaatt	540
tagttcttca ccagattctt taagtaccat caatgattgg gatgatatgg atgactttga	600
tacttctgag acttcaaaat catttgttac accaccccaa agtcactttg taagagtaag	660
cactgctcag aaatcaaaaa agggtaagag aaactttttt aaagcacagc tttatacaac	720
aaacacagta aagactgatt tgcctccacc ctctctgaa agcgagcaaa tagatttgac	780
tgaggaacag aaggatgact cagaatggtt aagcagcgat gtgatttgca tcgatgatgg	840
ccccattgct gaagtgcata taaatgaaga tgctcaggaa agtgactctc tgaaaactca	900
tttggagat gaaagagata atagcgaaaa gaagaagaat ttggaagaag ctgaattaca	960
ttcaactgag aaagttccat gtattgaatt tgatgatgat gattatgata cggattttgt	1020
tccaccttct ccagaagaaa ttatttctgc ttcttcttcc tcttcaaaat gccttagtac	1080
gttaaaggac cttgacacct ctgacagaaa agaggatggt cttagcacat caaaagatct	1140
tttgtcaaaa cctgagaaaa tgagtatgca ggagctgaat ccagaaacca gcacagactg	1200
tgacgctaga cagataagtt tacagcagca gcttattcat gtgatggagc acatctgtaa	1260
attaattgat actattcctg atgataaact gaaacttttg gattgtggga acgaactgct	1320
tcagcagcgg aacataagaa ggaaacttct aacggaagta gattttaata aaagtgatgc	1380
cagtcttctt ggctcattgt ggagatacag gcctgattca cttgatggcc ctatggaggg	1440
tgattcctgc cctacagga attctatgaa ggagttaaat ttttcacacc ttccctcaaa	1500
ttctgtttct cctggggact gtttactgac taccacccta ggaaagacag gattctctgc	1560

caccaggaag aatctttttg aaaggccttt attcaatacc catttacaga agtcctttgt	1620
aagtagcaac tgggctgaaa caccaagact aggaaaaaaa aatgaaagct cttatttccc	1680
aggaaatgtt ctcacaagca ctgctgtgaa agatcagaat aaacatactg cttcaataaa	1740
tgacttagaa agagaaaccc aaccttccta tgatattgat aattttgaca tagatgactt	1800
tgatgatgat gatgactggg aagacataat gcataattta gcagccagca aatcttccac	1860
agctgcctat caacccatca aggaaggctg gccaatataa tcagtatcag aaagactttc	1920
ctcagccaag acagactgtc ttccagtgtc atctactgct caaaatataa acttctcaga	1980
gtcaattcag aattatactg acaagtcagc acaaaattta gcatccagaa atctgaaaca	2040
tgagcgtttc caaagtctta gttttcctca tacaaaggaa atgatgaaga tttttcataa	2100
aaaatttggc ctgcataatt ttagaactaa tcagctagag gcgatcaatg ctgcactgct	2160
tggatgaagac tgttttatcc tgatgccgac tggaggtggg aagagtttgt gttaccagct	2220
ccctgcctgt gtttctcctg gggtcactgt tgtcatttct cccttgagat cacttatcgt	2280
agatcaagtc caaaagctga cttccttgga tattccagct acatatctga caggtgataa	2340
gactgactca gaagctacaa atatttacct ccagttatca aaaaagacc caatcataaa	2400
acttctatat gtcactccag aaaagatctg tgcaagtaac agactcattt ctactctgga	2460
gaatctctat gagaggaagc tcttggcacg ttttgttatt gatgaagcac attgtgtcag	2520
tcagtgggga catgattttc gtcaagatta caaaagaatg aatatgcttc gccagaagtt	2580
tccttctgtt ccggtgatgg ctcttacggc cacagctaata cccagggtac agaaggacat	2640
cctgactcag ctgaagattc tcagacctca ggtgttttagc atgagcttta acagacataa	2700
tctgaaatac tatgtattac cgaaaaagcc taaaaagggtg gcatttgatt gcctagaatg	2760
gatcagaaaag caccacccat atgattcagg gataatttac tgcctctcca ggcgagaatg	2820
tgacaccatg gctgacacgt tacagagaga tgggctcgct gctcttgctt accatgctgg	2880
cctcagtgat tctgccagag atgaagtgca gcagaagtgg attaatcagg atggctgtca	2940
ggttatctgt gctacaattg catttggaat ggggattgac aaaccggacg tgcgatttgt	3000
gattcatgca tctctcccta aatctgtgga gggttactac caagaatctg gcagagctgg	3060
aagagatggg gaaatatctc actgcctgct tttctatacc tatcatgatg tgaccagact	3120
gaaaagactt ataatgatgg aaaaagatgg aaaccatcat acaagagaaa ctcacttcaa	3180

taatttgtat agcatggtac attactgtga aaatataacg gaatgcagga gaatacagct	3240
tttggcctac tttggtgaaa atggatttaa tcctgatttt tgtaagaaac acccagatgt	3300
ttcttgtgat aattgctgta aaacaaagga ttataaaaca agagatgtga ctgacgatgt	3360
gaaaagtatt gtaagatttg ttcaagaaca tagttcatca caaggaatga gaaatataaa	3420
acatgtaggt ctttctggaa gatttactat gaatatgctg gtcgacattt tcttggggag	3480
taagagtgca aaaatccagt caggtatatt tggaaaagga tctgcttatt cacgacacaa	3540
tgccgaaaga ctttttaaaa agctgatact tgacaagatt ttggatgaag acttatatat	3600
caatgccaat gaccaggcga tcgcttatgt gatgctcgga aataaagccc aaactgtact	3660
aaatggcaat ttaaaggtag actttatgga aacagaaaat tccagcagtg tgaaaaaaca	3720
aaaagcgta gtagcaaaag tgtctcagag ggaagagatg gttaaaaaat gtcttggaga	3780
acttacagaa gtctgcaaat ctctggggaa agtttttgggt gtccattact tcaatatttt	3840
taataccgtc actctcaaga agcttgcaga atctttatct tctgacctg aggttttgct	3900
tcaaattgat ggtgttactg aagacaaact ggaaaaatat ggtgcggaag tgatttcagt	3960
attacagaaa tactctgaat ggacatcgcc agctgaagac agttccccag ggataagcct	4020
gtccagcagc agaggccccg gaagaagtgc cgctgaggag ctcgacgagg aaatacccg	4080
atcttccac tactttgcaa gtaaaaccag aaatgaaagg aagaggaaaa agatgccagc	4140
ctcccaaagg tctaagagga gaaaaactgc ttccagtgggt tccaaggcaa agggggggctc	4200
tgccacatgt agaaagatat cttccaaaac gaaatcctcc agcatcattg gatccagttc	4260
agcctcacat acttctcaag cgacatcagg agccaatagc aaattgggga ttatggctcc	4320
accgaagcct ataaatagac cgtttcttaa gccttcatat gcattctcat aacaaccgaa	4380
tctcaatgta catagaccct ctttcttggt tgtcagcatc tgaccatctg tgactataaa	4440
gctgttattc ttgttatacc atttgaagtt ttactcgtc tctattaata tttaaataaa	4500
tgctgggggg tgatagttct tctttttaaa ataaacattt tcttttgaat aagca	4555

<210> 11

<211> 2480

<212> DNA

<213> Homo sapiens

<220>

<223> >BRAF|ENSG00000157764|ENST00000288602|2480

<400> 11

cgctccctt cccctcccc gcccgacagc ggccgctcgg gcccggctc tcggttataa	60
gatggcggcg ctgagcgggtg gcggtggtgg cggcgcgag cggggccagg ctctgttcaa	120
cggggacatg gagcccgagg ccggcgccgg cgccggcgcc gcggcctctt cggctgcgga	180
ccctgccatt ccggaggagg tgtggaatat caaacaatg attaagttga cacaggaaca	240
tatagaggcc ctattggaca aatttggtgg ggagcataat ccaccatcaa tatatctgga	300
ggcctatgaa gaatacacca gcaagctaga tgcactcaa caaagagaac aacagttatt	360
ggaatctctg gggaacggaa ctgatttttc tgtttctagc tctgcatcaa tggataccgt	420
tacatcttct tcctcttcta gcctttcagt gctaccttca tctctttcag tttttcaaaa	480
tcccacagat gtggcacgga gcaaccccaa gtcaccacaa aaacctatcg ttagagtctt	540
cctgcccac aaacagagga cagtgtacc tgcaaggtgt ggagttacag tccgagacag	600
tctaaagaaa gactgatga tgagaggtct aatcccagag tgctgtgctg ttacagaat	660
tcaggatgga gagaagaaac caattggttg ggacactgat atttcctggc ttactggaga	720
agaattgcat gtggaagtgt tggagaatgt tccacttaca acacacaact ttgtacgaaa	780
aacgtttttc acctagcat tttgtgactt ttgtcgaaag ctgcttttcc agggtttccg	840
ctgtcaaaca tgtggttata aatttcacca gcgttgtagt acagaagttc cactgatgtg	900
tgtaattat gaccaacttg atttgctgtt tgtctccaag ttctttgaac accaccaat	960
accacaggaa gaggcgtcct tagcagagac tgccctaaca tctggatcat ccccttcgc	1020
acccgcctcg gactctattg ggcccaaat tctcaccagt ccgtctcctt caaatccat	1080
tccaattcca cagcccttcc gaccagcaga tgaagatcat cgaaatcaat ttgggcaacg	1140
agaccgatcc tcatcagctc ccaatgtgca tataaacaca atagaacctg tcaatattga	1200
tgacttgatt agagaccaag gatttcgtgg tgatggagga tcaaccacag gtttgtctgc	1260
tacccccct gcctcattac ctggctcact aactaacgtg aaagccttac agaaatctcc	1320
aggacctcag cgagaaagga agtcatcttc atcctcagaa gacaggaatc gaatgaaaac	1380
acttggtaga cgggactcga gtgatgattg ggagattcct gatgggcaga ttacagtggg	1440
acaaagaatt ggatctggat catttggaac agtctacaag ggaaagtggc atggtgatgt	1500

ggcagtga	atgttgaatg	tgacagcacc	tacacctcag	cagttacaag	ccttcaaaaa	1560
tgaagtagga	gtactcagga	aaacacgaca	tgtgaatatc	ctactcttca	tgggctattc	1620
cacaaagcca	caactggcta	ttgttaccca	gtggtgtgag	ggctccagct	tgtatcacca	1680
tctccatatc	attgagacca	aatttgagat	gatcaaactt	atagatat	cacgacagac	1740
tgcacagggc	atggattact	tacacgccaa	gtcaatcatc	cacagagacc	tcaagagtaa	1800
taatataatt	cttcatgaag	acctcacagt	aaaaataggt	gattttggtc	tagctacagt	1860
gaaatctcga	tggagtgggt	cccatcagtt	tgaacagttg	tctggatcca	ttttgtggat	1920
ggcaccagaa	gtcatcagaa	tgcaagataa	aatccatac	agctttcagt	cagatgtata	1980
tgcatttgga	attgttctgt	atgaattgat	gactggacag	ttaccttatt	caaacatcaa	2040
caacagggac	cagataattt	ttatgggtggg	acgaggatac	ctgtctccag	atctcagtaa	2100
ggtacggagt	aactgtccaa	aagccatgaa	gagattaatg	gcagagtgcc	tcaaaaagaa	2160
aagagatgag	agaccactct	ttccccaat	tctcgcctct	attgagctgc	tggcccgctc	2220
attgccaaaa	attcaccgca	gtgcatcaga	accctccttg	aatcgggctg	gtttccaaac	2280
agaggatttt	agtctatatg	cttgtgcttc	tccaaaaaca	cccatccagg	cagggggata	2340
tgggtgcgtt	cctgtccact	gaaacaaatg	agtgagagag	ttcaggagag	tagcaacaaa	2400
aggaaaataa	atgaacatat	gtttgcttat	atgttaaatt	gaataaaata	ctctcttttt	2460
ttttaaggtg	aaccaaagaa					2480

<210> 12

<211> 5936

<212> DNA

<213> Homo sapiens

<220>

<223> >BRCA1|ENSG0000012048|ENST00000471181|5936

<400> 12

gtaccttgat	ttcgtattct	gagaggctgc	tgcttagcgg	tagccccttg	gtttccgtgg	60
caacggaaaa	gcgcgggaat	tacagataaa	ttaaaactgc	gactgcgcgg	cgtgagctcg	120
ctgagacttc	ctggacgggg	gacaggctgt	ggggtttctc	agataactgg	gccctgcgc	180
tcaggaggcc	ttcaccctct	gctctgggta	aagttcattg	gaacagaaag	aatggattt	240
atctgctctt	cgcgttgaag	aagtacaaaa	tgtcattaat	gctatgcaga	aatcttaga	300

gtgtcccatc tgtctggagt tgatcaagga acctgtctcc acaaagtgtg accacatatt	360
ttgcaaattt tgcattgctga aactttctcaa ccagaagaaa gggccttcac agtgtccttt	420
atgtaagaat gatataacca aaaggagcct acaagaaagt acgagattta gtcaacttgt	480
tgaagagcta ttgaaaatca tttgtgcttt tcagcttgac acaggtttgg agtatgcaaa	540
cagctataat tttgcaaaaa aggaaaataa ctctcctgaa catctaaaag atgaagtttc	600
tatcatccaa agtatgggct acagaaaccg tgccaaaaga cttctacaga gtgaaccgga	660
aaatccttcc ttgcaggaaa ccagtctcag tgtccaactc tctaaccttg gaactgtgag	720
aactctgagg acaaagcagc ggatacaacc tcaaaagacg tctgtctaca ttgaattggg	780
atctgattct tctgaagata ccgttaataa ggcaacttat tgcagtgtgg gagatcaaga	840
attgttacia atcacccctc aaggaaccag ggatgaaatc agtttggatt ctgcaaaaaa	900
ggctgcttgt gaattttctg agacggatgt aacaaatact gaacatcatc aaccagtaa	960
taatgatttg aacaccactg agaagcgtgc agctgagagg catccagaaa agtatcaggg	1020
tagttctgtt tcaaacttgc atgtggagcc atgtggcaca aatactcatg ccagctcatt	1080
acagcatgag aacagcagtt tattactcac taaagacaga atgaatgtag aaaaggctga	1140
attctgtaat aaaagcaaac agcctggctt agcaaggagc caacataaca gatgggctgg	1200
aagtaaggaa acatgtaatg ataggcggac tcccagcaca gaaaaaaagg tagatctgaa	1260
tgctgatccc ctgtgtgaga gaaaagaatg gaataagcag aaactgccat gctcagagaa	1320
tcctagagat actgaagatg ttccttggat aacactaaat agcagcattc agaaagttaa	1380
tgagtggttt tccagaagtg atgaactgtt aggttctgat gactcacatg atggggagtc	1440
tgaatcaaat gccaaagtag ctgatgtatt ggacgttcta aatgaggtag atgaatattc	1500
tggttcttca gagaaaatag acttactggc cagtgatcct catgaggctt taatatgtaa	1560
aagtgaaga gttcactcca aatcagtaga gagtaatatt gaagacaaaa tatttgggaa	1620
aacctatcgg aagaaggcaa gcctcccaa cttaagccat gtaactgaaa atctaattat	1680
aggagcattt gttactgagc cacagataat acaagagcgt cccctcacia ataaattaaa	1740
gcgtaaaagg agacctacat caggccttca tcctgaggat tttatcaaga aagcagattt	1800
ggcagttcaa aagactcctg aaatgataaa tcagggaact aaccaaacgg agcagaatgg	1860
tcaagtgatg aatattacta atagtgggtca tgagaataaa aaaaaagtg attctattca	1920

gaatgagaaa aatcctaacc caatagaatc actcgaaaaa gaatctgctt tcaaaacgaa	1980
agctgaacct ataagcagca gtataagcaa tatggaactc gaattaaata tccacaattc	2040
aaaagcacct aaaaagaata ggctgaggag gaagtcttct accaggcata ttcatgcgct	2100
tgaactagta gtcagtagaa atctaagccc acctaattgt actgaattgc aaattgatag	2160
ttgttctagc agtgaagaga taaagaaaaa aaagtacaac caaatgccag tcaggcacag	2220
cagaaaccta caactcatgg aaggtaaaga acctgcaact ggagccaaga agagtaacaa	2280
gccaaatgaa cagacaagta aaagacatga cagcgatact ttcccagagc tgaagttaac	2340
aaatgcacct ggttctttta ctaagtgttc aaataccagt gaacttaaag aatttgtcaa	2400
tcctagcctt ccaagagaag aaaaagaaga gaaactagaa acagttaaag tgtctaataa	2460
tgctgaagac cccaaagatc tcatgttaag tggagaaagg gttttgcaaa ctgaaagatc	2520
tgtagagagt agcagtatct cattgggtacc tgggtactgat tatggcactc aggaaagtat	2580
ctcgttactg gaagttagca ctctagggaa ggcaaaaaca gaaccaaata aatgtgtgag	2640
tcagtgtgca gcatttgaaa accccaaggg actaattcat gggtgtttcca aagataatag	2700
aaatgacaca gaaggcttta agtatccatt gggacatgaa gttaaccaca gtcgggaaac	2760
aagcatagaa atggaagaaa gtgaacttga tgctcagtat ttgcagaata cattcaaggt	2820
ttcaaagcgc cagtcatttg ctccgttttc aaatccagga aatgcagaag aggaatgtgc	2880
aacatttctt gccactctg ggtccttaaa gaaacaaagt ccaaaagtca cttttgaatg	2940
tgaacaaaag gaagaaaatc aaggaaagaa tgagtctaata atcaagcctg tacagacagt	3000
taatatcact gcaggctttc ctgtggttg tcaaaaagat aagccagttg ataatgccaa	3060
atgtagtata aaaggaggct ctaggttttg tctatcatct cagttcagag gcaacgaaac	3120
tggactcatt actccaaata aacatggact tttaaaaaac ccatatcgta taccaccact	3180
ttttcccatc aagtcatttg ttaaaactaa atgtaagaaa aatctgctag aggaaaactt	3240
tgaggaacat tcaatgtcac ctgaaagaga aatgggaaat gagaacattc caagtacagt	3300
gagcacaatt agccgtaata acattagaga aaatgttttt aaagaagcca gctcaagcaa	3360
tattaatgaa gtaggttcca gtactaatga agtgggctcc agtattaatg aaataggttc	3420
cagtgatgaa aacattcaag cagaactagg tagaaacaga gggccaaaat tgaatgctat	3480
gcttagatta ggggttttgc aacctgaggt ctataaacia agtcttcctg gaagtaattg	3540

taagcatcct gaaataaaaa agcaagaata tgaagaagta gttcagactg ttaatacaga	3600
tttctctcca tatctgattt cagataactt agaacagcct atgggaagta gtcatgcatc	3660
tcaggtttgt tctgagacac ctgatgacct gttagatgat ggtgaaataa aggaagatac	3720
tagttttgct gaaaatgaca ttaaggaaag ttctgctgtt tttagcaaaa gcgtccagaa	3780
aggagagctt agcaggagtc ctagcccttt caccataca catttggtc agggttacg	3840
aagaggggcc aagaaattag agtcctcaga agagaactta tctagtgagg atgaagagct	3900
tccctgcttc caacacttgt tatttggtta agtaaacaat ataccttctc agtctactag	3960
gcatagcacc gttgctaccg agtgtctgtc taagaacaca gaggagaatt tattatcatt	4020
gaagaatagc ttaaatgact gcagtaacca ggtaatatg gcaaaggcat ctcaggaaca	4080
tcaccttagt gaggaaacaa aatgttctgc tagcttgttt tcttcacagt gcagtgaatt	4140
ggaagacttg actgcaaata caaacacca ggatcctttc ttgattgggtt cttccaaaca	4200
aatgaggcat cagtctgaaa gccagggagt tggctctgagt gacaaggaat tggtttcaga	4260
tgatgaagaa agaggaacgg gcttggaaga aaataatcaa gaagagcaaa gcatggattc	4320
aaacttaggt gaagcagcat ctgggtgtga gagtgaacaa agcgtctctg aagactgctc	4380
agggctatcc tctcagagt acattttaac cactcagcag agggatacca tgcaacataa	4440
cctgataaag ctccagcagg aaatggctga actagaagct gtgttagaac agcatgggag	4500
ccagccttct aacagctacc cttccatcat aagtgactct tctgcccttg aggacctgcg	4560
aaatccagaa caaagcacat cagaaaaaga ttcgcatata catggccaaa ggaacaactc	4620
catgttttct aaaaggccta gagaacatat atcagtatta acttcacaga aaagtagtga	4680
ataccctata agccagaatc cagaaggcct ttctgctgac aagtttgagg tgtctgcaga	4740
tagttctacc agtaaaaaata aagaaccagg agtggaaagg tcatcccctt ctaaagcccc	4800
atcattagat gataggtggt acatgcacag ttgctctggg agtcttcaga atagaaacta	4860
cccatctcaa gaggagctca ttaaggttgt tgatgtggag gagcaacagc tggaagagtc	4920
tgggccacac gatttgacgg aaacatctta cttgccaagg caagatctag aggggaacccc	4980
ttacctggaa tctggaatca gcctcttctc tgatgacctt gaatctgatc cttctgaaga	5040
cagagcccca gagtacgctc gtgttggaac cataccatct tcaacctctg cattgaaagt	5100
tccccaattg aaagttgcag aatctgcca gagtccagct gctgctcata ctactgatac	5160

tgctgggtat aatgcaatgg aagaaagtgt gagcagggag aagccagaat tgacagcttc	5220
aacagaaagg gtcaacaaaa gaatgtccat ggtgggtgtct ggcctgaccc cagaagaatt	5280
tatgctcgtg tacaagtttg ccagaaaaca ccacatcact ttaactaatc taattactga	5340
agagactact catgtttgta tgaaaacaga tgctgagttt gtgtgtgaac ggacactgaa	5400
atattttcta ggaattgcgg gaggaaaatg ggtagttagc tatttctggg tgacccagtc	5460
tattaaagaa agaaaaatgc tgaatgagca tgattttgaa gtcagaggag atgtgggtcaa	5520
tggaagaaac caccaaggtc caaagcgagc aagagaatcc caggacagaa agatcttcag	5580
ggggctagaa atctgttgct atgggccctt caccaacatg cccacagatc aactggaatg	5640
gatggtacag ctgtgtgggt cttctgtggg gaaggagctt tcatcattca cccttggcac	5700
aggtgtccac ccaattgtgg ttgtgcagcc agatgcctgg acagaggaca atggcttcca	5760
tgcaattggg cagatgtgtg aggcacctgt ggtgacccga gagtgggtgt tggacagtgt	5820
agcactctac cagtgccagg agctggacac ctacctgata cccagatcc cccacagcca	5880
ctactgactg cagccagcca caggtacaga gccacaggac cccaagaatg agctta	5936

<210> 13
 <211> 10984
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >BRCA2|ENSG00000139618|ENST00000544455|10984

<400> 13	
gtggcgcgag cttctgaaac taggcggcag aggcggagcc gctgtggcac tgctgcgcct	60
ctgctgcgcc tcgggtgtct tttgcggcgg tgggtcgccg ccgggagaag cgtgagggga	120
cagatttgtg accggcgcgg tttttgtcag cttactccgg ccaaaaaaga actgcacctc	180
tggagcggac ttatttacca agcattggag gaatatcgta ggtaaaaatg cctattggat	240
ccaaagagag gccaacattt tttgaaattt ttaagacacg ctgcaacaaa gcagatttag	300
gaccaataag tcttaattgg tttgaagaac tttcttcaga agctccaccc tataattctg	360
aacctgcaga agaattctgaa cataaaaaaca acaattacga accaaaccta tttaaaactc	420
cacaaaggaa accatcttat aatcagctgg cttcaactcc aataatattc aaagagcaag	480

ggctgactct gccgctgtac caatctcctg taaaagaatt agataaattc aaattagact	540
taggaaggaa tgttcccaat agtagacata aaagtcttcg cacagtgaaa actaaaatgg	600
atcaagcaga tgatgtttcc tgtccacttc taaattcttg tcttagtgaa agtcctgttg	660
ttctacaatg tacacatgta acaccacaaa gagataagtc agtggtatgt gggagtttgt	720
ttcatacacc aaagtttgtg aagggtcgtc agacaccaaa acatatttct gaaagtctag	780
gagctgaggt ggatcctgat atgtcttggt caagttcttt agctacacca cccaccctta	840
gttctactgt gtcatagtc agaaatgaag aagcatctga aactgtattt cctcatgata	900
ctactgctaa tgtgaaaagc tatttttcca atcatgatga aagtcctgaag aaaaatgata	960
gattttatcgc ttctgtgaca gacagtgaaa acacaaatca aagagaagct gcaagtcatg	1020
gatttggaag aacatcaggg aattcattta aagtaaataag ctgcaaagac cacattggaa	1080
agtcaatgcc aaatgtccta gaagatgaag tatatgaaac agttgtagat acctctgaag	1140
aagatagttt ttcattatgt ttttctaaat gtagaacaaa aaatctacaa aaagtaagaa	1200
ctagcaagac taggaaaaaa attttccatg aagcaaacgc tgatgaatgt gaaaaatcta	1260
aaaaccaagt gaaagaaaaa tactcatttg tatctgaagt ggaaccaaat gatactgac	1320
cattagattc aaatgtagca aatcagaagc ctttgagag tggaagtgaac aaaatctcca	1380
aggaagttgt accgtctttg gcctgtgaat ggtctcaact aaccctttca ggtctaaatg	1440
gagcccagat ggagaaaata cccctattgc atatttcttc atgtgaccaa aatatttcag	1500
aaaaagacct attagacaca gagaacaaaa gaaagaaaga ttttcttact tcagagaatt	1560
ctttgccacg tatttctagc ctacaaaaat cagagaagcc attaaatgag gaaacagtgg	1620
taaataagag agatgaagag cagcatcttg aatctcatac agactgcatt cttgcagtaa	1680
agcaggcaat atctggaact tctccagtgg cttcttcatt tcagggtatc aaaaagtcta	1740
tattcagaat aagagaatca cctaaagaga ctttcaatgc aagtttttca ggtcatatga	1800
ctgatccaaa ctttaaaaaa gaaactgaag cctctgaaag tggactggaa atacatactg	1860
tttgctcaca gaaggaggac tccttatgtc caaatttaat tgataatgga agctggccag	1920
ccaccaccac acagaattct gtagctttga agaatgcagg tttaatatcc actttgaaaa	1980
agaaaacaaa taagtttatt tatgctatac atgatgaaac atcttataaa ggaaaaaaaa	2040
taccgaaaga ccaaaaatca gaactaatta actgttcagc ccagtttgaa gcaaatgctt	2100

ttgaagcacc	acttacat	ttt gcaa	atgctg	attcaggt	ttt attgc	atttct	tctgtg	aaaa	2160
gaagctgttc	acagaatgat	tctgaaga	ac	actttgtc	cttaactagc	tctttt	ggga		2220
caattctgag	gaaatgttct	agaaatgaaa	catgttctaa	taatacagta	atctctcagg				2280
atcttgatta	taaagaagca	aaatgtaata	aggaaaaact	acagttat	ttt attac	cccgag			2340
aagctgattc	tctgtcatgc	ctgcaggaag	gacagtgtga	aaatgatcca	aaaagcaaaa				2400
aagtttcaga	tataaaagaa	gaggtcttgg	ctgcagcatg	tcaccagta	caacattcaa				2460
aagtgggaata	cagtgatact	gactttcaat	cccagaaaag	tcttttatat	gatcatgaaa				2520
atgccagcac	tcttatttta	actcctactt	ccaaggatgt	tctgtcaaac	ctagtcatga				2580
tttctagagg	caaagaatca	tacaaaatgt	cagacaagct	caaaggtaac	aattatgaat				2640
ctgatgttga	attaaccaa	aatattccca	tggaaaagaa	tcaagatgta	tgtgctttaa				2700
atgaaaatta	taaaaacgtt	gagctgttgc	cacctgaaaa	atacatgaga	gtagcatcac				2760
cttcaagaaa	ggtacaattc	aaccaaaca	caaatactag	agtaatcca	aaaaatcaag				2820
aagaaactac	ttcaatttca	aaaataactg	tcaatccaga	ctctgaagaa	cttttctcag				2880
acaatgagaa	taattttgtc	ttccaagtag	ctaatagaag	gaataatctt	gctttaggaa				2940
atactaagga	acttcatgaa	acagacttga	cttgtgtaaa	cgaaccatt	ttcaagaact				3000
ctaccatggg	tttatatgga	gacacaggtg	ataaacaagc	aaccaagt	gtcaattaaaa				3060
aagatttggg	ttatgttctt	gcagaggaga	acaaaaatag	tgtaaagcag	catataaaaa				3120
tgactctagg	tcaagattta	aaatcggaca	tctccttgaa	tatagataaa	ataccagaaa				3180
aaaataatga	ttacatgaac	aaatgggcag	gactcttagg	tccaatttca	aatcacagtt				3240
ttggaggtag	cttcagaaca	gcttcaaata	aggaaatcaa	gctctctgaa	cataacatta				3300
agaagagcaa	aatgttcttc	aaagatat	tg aaga	acaata	tcctactagt	ttagctt	gtg		3360
ttgaaattgt	aaataccttg	gcattagata	atcaaaagaa	actgagcaag	cctcagtcaa				3420
ttaatactgt	atctgcacat	ttacagagta	gtgtagttgt	ttctgattgt	aaaaatagtc				3480
atataacccc	tcagatgtta	ttttccaagc	aggattttta	ttcaaaccat	aatttaacac				3540
ctagccaaaa	ggcagaaatt	acagaacttt	ctactatatt	agaagaatca	ggaagtcagt				3600
ttgaatttac	tcagtttaga	aaaccaagct	acatat	tgca	gaagagtaca	tttgaagtgc			3660
ctgaaaacca	gatgactatc	ttaaagacca	cttctgagga	atgcagagat	gctgatcttc				3720

atgtcataat gaatgcccc	tcgattggtc aggtagacag	cagcaagcaa tttgaaggta	3780
cagttgaaat taaacggaag	tttgctggcc tgttgaaaa	tgactgtaac aaaagtgctt	3840
ctggttattt aacagatgaa	aatgaagtgg ggtttagggg	cttttattct gctcatggca	3900
caaaactgaa tgtttctact	gaagctctgc aaaaagctgt	gaaactgttt agtgatattg	3960
agaatattag tgaggaaact	tctgcagagg tacatccaat	aagtttatct tcaagtaa	4020
gtcatgattc tgttgtttca	atgtttaaga tagaaaatca	taatgataaa actgtaagt	4080
aaaaaaataa taaatgccaa	ctgatattac aaaataatat	tgaaatgact actggcactt	4140
ttgttgaaga aattactgaa	aattacaaga gaaatactga	aaatgaagat aacaaatata	4200
ctgctgccag tagaaattct	cataacttag aatttgatgg	cagtgattca agtaaaaatg	4260
atactgtttg tattcataaa	gatgaaacgg acttgctatt	tactgatcag cacaacatat	4320
gtcttaaaatt atctggccag	tttatgaagg agggaaacac	tcagattaaa gaagatttgt	4380
cagatttaac ttttttgga	gttgcgaaag ctcaagaagc	atgtcatggt aatacttcaa	4440
ataaagaaca gttaactgct	actaaaacgg agcaaaatat	aaaagatttt gagacttctg	4500
atacattttt tcagactgca	agtgggaaaa atattagtgt	cgccaaagag tcatttaata	4560
aaattgtaaa tttctttgat	cagaaaccag aagaattgca	taacttttcc ttaaattctg	4620
aattacattc tgacataaga	aagaacaaaa tggacattct	aagttatgag gaaacagaca	4680
tagttaaaca caaaatactg	aaagaaagtg tcccagttgg	tactggaaat caactagtga	4740
ccttccaggg acaacccgaa	cgtgatgaaa agatcaaaga	acctactcta ttgggttttc	4800
atacagctag cgggaaaaaa	gttaaaattg caaaggaatc	tttgacaaa gtgaaaaacc	4860
tttttgatga aaaagagcaa	ggtactagtg aaatcaccag	ttttagccat caatgggcaa	4920
agaccctaaa gtacagagag	gcctgtaaag acctgaatt	agcatgtgag accattgaga	4980
tcacagctgc cccaaagtgt	aaagaaatgc agaattctct	caataatgat aaaaacctg	5040
tttctattga gactgtggtg	ccacctaagc tcttaagtga	taatttatgt agacaaactg	5100
aaaatctcaa aacatcaaaa	agtatctttt tgaaagttaa	agtacatgaa aatgtagaaa	5160
aagaaacagc aaaaagtcct	gcaacttggt acacaaatca	gtccccttat tcagtcattg	5220
aaaattcagc cttagctttt	tacacaagtt gtagtagaaa	aacttctgtg agtcagactt	5280
cattacttga agcaaaaaaa	tggcttagag aaggaatatt	tgatggtcaa ccagaaagaa	5340

taaatactgc agattatgta ggaaattatt tgtatgaaaa taattcaaac agtactatag	5400
ctgaaaatga caaaaatcat ctctccgaaa aacaagatac ttattttaagt aacagtagca	5460
tgtctaacag ctattcctac cattctgatg aggtatataa tgattcagga tatctctcaa	5520
aaaataaact tgattctggt attgagccag tattgaagaa tgttgaagat caaaaaaaca	5580
ctagtttttc caaagtaata tccaatgtaa aagatgcaaa tgcataccca caaactgtaa	5640
atgaagatat ttgcgttgag gaacttgtga ctagctcttc accctgcaaa aataaaaaatg	5700
cagccattaa attgtccata tctaatagta ataattttga ggtagggcca cctgcattta	5760
ggatagccag tggtaaaatc gtttgtgttt cacatgaaac aattaaaaaa gtgaaagaca	5820
tatttacaga cagtttcagt aaagtaatta aggaaaacaa cgagaataaa tcaaaaattt	5880
gccaaacgaa aattatggca ggttgttacg aggcatggga tgattcagag gatattcttc	5940
ataactctct agataatgat gaatgtagca cgcattcaca taagggtttt gctgacattc	6000
agagtgaaga aattttacaa cataaccaa atatgtctgg attggagaaa gtttctaaaa	6060
tatcaccttg tgatgttagt ttggaaactt cagatatatg taaatgtagt atagggaagc	6120
ttcataagtc agtctcatct gcaaatactt gtgggatttt tagcacagca agtggaaaat	6180
ctgtccaggt atcagatgct tcattacaaa acgcaagaca agtgttttct gaaatagaag	6240
atagtaccaa gcaagtcttt tccaaagtat tgtttaaaag taacgaacat tcagaccagc	6300
tcacaagaga agaaaatact gctatacgta ctccagaaca tttaatatcc caaaaaggct	6360
tttcatataa tgtggtaaat tcatctgctt tctctggatt tagtacagca agtggaaagc	6420
aagtttccat tttagaaagt tccttacaca aagttaaggg agtgtttagag gaatttgatt	6480
taatcagaac tgagcatagt cttcactatt cacctacgtc tagacaaaat gtatcaaaaa	6540
tacttcctcg tgttgataag agaaaccag agcactgtgt aaactcagaa atggaaaaaa	6600
cctgcagtaa agaatttaaa ttatcaaata acttaaatgt tgaaggtggt tcttcagaaa	6660
ataatcactc tattaagtt tctccatatc tctctcaatt tcaacaagac aaacaacagt	6720
tggtattagg aaccaaagtg tcacttgttg agaacattca tgttttggga aaagaacagg	6780
cttcacctaa aaacgtaaaa atggaaattg gtaaaactga aactttttct gatgttcctg	6840
tgaaaacaaa tatagaagtt tgttctactt actccaaaga ttcagaaaac tactttgaaa	6900
cagaagcagt agaaattgct aaagctttta tggaagatga tgaactgaca gattctaaac	6960

tgccaagtca tgccacacat tctcttttta catgtcccga aaatgaggaa atggttttgt	7020
caaattcaag aattggaaaa agaagaggag agccccttat cttagtggga gaaccctcaa	7080
tcaaaagaaa cttattaaat gaatttgaca ggataataga aaatcaagaa aaatccttaa	7140
aggcttcaaa aagcactcca gatggcacia taaaagatcg aagattgttt atgcatcatg	7200
tttctttaga gccgattacc tgtgtaccct ttcgcacaac taaggaacgt caagagatac	7260
agaatccaaa ttttaccgca cctgggtcaag aatttctgtc taaatctcat ttgtatgaac	7320
atctgacttt ggaaaaatct tcaagcaatt tagcagtttc aggacatcca ttttatcaag	7380
tttctgttac aagaaatgaa aaaatgagac acttgattac tacaggcaga ccaaccaaag	7440
tctttgttcc accttttaaa actaaatcac attttcacag agttgaacag tgtgttagga	7500
atattaactt ggaggaaaac agacaaaagc aaaacattga tggacatggc tctgatgata	7560
gtaaaaataa gattaatgac aatgagattc atcagtttaa caaaaacaac tccaatcaag	7620
cagtagctgt aactttcaca aagtgtgaag aagaaccttt agatttaatt acaagtcttc	7680
agaatgccag agatatacag gatatgcgaa ttaagaagaa acaaaggcaa cgcgtctttc	7740
cacagccagg cagtctgtat cttgcaaaaa catccactct gcctcgaatc tctctgaaag	7800
cagcagtagg aggccaagtt ccctctgcgt gttctcataa acagctgtat acgtatggcg	7860
tttctaaaca ttgcataaaa attaacagca aaaatgcaga gtcttttcag tttcacactg	7920
aagattatth ttgtaaggaa agtttatgga ctggaaaagg aatacagttg gctgatgggtg	7980
gatggctcat accctccaat gatggaaagg ctggaaaaga agaattttat agggctctgt	8040
gtgacactcc aggtgtggat ccaaagctta tttctagaat ttgggtttat aatcactata	8100
gatggatcat atggaaactg gcagctatgg aatgtgcctt tcctaaggaa tttgctaata	8160
gatgcctaag cccagaaagg gtgcttcttc aactaaaata cagatatgat acggaaattg	8220
atagaagcag aagatcggct ataaaaaaga taatggaaag ggatgacaca gctgcaaaaa	8280
cacttgttct ctgtgtttct gacataatth cattgagcgc aaatatatct gaaacttcta	8340
gcaataaaaac tagtagtgca gatacccaa aagtggccat tattgaactt acagatgggt	8400
ggtatgctgt taaggcccag ttagatcctc ccctcttagc tgtcttaaag aatggcagac	8460
tgacagttgg tcagaagatt attcttcatg gagcagaact ggtgggctct cctgatgcct	8520
gtacacctct tgaagcccca gaatctctta tgttaaagat ttctgctaac agtactcggc	8580

ctgctcgctg gtataccaaa cttggattct ttcctgaccc tagacctttt cctctgcctt	8640
tatcatcgct tttcagtgat ggaggaaatg ttggttgtgt tgatgtaatt attcaaagag	8700
cataccctat acagtggatg gagaagacat catctggatt atacatattt cgcaatgaaa	8760
gagaggaaga aaaggaagca gcaaaatatg tggaggccca acaaaagaga ctagaagcct	8820
tattcactaa aattcaggag gaatttgaag aacatgaaga aaacacaaca aaacctatatt	8880
taccatcacg tgcactaaca agacagcaag ttcgtgcttt gcaagatggt gcagagcttt	8940
atgaagcagt gaagaatgca gcagaccag cttaccttga gggttatttc agtgaagagc	9000
agttaagagc cttgaataat cacaggcaaa tgttgaatga taagaaacaa gctcagatcc	9060
agttggaaat taggaaggcc atggaatctg ctgaacaaaa ggaacaaggt ttatcaaggg	9120
atgtcacaac cgtgtggaag ttgcgtattg taagctattc aaaaaaagaa aaagattcag	9180
ttatactgag tatttggcgt ccatcatcag atttatattc tctgttaaca gaaggaaaga	9240
gatacagaat ttatcatctt gcaacttcaa aatctaaaag taaatctgaa agagctaaca	9300
tacagttagc agcgacaaaa aaaactcagt atcaacaact accggtttca gatgaaattt	9360
tatttcagat ttaccagcca cgggagcccc ttcacttcag caaattttta gatccagact	9420
ttcagccatc ttgttctgag gtggacctaa taggatttgt cgtttctgtt gtgaaaaaaa	9480
caggacttgc ccttttcgtc tatttgtcag acgaatgtta caatttactg gcaataaagt	9540
tttggataga ccttaatgag gacattatta agcctcatat gttaattgct gcaagcaacc	9600
tccagtggcg accagaatcc aaatcaggcc ttcttacttt atttgctgga gatTTTTctg	9660
tgTTTTctgc tagtccaaaa gagggccact ttcaagagac attcaacaaa atgaaaaata	9720
ctgttgagaa tattgacata ctttgcaatg aagcagaaaa caagcttatg catatactgc	9780
atgcaaatga tccaagtgg tccaccccaa ctaaagactg tacttcaggg ccgtacactg	9840
ctcaaatcat tcctggtaca ggaaacaagc ttctgatgtc ttctcctaatt tgtgagatat	9900
attatcaaag tcctttatca ctttgtatgg ccaaaaggaa gtctgtttcc acacctgtct	9960
cagcccagat gacttcaaag tcttgtaaag gggagaaaaga gattgatgac caaaagaact	10020
gcaaaaagag aagagccttg gatttcttga gtagactgcc ttacctcca cctgttagtc	10080
ccatttgtac atttgtttct ccggctgcac agaaggcatt tcagccacca aggagtgtg	10140
gcaccaaata cgaaacaccc ataaagaaaa aagaactgaa ttctcctcag atgactccat	10200

ttaaaaaatt caatgaaatt tctcttttgg aaagtaattc aatagctgac gaagaacttg	10260
cattgataaa tacccaagct cttttgtctg gttcaacagg agaaaaacaa tttatatctg	10320
tcagtgaatc cactaggact gctcccacca gttcagaaga ttatctcaga ctgaaacgac	10380
gttggtactac atctctgac aaagaacagg agagttccca ggccagtagc gaagaatgtg	10440
agaaaaataa gcaggacaca attacaacta aaaaatatat ctaagcattt gcaaaggcga	10500
caataaatta ttgacgctta acctttccag tttataagac tggaatataa tttcaaacca	10560
cacattagta cttatgttgc acaatgagaa aagaaattag tttcaaattt acctcagcgt	10620
ttgtgtatcg ggcaaaaatc gttttgcccg attccgtatt ggtatacttt tgcttcagtt	10680
gcatacttta aaactaaatg taatttatta actaatcaag aaaaacatct ttggctgagc	10740
tcggtggctc atgcctgtaa tcccaacact ttgagaagct gaggtgggag gaggcttga	10800
ggccaggagt tcaagaccag cctgggcaac ataggagagac ccccatcttt acaaagaaaa	10860
aaaaaagggg aaaagaaaat cttttaaatc tttggatttg atcactacaa gtattatctt	10920
acaatcaaca aaatggcat ccaaactcaa acttgagaaa atatcttgct ttcaaattgg	10980
cact	10984

<210> 14
 <211> 6048
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >BRIP1|ENSG00000136492|ENST00000259008|6048

<400> 14	
aattcgtctc gggttgtgtg gttgaggggt ctggtgggtc gaggaaggt aacggcggcc	60
ccagtcctgc acacaaggcc ggggaagtag cagcaccccc aggaagaggg aggaggaagg	120
gctcgtgccc tttcttctct tccagggtc cgctttatct gctctcagaa gtcggtttcc	180
tttccttttc ttcagtgaat cggagctcag agcgttgctt cggtttccct ccagacagtt	240
aggaatctga aataaacagg aaagcactat gtcttcaatg tggctctgaat atacaattgg	300
tgggggtgaag atttactttc cttataaagc ttaccgtca cagcttgcta tgatgaattc	360
tattctcaga ggattaaaca gcaagcaaca ttgtttgttg gagagtcca caggaagtgg	420
aaaaagctta gccttacttt gttctgcttt agcatggcaa caatctctta gtgggaaacc	480

agcagatgag ggcgtaagtg aaaaagctga agtacaattg tcatgtttgtt gtgcatgcca	540
ttcaaaggat tttacaaaca atgacatgaa ccaaggaact tcacgtcatt tcaactatcc	600
aagcacacca ctttctgaaa gaaatggcac ttcatcaact tgtcaagact cccctgaaaa	660
aaccactctg gctgcaaagt tatctgctaa gaaacaggca tccatataca gagatgaaaa	720
tgatgatttt caagtagaga agaaaagaat tcgaccctta gaaactacac agcagattag	780
aaaacgtcat tgctttggaa cagaagtaca caatttggat gcaaaagttg attcaggaaa	840
gactgtaaaa ctcaactctc cactggaaaa gataaactcc ttttcgccac agaaaccccc	900
tggccactgt tctaggtgct gttgttctac taaacaagga aacagtcaag agtcatcgaa	960
taccattaag aaggatcata cagggaatc caagataccc aaaatatatt ttgggacacg	1020
cacacacaag cagattgctc agattactag agagctccgg aggacggcat attcaggggt	1080
tccaatgact attctttcca gcagggatca tacttgtgtc catcctgagg tagtcggtaa	1140
cttcaacaga aatgagaagt gcatggaatt gctagatggg aaaaacggaa aatcctgcta	1200
tttttatcat ggagttcata aaattagtga tcagcacaca ttacagactt tccaagggat	1260
gtgcaaagcc tgggatatag aagaacttgt cagcctgggg aagaaactaa aggcctgtcc	1320
atattacaca gcccgagaac taatacaaga tgctgacatc atattttgtc cctacaacta	1380
tcttctagat gcacaaataa gggaaagtat ggattttaat ctgaaagaac aggttgtcat	1440
tttagatgaa gctcataaca tcgaggactg tgctcgggaa tcagcaagtt acagtgtaac	1500
agaagttcag cttcggtttg ctcgggatga actagatagt atggtcaaca ataataaag	1560
gaagaaagat catgaacccc tacgagctgt gtgctgtagc ctcattaatt ggtagaagc	1620
aaacgctgaa tatcttgtag aaagagatta tgaatcagct tgtaaaatat ggagtggaaa	1680
tgaaatgctc ttaactttac aaaaaatggg tatcaccact gctacttttc ccattttgca	1740
gggacatttt tctgctgttc ttcaaaaaga ggaaaaaatc tcaccaattt atggtaaaga	1800
ggaggcaaga gaagtacctg ttattagtgc atcaactcaa ataatgctta aaggactttt	1860
tatggtactt gactatcttt ttaggcaaaa tagcagattt gcagatgatt ataaaattgc	1920
gattcaacag acttactcct ggacaaatca gattgatatt tcagacaaaa atgggttggt	1980
ggttctacca aaaaataaga aacgttcacg acagaaaact gcagttcatg tgctaaactt	2040
ttggtgctta aatccagctg tggccttttc agatattaat ggcaaagttc agaccattgt	2100

tttgacatct ggtacattat caccaatgaa atccttttcg tcagaacttg gtgttacatt	2160
tactatccag ctggaggcta atcatatcat taaaaattca caggtttggg ttggtacat	2220
tgggtcaggc cccaagggtc ggaatctctg tgctaccttc cagaatactg aaacatttga	2280
gttccaagat gaagtgggag cacttttggt atctgtgtgc cagactgtga gccaaaggaat	2340
tttgtgtttc ttgccatctt acaagttatt agaaaaatta aaagaacgtt ggctctctac	2400
tggtttatgg cataatctgg agttggtgaa gacagtcatt gtagaaccac agggaggaga	2460
aaaaacaaat tttgatgaat tactgcaggt gtactatgac gcaatcaaat acaaaggaga	2520
gaaagatgga gctctcctgg tagcagtttg tcgtggtaaa gtgagtgagg gtctggattt	2580
ctcagatgac aatgcccgtg ctgtcataac aataggaatt ccttttccaa atgtgaaaga	2640
tctacagggt gaactaaaac gacaatacaa tgaccacat tcaaaattga gaggtcttct	2700
acctggccgt cagtggatat aaattcaagc atacagggcc ttaaaccagg cccttggtag	2760
atgtattaga cacagaaatg attggggagc tcttattcta gtggatgac gctttaggaa	2820
taaccaagt cgctatatat ctggactttc taaatgggta cggcagcaga ttcagcacca	2880
ttcaaccttt gaaagtgac tggaatcctt ggctgaattt tccaaaagc atcaaaaagt	2940
tcttaatgta tccataaagg acagaaccaa tatacaggac aatgagtcta cacttgaagt	3000
gacctcttta aagtacagta cctcacctta tttactggaa gcagcaagtc atctatcacc	3060
agaaaatttt gtggaagatg aagcaaagat atgtgtccag gaactacagt gtcctaaat	3120
tattacaaa aattcacctc taccaagtag cattatctcc agaaaggaga aaaatgatcc	3180
agtattcctg gaagaagcag ggaaagcaga aaaaattgtg atttccagat ccacaagccc	3240
aactttcaac aaacaaaca agagagttag ctggtcaagc ttttaattctt tgggacagta	3300
ttttactggg aaaataccga aggcaacacc tgagctcggg tcatcagaga atagtcctc	3360
tagtcctccc cgtttcaaaa cagagaagat ggaaagtaaa actgttttgc ctttactga	3420
taaatgtgaa tcctcaaadc tgacagtaaa cacatcgttt ggatcatgcc ctcaatcaga	3480
aaccattatt tcatcattaa agattgatgc cacccttact agaaaaaatc attctgaaca	3540
tccgctctgt tctgaagaag ccctggatcc agacattgaa ttgtctctag taagtgaaga	3600
agataaacag tccacttcaa atagagattt tgaaacagaa gcagaagatg aatctatcta	3660
ttttacacct gaactttatg atcctgaaga tacagatgaa gaaaaaaatg acctagctga	3720

aactgataga ggaaatagat tggctaacaa ttcagattgc attttagcta aagacctttt	3780
tgaaattaga actataaaag aagtagattc agccagagaa gtgaaagctg aggattgcat	3840
agatacaaag ttgaatggaa ttctgcatat tgaagaaagt aaaattgatg acattgatgg	3900
taatgtaaaa acaacttgga taaatgaact ggaactggga aaaactcatg aaatagaaat	3960
aaagaacttt aaaccatctc cttccaaaaa taaaggcatg tttcctggtt ttaagtaata	4020
atacttaact ctcaagctaa gtaaaaatat gtcacatgc ttatgttaaa ctctgttgta	4080
agtaataatt tgtaaattga ataagtggca tacttttttaaaa aaactatgtt tatgttcaga	4140
aatgtaaatg ttattattct tgagtttttg ggttttttttt tttgagacag agtcttggtc	4200
tgttgccag gctggaatgc agtggtgtgc tctgggctca ctgcaacctt cacctccagg	4260
ttcaagtgat tctcctgcct cagccttctg agtagctggg actacaggtg tgcaccacca	4320
tgcccagcta gtttttgtat ttttagtaga gatgggggtt caccatgttg gccaggctgg	4380
tctcgaactc ctggcctcgt gatctgccct tctctgcctc cctaagttgc tgggattaca	4440
ggtgtgagcc acagtgcctg gccattctt gagttttgat aaagtaattc atacaaagta	4500
ctgtcctcaa ataagtcttc cttagctaaa tgcaatttaa aattattcaa agatcctagg	4560
gcacttctag tttcacgtaa atattcatat taggtggttc tcttcatcca tttgttttca	4620
cactgatata taaaaattaa cagcagtcta atctagtac acctcagtca tatttcgcta	4680
tagattttac ctcaaatcag tccaagactt tttcagagat caccatttgt cttgaaaggt	4740
ttatttcgtt attaaactgc ctacttataa gtaattaaga gaaattaaga aagtagtatg	4800
catttttaat tgaaattggt ttacattctt tgtataataa acctaaaacc aaacatgtca	4860
taaacaaatt gacgtaaaga tataaaatgc caaatgaagt attccaaatt ttctattcta	4920
attattttagc ttcacatca ttgtggaaaa aaatactaga tcctgcttag tattatatat	4980
ttttcctagt ggatcagtga gtaataagta ccaaacta gactagaagg taatttctac	5040
attgtttaga aagggtgaaa caatttatcc cctctggtat tgttctagca taagctttag	5100
ttatacaatg attaagatag aaaacttcat atataaattt gataagcaaa cccacattta	5160
tagctgcagc taaaatatgt ttccttaggg cacagtaatc ctttctgtga attttgacct	5220
tgtttggtt tttgtgaatg aagctatatg tctaataaaa aatgattata aaagaggctc	5280
atctctgaca tcattccaaa aatacattca ttgatctctt ttaagaaac atctgttatt	5340

cactgggcat tgggactttt tgtgagtaat ttgaattgaa attttatgag ctatccaaga	5400
attctgtatg gtctattatt ttcaagtcaa aatttccagt aaggatttac ttacatttc	5460
atttggataa atgaatcatt atataggtat gtctttgctt ccattttgag acatttagat	5520
ttttacagcc tgtttctata gcatttgatg ttacaactct aagcgtagtt caaagacatt	5580
taaattgaca agttaccagt taaagaattt agaatatatt agatcccatc tagtattata	5640
tatttttttct agttgatcat tgagcagtaa ataccaaata ctcgattaga aggtaatttt	5700
tacattgttt tgaaagggtg aaacaattta tctcctctgg tattattctt aaaccacaga	5760
tagggatagt agggtagtga aacgaataaa tacctggtag aagacaagag acttgggctc	5820
tacacctggc tctgccactg atttgctaag tcatattggc aatcaccaca cccttcaggg	5880
aattagtttc atctgtaaaa tgcagcgggt agtactataa aatcatacaa atttctttgt	5940
gctttgagaa tctataaagg aatgtctgtt gatattctga gtcgattttc atttgctttt	6000
gttccagaac ggttaaaata aagcatatta tttcatttaa aaagtaaa	6048

<210> 15
 <211> 4307
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >CCND1|ENSG00000110092|ENST00000227507|4307

<400> 15	
gcttaacaac agtaacgtca cacggactac aggggagttt tgttgaagtt gcaaagtcct	60
ggagcctcca gagggctgtc ggcgtagtag cagcgagcag cagagtccgc acgctccggc	120
gaggggcaga agagcgcgag ggagcgcggg gcagcagaag cgagagccga gcgcggaccc	180
agccaggacc cacagccctc cccagctgcc caggaagagc cccagccatg gaacaccagc	240
tcctgtgctg cgaagtggaa accatccgcc gcgcgtaccc cgatgccaac ctctcaacg	300
accgggtgct gcgggccatg ctgaaggcgg aggagacctg cgcgccctcg gtgtcctact	360
tcaaattgtgt gcagaaggag gtcctgccgt ccatgcggaa gatcgtcgcc acctggatgc	420
tggaggtctg cgaggaacag aagtgcgagg aggaggtctt cccgctggcc atgaactacc	480
tggaccgctt cctgtcgctg gagcccgtga aaaagagccg cctgcagctg ctggggggcca	540

cttgcattgtt cgtggcctct aagatgaagg agaccatccc cctgacggcc gagaagctgt	600
gcatctacac cgacaactcc atccggcccg aggagctgct gcaaattggag ctgctcctgg	660
tgaacaagct caagtggaac ctggccgcaa tgaccccgca cgatttcatt gaacacttcc	720
tctccaaaat gccagaggcg gaggagaaca aacagatcat ccgcaaacac gcgcagacct	780
tcgttgccct ctgtgccaca gatgtgaagt tcattttcaa tccgccctcc atggtggcag	840
cggggagcgt ggtggccgca gtgcaaggcc tgaacctgag gagccccaac aacttcctgt	900
cctactaccg cctcacacgc ttcctctcca gagtgatcaa gtgtgaccg gactgcctcc	960
gggcctgcca ggagcagatc gaagccctgc tggagtcaag cctgcgccag gccagcaga	1020
acatggaccc caaggccgcc gaggaggagg aagaggagga ggaggagggtg gacctggctt	1080
gcacaccac cgacgtgcgg gacgtggaca tctgaggcg ccaggcaggc gggcgccacc	1140
gccacccgca gcgaggcggg agccggcccc aggtgctccc ctgacagtcc ctctctccg	1200
gagcattttg ataccagaag ggaaagcttc attctccttg ttgttggttg tttttcctt	1260
tgctctttcc cccttccatc tctgacttaa gcaaaagaaa aagattacc aaaaactgtc	1320
tttaaaagag agagagagaa aaaaaaata gtatttgcatt aaccctgagc ggtgggggag	1380
gagggttggtg ctacagatga tagaggattt tatacccaa taatcaactc gtttttatat	1440
taatgtactt gtttctctgt tgtaagaata ggcattaaca caaaggaggc gtctcgggag	1500
aggattaggt tccatccttt acgtgtttta aaaaaagcat aaaaacattt taaaaacata	1560
gaaaaattca gcaaaccatt tttaaagtag aagagggttt taggtagaaa aacatattct	1620
tgtgcttttc ctgataaagc acagctgtag tggggttcta ggcatctctg tactttgctt	1680
gctcatatgc atgtagtcac ttataagtc attgtatgtt attatattcc gtaggtagat	1740
gtgtaacctc ttcaccttat tcatggctga agtcacctct tggttacagt agcgtagcgt	1800
gcccgtgtgc atgtcctttg cgctgtgac caccaccca acaaaccatc cagtgacaaa	1860
ccatccagtg gaggtttgtc gggcaccagc cagcgtagca gggtcgggaa aggccacctg	1920
tcccactcct acgatacgt actataaaga gaagacgaaa tagtgacata atatattcta	1980
tttttatact cttcctattt ttgtagtgac ctgtttatga gatgctgggtt ttctaccaa	2040
cggccctgca gccagctcac gtccagggtt aaccacagc tacttggttt gtgttcttct	2100
tcatattcta aaaccattcc atttccaagc actttcagtc caataggtgt aggaaatagc	2160

gctgtttttg ttgtgtgtgc agggagggca gttttctaataat ggaatggttt gggaatatcc	2220
atgtacttgt ttgcaagcag gactttgagg caagtgtggg ccactgtggg ggcagtgagg	2280
gtgggggtgtt tgggaggctg cgtgccagtc aagaagaaaa aggtttgcat tctcacattg	2340
ccaggatgat aagttccttt ctttttcttt aaagaagttg aagtttagga atcctttggg	2400
gccaactggg gtttgaaagt agggacctca gaggtttacc tagagaacag gtgggttttta	2460
agggttatct tagatgtttc acaccggaag gtttttaaac actaaaatat ataatttata	2520
gttaaggcta aaaagtatat ttattgcaga ggatgttcat aaggccagta tgatttataa	2580
atgcaatctc cccttgattt aaacacacag atacacacac acacacacac acacacaaac	2640
cttctgcctt tgatgttaca gatttaatac agtttatttt taaagataga tccttttata	2700
ggtgagaaaa aaacaatctg gaagaaaaaa accacacaaa gacattgatt cagcctgttt	2760
ggcgtttccc agagtcactt gattggacag gcatgggtgc aaggaaaatt aggggtactca	2820
acctaagttc ggttccgatg aattcttatc ccctgcccct tcctttaaaa aacttagtga	2880
caaaatagac aatttgcaca tcttggctat gtaattcttg taatttttat ttaggaagtg	2940
ttgaaggagg gtggcaagag tgtggaggct gacgtgtgag ggaggacagg cgggaggagg	3000
tgtgaggagg aggctcccga ggggaagggg cggtgccac accggggaca ggccgcagct	3060
ccattttctt attgcgctgc taccgttgac ttccaggcac ggtttggaat tattcacatc	3120
gcttctgtgt atctctttca cattgtttgc tgctattgga ggatcagttt tttgttttac	3180
aatgtcatat actgcatgt actagtttta gttttctctt agaacattgt attacagatg	3240
ctttttttgt agtttttttt ttttttatgt gatcaatttt gacttaatgt gattactgct	3300
ctattccaaa aaggttgctg tttcacaata cctcatgctt cacttagcca tgggtggacc	3360
agcgggcagg ttctgcctgc tttggcgggc agacacgcgg gcgcgatccc acacaggctg	3420
gcgggggccc gccccgaggc cgcgtgcgtg agaaccgcgc cggtgtcccc agagaccagg	3480
ctgtgtccct cttctcttcc ctgvcctgt gatgctgggc acttcatctg atcggggg	3540
tagcatcata gtagttttta cagctgtgtt attctttgag ttagctatg gaagttgcat	3600
aattattatt attattatta taacaagtgt gtcttacgtg ccaccacggc gttgtacctg	3660
taggactctc attcgggatg attggaatag cttctggaat ttgttcaagt tttgggtatg	3720
tttaatctgt tatgtactag tgttctgttt gttattgttt tgttaattac accataatgc	3780

taatttaaag agactccaaa tctcaatgaa gccagctcac agtgctgtgt gccccggtca	3840
cctagcaagc tgccgaacca aaagaatttg caccgctg cgggcccacg tggttggggc	3900
cctgccctgg cagggtcatc ctgtgctcgg aggccatctc gggcacaggc ccaccccgcc	3960
ccacccctcc agaacacggc tcacgcttac ctcaaccatc ctggctgcgg cgtctgtctg	4020
aaccacgcgg gggccttgag ggacgctttg tctgtcgtga tggggcaagg gcacaagtcc	4080
tggatgttgt gtgtatcgag aggccaaagg ctggtggcaa gtgcacgggg cacagcggag	4140
tctgtcctgt gacgcgcaag tctgagggtc tgggcggcgg gcggctgggt ctgtgcattt	4200
ctggttgac cgcggcgctt cccagcacca acatgtaacc ggcatgtttc cagcagaaga	4260
caaaaagaca aacatgaaag tctagaaata aaactggtaa aaccca	4307

<210> 16
 <211> 2043
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >CCNE1|ENSG00000105173|ENST00000262643|2043

<400> 16	
gaggggctgg gagccgcggc ggggcggtgc gagggcgggc cggggccggt tccgcgcgca	60
gggattttta atgtcccgt ctgagccggg cgcaggagca gccggcgcg cgcagcgc	120
ggtgtagggg gcaggcgcg atcccgccac cgccgcgcgc tcggcccgcc gactcccggc	180
gccgccgccc cactgccgt cgccgccgcc gcctgccggg actggagcgc gccgtccgcc	240
gcggacaaga ccctggcctc agggcggagc agcccatca tgccgaggga gcgcagggag	300
cgggatgcga aggagcggga caccatgaag gaggacggcg gcgcggagtt ctcggctcgc	360
tccaggaaga ggaaggcaaa cgtgaccgtt tttttgcagg atccagatga agaaatggcc	420
aaaatcgaca ggacggcgag ggaccagtgt gggagccagc cttgggacaa taatgcagtc	480
tgtgcagacc cctgctccct gatccccaca cctgacaaag aagatgatga ccgggtttac	540
ccaaactcaa cgtgcaagcc tcggattatt gcaccatcca gaggctcccc gctgcctgta	600
ctgagctggg caaatagaga ggaagtctgg aaaatcatgt taaacaagga aaagacatac	660
ttaagggatc agcactttct tgagcaacac cctcttctgc agccaaaaat gcgagcaatt	720
cttctggatt ggttaatgga ggtgtgtgaa gtctataaac ttcacaggga gaccttttac	780

ttggcacaag atttctttga ccggtatatg gcgacacaag aaaatgttgt aaaaactctt	840
ttacagctta ttgggatttc atctttatctt attgcagcca aacttgagga aatctatcct	900
ccaaagttgc accagtttgc gtatgtgaca gatggagctt gttcaggaga tgaaattctc	960
accatggaat taatgattat gaaggccctt aagtggcgtt taagtcccct gactattgtg	1020
tcctggctga atgtatacat gcaggttgca tatctaaatg acttacatga agtgctactg	1080
ccgcagtatc cccagcaaat ctttatacag attgcagagc tgttgatct ctgtgtcctg	1140
gatgttgact gccttgaatt tccttatggt atacttgctg cttcggcctt gtatcatttc	1200
tcgtcatctg aattgatgca aaaggtttca gggatcagc ggtgcgacat agagaactgt	1260
gtcaagtgga tggttccatt tgccatggtt ataagggaga cggggagctc aaaactgaag	1320
cacttcaggg gcgtcgctga tgaagatgca cacaacatac agaccacag agacagcttg	1380
gatttgctgg acaaagcccg agcaaagaaa gccatgttgt ctgaacaaaa tagggcttct	1440
cctctcccca gtgggctcct caccgccca cagagcggtg agaagcagag cagcgggccg	1500
gaaatggcgt gaccaccca tccttctcca ccaaagacag ttgcgcgcct gctccacgtt	1560
ctcttctgtc tggtgcagcg gaggcgtgcg tttgctttta cagatatctg aatggaagag	1620
tgtttcttcc acaacagaag tatttctgtg gatggcatca aacagggcaa agtggttttt	1680
attgaatgct tataggtttt ttttaataa gtgggtcaag tacaccagcc acctccagac	1740
accagtgcgt gctcccgatg ctgctatgga aggtgctact tgacctagg gactcccaca	1800
acaacaaaag cttgaagctg tggagggcca cgggtggcgtg gctctcctcg caggtgttct	1860
gggctccgtt gtaccaagtg gagcaggtgg ttgcgggcaa gcgttgtgca gagcccatag	1920
ccagctgggc agggggctgc cctctccaca ttatcagttg acagtgtaca atgcctttga	1980
tgaactgttt tgtaagtgtc gctatatcta tccatttttt aataaagata atactgtttt	2040
tga	2043

<210> 17

<211> 3330

<212> DNA

<213> Homo sapiens

<220>

<223> >CCNE2|ENSG00000175305|ENST00000520509|3330

<400> 17

ccgcgaggct ccgtctcccc ggcgcgggccc gctcgccgtc cggctcgctc agggcctggg	60
cagacgcggc ccgcccgagc taccgcgggt tccgagacgc cttcgactg ctcctccacc	120
cgggggatct ttgttccccg agctgttccc cgcctcgctg ctcccgccgc aaaacctgtt	180
tgcggaatac ccgccgaag tctgacttga cgatgtgcag ttttgggagg ttttatacac	240
ctgaaagaag agaattgtcaa gacgaagtag ccgtttacaa gctaagcagc agccccagcc	300
cagccagacg gaatcccccc aagaagccca gataatccag gccagaaga ggaaaactac	360
ccaggatgtc aaaaaaagaa gagaggaggt caccaagaaa catcagtatg aaattaggaa	420
ttgttggcca cctgtattat ctggggggat cagtccttgc attatcattg aaacacctca	480
caaagaaata ggaacaagtg atttctccag atttacaat tacagattta aaaatctttt	540
tattaatcct tcacctttgc ctgatttaag ctggggatgt tcaaaagaag tctggctaaa	600
catgttaaaa aaggagagca gatatgttca tgacaaacat tttgaagttc tgcattctga	660
cttggacca cagatgaggt ccatacttct agactggctt ttagaggat gtgaagtata	720
cacacttcat agggaaacat tttatcttgc acaagacttt tttgatagat ttatgttgac	780
acaaaaggat ataaataaaa atatgcttca actcattgga attacctcat tattcattgc	840
ttccaaactt gaggaatct atgctcctaa actccaagag tttgcttacg tcaactgatgg	900
tgcttgcagt gaagaggata tcttaaggat ggaactcatt atattaaagg ctttaaaatg	960
ggaactttgt cctgtaacaa tcatctcctg gctaaatctc tttctccaag ttgatgctct	1020
taaagatgct cctaaagtgc ttctacctca gtattctcag gaaacattca ttcaaatagc	1080
tcagctttta gatctgtgta ttctagccat tgattcatta gagttccagt acagaatact	1140
gactgctgct gccttgtgcc attttacctc cattgaagtg gttaagaaag cctcaggttt	1200
ggagtgggac agtatttcag aatgtgtaga ttggatggta cttttgtca atgtagtaaa	1260
aagtactagt ccagtgaagc tgaagacttt taagaagatt cctatggaag acagacataa	1320
tatccagaca catacaaact atttggctat gctggaggaa gttaaattaca taaacacctt	1380
cagaaaaggg ggacagtgtg caccagtgtg caatggaggc attatgacac caccgaagag	1440
cactgaaaaa ccaccaggaa aacactaaag aagataacta agcaaacaag ttggaattca	1500
ccaagattgg gtagaactgg tatcactgaa ctactaaagt ttacagaaa gtagtgctgt	1560

gattgattgc cctagccaat tcacaagtta cactgccatt ctgattttta aacttacaat	1620
tggcactaaa gaatacatTTT aattatttcc tatgttagct gttaaagaaa cagcaggact	1680
tgtttacaaa gatgtcttca ttcccaaggt tactggatag aagccaacca cagtctatac	1740
catagcaatg tttttccttt aatccagtgt tactgtgttt atcttgataa actaggaatt	1800
ttgtcactgg agttttggac tggataagtg ctaccttaaa gggataacta agtgatacag	1860
tactttgaat ctagttgtta gattctcaaa attcctacac tcttgactag tgcaatttgg	1920
ttcttgaaaa ttaaatttaa acttgtttac aaaggtttag ttttgaata aggtgactaa	1980
tttatctata gctgctatag caagctatta taaaacttga atttctacaa atggtgaaat	2040
ttaatgtttt ttaaactagt ttatttgcct tgccataaca ctttttttaa ctaataaggc	2100
ttagatgaac atggtgttca acctgtgctc taaacagtgg gagtaccaa gaaattataa	2160
acaagataaa tgctgtggct ccttcctaac tggggctttc ttgacatgta ggttgcttgg	2220
taataacctt tttgtatatc acaatttggg tgaaaaactt aagtaccctt tcaaactatt	2280
tatatgagga agtcacttta ctactctaag atatccctaa ggaatTTTT ttttaattt	2340
agtgtgacta aggctttatt tatgtttgtg aaactgttaa ggtcctttct aaattcctcc	2400
attgtgagat aaggacagtg tcaaagtgat aaagcttaac acttgacctt aacttctatt	2460
ttcttaagga agaagagtat taaatatata ctgactccta gaaatctatt tattaaaaaa	2520
agacatgaaa acttgctgta cataggctag ctatttctaa atattttaaa ttagcttttc	2580
taaaaaaaa atccagcctc ataaagtaga ttagaaaact agattgctag tttattttgt	2640
tatcagatat gtgaatctct tctcccttg aagaaactat acatttattg ttacggtatg	2700
aagtcttctg tatagtttgt ttttaaacta atatttgttt cagtattttg tctgaaaaga	2760
aaacaccact aattgtgtac atatgtatta tataaactta accttttaac actgtttatt	2820
tttagcccat tgtttaaaaa ataaaagtta aaaaaattta actgcttaaa agtaaagttt	2880
tgccattgct tggagaaaact tttttttcct tctctgcgct gccagctgta acacttcttc	2940
tggattgctt gcattcaact ctgtctggcc gatggctttg atcttccaaa acagaaaagt	3000
gatgttatta gaggtgtgtc aaaaattaag ttttgttggg acaagtaata taaagctacc	3060
tacgtgctaa caacgataca gtttaatgat taactgaacc tcttaactgt aaaaccagg	3120
agtcttgga aaaaattaac ataaagatta accagggcc ctctcaagga aagatggact	3180

gctgagccat agtttatgga attacttaag tggcatttta atatattaag taaattcatg	3240
taaattttatt tgaaagtata agtaagctct acatggcgat ttttagagtt aaactagggc	3300
tccacttggt aatgtgcaaa aataactggt	3330

<210> 18
 <211> 3313
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >CDC7|ENSG00000097046|ENST00000428239|3313

<400> 18	
atctggttcg gtctctggcc cgagggaagc cggtccttcc cggctgagct cgcggccagc	60
gctggccggc ggattcccat tcattcacc ttctcctcct ccgccagta ctctgtggcca	120
gggtcgtatc agttctccgt caacttgctt ggggccttgg acgagcctcc tggcgcttcc	180
tgtcagtggc gaaaagctgc tttgctcccc ctgtggatgt aacccttag ctggcatttt	240
gcatctcaat tggcttgtga tggaggcgtc tttggggatt cagatggatg agccaatggc	300
tttttctccc cagcgtgacc ggtttcaggc tgaaggctct ttaaaaaaaaa acgagcagaa	360
ttttaaaactt gcagggtgta aaaaagatat tgagaagctt tatgaagctg taccacagct	420
tagtaatgtg ttttaagattg aggacaaaat tggagaaggc actttcagct ctgtttattt	480
ggccacagca cagttacaag taggacctga agagaaaatt gctctaaaac acttgattcc	540
aacaagtcac cctataagaa ttgcagctga acttcagtgc ctaacagtgg ctgggggggca	600
agataatgtc atgggagtta aatactgctt taggaagaat gatcatgtag ttattgctat	660
gccatatctg gagcatgagt cgtttttggc cattctgaat tctctttcct ttcaagaagt	720
acgggaatat atgcttaatc tgttcaaagc tttgaaacgc attcatcagt ttggtattgt	780
tcaccgtgat gttaagccca gcaatTTTTT atataatagg cgctgaaaa agtatgcctt	840
ggtagacttt ggtttggccc aaggaacca tgatacgaaa atagagcttc ttaaatttgt	900
ccagtctgaa gctcagcagg aaagggtgtt acaaaacaaa tcccacataa tcacaggaaa	960
caagattcca ctgagtggcc cagtacctaa ggagctggat cagcagtcca ccacaaaagc	1020
ttctgttaaa agaccctaca caaatgcaca aattcagatt aaacaaggaa aagacggaaa	1080
ggagggatct gtaggccttt ctgtccagcg ctctgttttt ggagaaagaa atttcaatat	1140

acacagctcc atttcacatg agagccctgc agtgaaactc atgaagcagt caaagactgt	1200
ggatgtactg tctagaaagt tagcaacaaa aaagaaggct atttctacaa aagttatgaa	1260
tagtgctgtg atgaggaaaa ctgccagttc ttgccagct agcctgacct gtgactgcta	1320
tgcaacagat aaagtttgta gtatttgcct ttcaaggcgt cagcaggttg cccctagggc	1380
aggtacacca ggattcagag caccagaggt cttgacaaag tgccccaatc aaactacagc	1440
aattgacatg tgggtctgcag gtgtcatatt tctttctttg cttagtggac gatatccatt	1500
ttataaagca agtgatgatt taactgcttt ggcccaaatt atgacaatta ggggatccag	1560
agaaactatc caagctgcta aaacttttgg gaaatcaata ttatgtagca aagaagttcc	1620
agcacaagac ttgagaaaac tctgtgagag actcaggggt atggattcta gcactcccaa	1680
gttaacaagt gatatacaag ggcatgcttc tcatcaacca gctatttcag agaagactga	1740
ccataaagct tcttgcctcg ttcaaacacc tccaggacaa tactcaggga attcatttaa	1800
aaagggggat agtaatagct gtgagcattg ttttgatgag tataatacca atttagaagg	1860
ctggaatgag gtacctgatg aagcttatga cctgcttgat aaacttctag atctaaatcc	1920
agcttcaaga ataacagcag aagaagcttt gttgcatcca ttttttaaag atatgagctt	1980
gtgataatgg atcttcattt aatgtttact gttatgaggt agaataaaaa agaatacttt	2040
gtaatagcca caagttcttg tttagagacc agagcaggat taataattta ttttaacatt	2100
ttagtgtttg gtggcacatt ctaaaatata gattaagaat acttaaaatg cctgggatag	2160
ttcttgggac taacaacatg atcttctttg agttaaacct acctaagtag attttaggtg	2220
ggttcctatt aggtcagatt tttagcttcc ctaattacct ttcactgaca tatacagaaa	2280
aaggagcagt tttagtttta attaattaaa attaacagat gtgatgagga ttaaatgaat	2340
caaaagactt aatttgtaga ttcttttaga gttatgagct aggtatagtt tggggaaact	2400
caacctggtg ctggtgctct taacaatttt gtaaataaag aagataattt ctttttctag	2460
aggtacatat taggcctttt atgaacacta aaacaatgag gaaatgttgg tcatggggca	2520
aagtatcact taaaattgaa ttcatccatt tttaaaaaac acttcatgaa agcattctgg	2580
tgtgaattgc catTTTTTt ttactggctt ctcaattttc ttccttctct gccctacct	2640
aaaacattct cctcggaat tacatggtgc tgaccacaaa gtttctggat gttttattaa	2700
atattgtacg tgtttacagt tgggaattta aaataatata tacactggtt gataaaggga	2760

agctgcagga ccaaggtgaa gattgatagt ccaaagtctt ttcttttttg agttgtatat	2820
tttttcacac catcttagat ataattaggt agctgctgaa aggaaaagtg aatacagaat	2880
tgacgggtatt attggagatt tttcctctgc gtagagccat ccagatctct gtatcctgtt	2940
ttgactaagt cttaggtggg ttgggaagac agataatgaa gtaggcaaag agaaaaggac	3000
ccaagataga ggtttatatt cagaaatggg atatatcaat gacagcatat caaacttcct	3060
atgggaaaaa gtctgggtggg tggtcagctg acagatttcc catttagtag tcatagaata	3120
cagaaatagt ttagggacat gtattcattt tgttattttg agcattgata ggtcagtata	3180
tctacctaat ctgtttggta agtataggat atataaacca ttaccattga tctgtcttat	3240
gccataatct taaaaaaaaa ttgaatgctc ttgaatttgt atattcaata aagttatcct	3300
tttatatttt tta	3313

<210> 19
 <211> 8336
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >CDK12|ENSG00000167258|ENST00000447079|8336

<400> 19	
cttttttccc ttcttcaggt caggggaaag ggaatgccca attcagagag acatgggggc	60
aagaaggacg ggagtggagg agcttctgga actttgcagc cgtcatcggg aggcggcagc	120
tctaacagca gagagcgtca ccgcttggtg tcgaagcaca agcggcataa gtccaaacac	180
tccaaagaca tgggggttggg gacccccgaa gcagcatccc tgggcacagt tatcaaacct	240
ttggtggagt atgatgatag cagctctgat tccgacacct tctccgatga catggccttc	300
aaactagacc gaaggagaa cgacgaacgt cgtggatcag atcggagcga ccgcctgcac	360
aaacatcgtc accaccagca caggcgttcc cgggacttac taaaagctaa acagaccgaa	420
aaagaaaaaa gccaaagaat ctccagcaag tcgggatcga tgaaggaccg gatatcggga	480
agttcaaagc gttcgaatga ggagactgat gactatggga aggcgaggt agccaaaagc	540
agcagcaagg aatccaggtc atccaagctc cacaaggaga agaccaggaa agaacgggag	600
ctgaagtctg ggcacaaaaga ccggagtaaa agtcatcgaa aaagggaac acccaaaagt	660

tacaaaacag tggacagccc aaaacggaga tccaggagcc cccacaggaa gtggtctgac	720
agctccaaac aagatgatag cccctcggga gcttcttatg gccaagatta tgaccttagt	780
ccctcacgat ctcatacctc gagcaattat gactcctaca agaaaagtcc tggaagtacc	840
tcgagaaggc agtcggtcag tcccccttac aaggagcctt cggcctacca gtccagcacc	900
cggtcaccga gcccctacag taggcgacag agatctgtca gtccctatag caggagacgg	960
tcgtccagct acgaaagaag tggctcttac agcgggcatg cgcccagtcc ctatggctga	1020
aggcgggtcca gcagcccttt cctgagcaag cggctcttga gtcggagtcc actccccagt	1080
aggaaatcca tgaagtccag aagtagaagt cctgcatatt caagacattc atcttctcat	1140
agtaaaaaga agagatccag ttcacgcagt cgtcattcca gtatctcacc tgtcaggctt	1200
ccacttaatt ccagtctggg agctgaactc agtaggaaaa agaaggaaag agcagctgct	1260
gctgctgcag caaagatgga tggaaaggag tccaagggtt cacctgtatt tttgcctaga	1320
aaagagaaca gttcagtaga ggctaaggat tcaggtttgg agtctaaaaa gttaccacaga	1380
agtgtaaaat tggaaaaatc tgccccagat actgaactgg tgaatgtaac acatctaaac	1440
acagaggtaa aaaattcttc agatacaggg aaagtaaagt tggatgagaa ctccgagaag	1500
catcttgta aagatttgaa agcacaggga acaagagact ctaaaccat agcactgaaa	1560
gaggagattg ttactccaaa ggagacagaa acatcagaaa aggagacccc tccacctctt	1620
cccacaattg cttctcccc acccctcta ccaactacta cccctccacc tcagacaccc	1680
cctttgccac ctttgctcc aataccagct cttccacagc aaccacctct gcctccttct	1740
cagccagcat ttagtcaggt tcctgcttcc agtacttcaa ctttgcccc ttctactcac	1800
tcaaagacat ctgctgtgtc ctctcaggca aattctcagc cccctgtaca ggtttctgtg	1860
aagactcaag tatctgtaac agctgctatt ccacacctga aaacttcaac gttgcctcct	1920
ttgcccctcc cacccttatt acctggagat gatgacatgg atagtccaaa agaaactctt	1980
ccttcaaaac ctgtgaagaa agagaaggaa cagaggacac gtcacttact cacagacctt	2040
cctctccctc cagagctccc tgggtggagat ctgtctcccc cagactctcc agaaccaaag	2100
gcaatcacac cacctcagca accatataaa aagagaccaa aaatttggtg tcctcgttat	2160
ggagaaagaa gacaaacaga aagcgactgg gggaaacgct gtgtggacaa gtttgacatt	2220
attgggatta ttggagaagg aacctatggc caagtatata aagccaagga caaagacaca	2280

ggagaactag tggctctgaa gaaggtgaga ctagacaatg agaaagaggg cttcccaatc	2340
acagccattc gtgaaatcaa aatccttcgt cagttaatcc accgaagtgt tgttaacatg	2400
aaggaaattg tcacagataa acaagatgca ctggatttca agaaggacaa aggtgccttt	2460
taccttgat ttgagtatat ggaccatgac ttaatgggac tgctagaatc tggtttggtg	2520
cacttttctg aggaccatat caagtcgttc atgaaacagc taatggaagg attggaatac	2580
tgtcacaaaa agaatttcct gcatcgggat attaagtgtt ctaacatttt gctgaataac	2640
agtgggcaaa tcaaactagc agattttgga cttgctcggc tctataactc tgaagagagt	2700
cgcccttaca caaacaagt cattactttg tggtagcgac ctccagaact actgctagga	2760
gaggaacgtt acacaccagc catagatgtt tggagctgtg gatgtattct tggggaacta	2820
ttcacaaaga agcctatttt tcaagccaat ctggaactgg ctgagctaga actgatcagc	2880
cgactttgtg gtagcccttg tccagctgtg tggcctgatg ttatcaaact gccctacttc	2940
aacaccatga aaccgaagaa gcaatatcga aggcgtctac gagaagaatt ctctttcatt	3000
ccttctgcag cacttgattt attggaccac atgctgacac tagatcctag taagcgggtgc	3060
acagctgaac agaccctaca gagcgacttc cttaaagatg tcgaactcag caaaatggct	3120
cctccagacc tccccactg gcaggattgc catgagttgt ggagtaagaa acggcgacgt	3180
cagcgacaaa gtggtgttgt agtcgaagag ccacctccat ccaaaacttc tcgaaaagaa	3240
actacctcag ggacaagtac tgagcctgtg aagaacagca gccagcacc acctcagcct	3300
gctcctggca aggtggagtc tggggctggg gatgcaatag gccttgctga catcacacaa	3360
cagctgaatc aaagtgaatt ggcagtgtta ttaaacctgc tgcagagcca aaccgacctg	3420
agcatccctc aaatggcaca gctgcttaac atccactcca acccagagat gcagcagcag	3480
ctggaagccc tgaaccaatc catcagtgcc ctgacggaag ctacttccca gcagcaggac	3540
tcagagacca tggccccaga ggagtctttg aaggaagcac cctctgcccc agtgatcctg	3600
ccttcagcag aacagacgac ccttgaagct tcaagcacac cagctgacat gcagaatata	3660
ttggcagttc tcttgagtca gctgatgaaa acccaagagc cagcaggcag tctggaggaa	3720
aacaacagtg acaagaacag tgggccacag gggccccgaa gaactccac aatgccacag	3780
gaggaggcag cagcatgtcc tcctcacatt cttccaccag agaagaggcc ccctgagccc	3840
cccggacctc caccgccgcc acctccacc cctctggttg aaggcgatct ttccagcgcc	3900

ccccaggagt tgaaccacgc cgtgacagcc gccttgctgc aacttttatc ccagcctgaa	3960
gcagagcctc ctggccacct gccacatgag caccaggcct tgagaccaat ggagtactcc	4020
acccgacccc gtccaaacag gacttatgga aacactgatg ggcctgaaac agggttcagt	4080
gccattgaca ctgatgaacg aaactctggg ccagccttga cagaatcctt ggtccagacc	4140
ctgggtgaaga acaggacctt ctcaggctct ctgagccacc ttggggagtc cagcagttac	4200
cagggcacag ggtcagtgc gtttccaggg gaccaggacc tccgttttgc cagggtcccc	4260
ttagcgttac acccggtggg cgggcaacca ttcctgaagg ctgagggaag cagcaattct	4320
gtggtacatg cagagaccaa attgcaaaac tatggggagc tggggccagg aaccactggg	4380
gccagcagct caggagcagg ctttactgg gggggcccaa ctcagtcctt tgcttatgga	4440
aaactctatc gggggcctac aagagtccca ccaagagggg gaagaggag aggagttcct	4500
tactaacca gagacttcag tgtcctgaaa gattcctttc ctatccatcc ttccatccag	4560
ttctctgaat ctttaatgaa atcatttgcc agagcgagg aatcatctgc atttggctac	4620
tgcaaagctg tccgttgat tccttgctca cttgctacta gcaggcgact tacgaaataa	4680
tgatgttggc accagttccc cctggatggg ctatagccag aacatttact tcaactctac	4740
cttagtagat acaagtagag aatatggaga ggatcattac attgaaaagt aaatgtttta	4800
ttagttcatt gcctgcactt actgatcgga agagagaaag aacagtttca gtattgagat	4860
ggctcaggag aggctctttg atttttaaaag ttttgggggtg ggggattgtg tgtggtttct	4920
ttcttttgaa ttttaattta ggtgttttgg gtttttttcc tttaaagaga atagtgttca	4980
caaaatttga gctgctcttt ggcttttgct ataagggaag cagagtggcc tggctgattt	5040
gaataaatgt ttctttcctc tccaccatct cacattttgc ttttaagtga acactttttc	5100
cccattgagc atcttgaaca tacttttttt ccaaataaat tactcatcct taaagtttac	5160
tccactttga caaaagatac gcccttctcc ctgcacataa agcaggttgt agaacgtggc	5220
attcttgggc aagtaggtag actttacca gtctctttcc ttttttgctg atgtgtgctc	5280
tctctctctc tttctctctc tctctctctc tctctctctc tctctctctc tctgtctcgc	5340
ttgctcgtc tcgctgtttc tctctctttg aggcatgtgt ttggaaaaaa tcgttgagat	5400
gccaagaac ctgggataat tctttacttt ttttgaaata aaggaaagga aattcagact	5460
cttacattgt tctctgtaac tcttcaattc taaaatgttt tgttttttta accatgttct	5520

gatggggaag ttgatttgta agtgtggaca gcttggacat tgctgctgag ctgtggttag	5580
agatgatgcc tccattccta gagggctaataacagcattt agcatattgt ttacacatat	5640
atTTTTatgt caaaaaaaaaa acaaaaacct ttcaaacaga gcattgtgat attgtcaaag	5700
agaaaaacaa atcctgaaga tacatggaaa tgtaacctag tttagggtgg gtatTTTTct	5760
gaagatacat caatacctga cTTTTTTTaa aaaaataatt ttaaaacagc atactgtgag	5820
gaagaacagt attgacatac ccacatccca gcatgtgtac cctgccagtt cTTTTaggga	5880
TTTTtctcc aaagagattt ggatttgggtt ttggtaaaag gggTTaaatt gtgcttccag	5940
gcaagaactt tgccttatca taaacaggaa atgaaaaagg gaagggtgt caggatggga	6000
taatttggga ggcttctcat tctggcttct atttctatgt gaggtagc atatagagt	6060
TTTTaaaaac agatacatgt catataattt atctgcacag acttagacct tcaggaaaca	6120
taggttaagc cccTTTTac aaagaaaaag taaacatact tcagcatctt ggagggtagt	6180
ttcaaaaact caagtttcat gtttcaatgc caagttctta TTTTTaaaaa taaaatctac	6240
ttataagaga aaggtgcatt acttaaaaaa aaaaaacttt aaagaaatga aagaagaacc	6300
ctcttcagat acttacttga agactgtttt cccctgttaa tgagatatag ctagatatcg	6360
gtgtgtgtat ttctttatta ttctctgggtt ttgatctgg ccttgcctcc agggccaaac	6420
actgatttag aaagagagcc ttctagctat ttggcattg atggcttttt ataccagtgt	6480
gtccagttag atttactagg ctactgaca tgctattgggt aaatcgcat aaagttcatc	6540
tgaaccttct gtctgttgac ttcttagtcc tcagacatgg gcctttgtgt tttagaatat	6600
ttgaatttga gttattgggc cccactccct gTTTTtatt aaagaacgtg agcctgggat	6660
actttcagaa gtatctgttc aatgaaaaaa agttggtttc ccatcaaata tgaataaaat	6720
tctctatata ttctattgta ttttggttat cagcagtcac caataatgtt tttccctccc	6780
ctctcccacc tcttattttt aattatgcca aatatactaa ataataact taagcctcca	6840
ttccctcatc cctactaggg aagggggtga gtgtatgtgt gagtgtatgt gtatgtatga	6900
tcccatctca cccccaccc cattttggga gtcttttaaa atgaaaacaa agtttggtag	6960
ttttgactat ttctaaaagc agaggagaaa aaaaaactta tttaaatact ctggaatctg	7020
tatggaggaa gaaaaggtat ttgttaattt ttcagttacg ttatctataa acatgatgga	7080
agtaaagggtt tggcagaatt tcaccttgac tatttgaaaa ttacagaccc aattaattcc	7140

attcaaaagt ggttttcgtt ttgttttaaat tattgtacaa tgagagatat tgtctattaa	7200
atacattatt ttgaacagat gagaaatctg attctgttca tgagtgggag gcaaaactgg	7260
tttgaccgtg atcatttttg tggttttgaa aacaaatata cttgaccag tttccttagt	7320
tttttcttca actgtccata ggaacgataa gtatttgaaa gcaacatcaa atctatacgt	7380
ttaaagcagg gcagtttagca caaatgtgca agtagaactt ctattagctt atgccataga	7440
catcacccaa ccacttgtat gtgtgtgtgt atatataata tgcatatata gttaccgtgc	7500
taaaatgggt accagcaggt ttgagagag aatgctgcat cagaaaagtg tcagttgcca	7560
cctcattctc cctgatttag gttcctgaca ctgattcctt tctctctcgt ttttgacccc	7620
cattgggtgt atcttgtcta tgtacagata ttttgtaata tattaaattt ttttctttca	7680
gtttataaaa atggaaagtg gagattggaa aattaaatat ttcctgttac tataaccactt	7740
ttgctccatt gcatttactt cttaatctgt accccctgag catatctaata catgtataaa	7800
ggacgttttt cctccacttt atcttagggg ttctctgtct cagaatcatt atagactcat	7860
taactcccc tcccagcaaa aggttatcag gatttgaaga ggtgcttgaa aacgctagac	7920
taggaactag agaataaatg agttgggaaa aaccatgaaa tgtgattttt ttaaagtaga	7980
aaagttatac aaataatggt accaaacat caaaagagtt gagcttcatg taccctgact	8040
cctcctgaca ggagaggtaa gtgggtttga gctcaactgt catcaaggga agttggttaag	8100
aggctgttta gacccaaagg atagtcttaa accagacttc accaccacc ctacctcagt	8160
tcccatgtta ttacatgcag agtcagcatg gggattagtg tacctacctt tgctgagatt	8220
tcccgatgcg ttgccaatcc agaaagtga tcaaaaagtt gtttaaaagt taaaatctct	8280
attgtttcca aaatctttcc catctccacc tgaagacaga attgcttccc cttctc	8336

<210> 20

<211> 3502

<212> DNA

<213> Homo sapiens

<220>

<223> >CHEK1|ENSG00000149554|ENST00000534070|3502

<400> 20

gcaattggga ggaggcgtg ccgcaggcgt tttctgcccc ataccgtcc ctatatacctc	60
--	----

ttcctctctt cccagaccc ccacctctcc ctctctcttc cccagtcgtt cgccggaaag	120
--	-----

catttgtctc ccacctcttc ataacaacaa ttaatttcct ctggggcctg aggagggcag	180
aatttcaacc ttcggtgtgc ttgggagtg cgttgtgat ttacacgaca aaatgccgag	240
gtgctcggtg gagtcatggc agtgcccttt gtggaagact gggacttggt gcaaaccctg	300
ggagaaggtg cctatggaga agttcaactt gctgtgaata gagtaactga agaagcagtc	360
gcagtgaaga ttgtagatat gaagcgtgcc gtagactgtc cagaaaatat taagaaagag	420
atctgtatca ataaaatgct aaatcatgaa aatgtagtaa aattctatgg tcacaggaga	480
gaaggcaata tccaatattt atttctggag tactgtagtg gaggagagct ttttgacaga	540
atagagccag acataggcat gcctgaacca gatgctcaga gattcttcca tcaactcatg	600
gcaggggtgg tttatctgca tgggtattgga ataactcaca gggatattaa accagaaaat	660
cttctgttgg atgaaaggga taacctcaaa atctcagact ttggcttggc aacagtattt	720
cggataata atcgtgagcg tttgttgaac aagatgtgtg gtactttacc atatgttgct	780
ccagaacttc tgaagagaag agaatttcat gcagaaccag ttgatgtttg gtcctgtgga	840
atagtactta ctgcaatgct cgctggagaa ttgccatggg accaaccag tgacagctgt	900
caggagtatt ctgactggaa agaaaaaaaa acatacctca acccttgga aaaaatcgat	960
tctgctcctc tagctctgct gcataaaatc ttagttgaga atccatcagc aagaattacc	1020
attccagaca tcaaaaaaga tagatggtac aacaaacccc tcaagaaagg ggcaaaaagg	1080
ccccgagtca cttcaggtgg tgtgtcagag tctcccagtg gattttctaa gcacattcaa	1140
tccaatttgg acttctctcc agtaaacagt gcttctagtg aagaaaatgt gaagtactcc	1200
agttctcagc cagaaccccg cacaggtctt tccttatggg ataccagccc ctcatacatt	1260
gataaattgg tacaagggat cagcttttcc cagcccacat gtcctgatca tatgcttttg	1320
aatagtcagt tacttggcac cccaggatcc tcacagaacc cctggcagcg gttggtcaaa	1380
agaatgacac gattctttac caaattggat gcagacaaat cttatcaatg cctgaaagag	1440
acttgtgaga agttgggcta tcaatggaag aaaagttgta tgaatcaggt tactatatca	1500
acaactgata ggagaaacaa taaactcatt ttcaaagtga atttgttaga aatggatgat	1560
aaaatattgg ttgacttccg gctttctaag ggtgatggat tggagttcaa gagacacttc	1620
ctgaagatta aagggaagct gattgatatt gtgagcagcc agaagatttg gcttcctgcc	1680
acatgatcgg accatcggct ctggggaatc ctggtgaata tagtgctgct atgttgacat	1740

tattcttcct agagaagatt atcctgtcct gcaaactgca aatagtagtt cctgaagtgt	1800
tcacttcctt gtttatccaa acatcttcca atttattttg tttgttcggc atacaaataa	1860
tacctatata ttaattgtaa gcaaaacttt ggggaaagga tgaatagaat tcatttgatt	1920
atttcttcat gtgtgttttag tatctgaatt tgaaactcat ctggtggaaa ccaagtttca	1980
ggggacatga gttttccagc ttttatacac acgtatctca tttttatcaa aacattttgt	2040
ttaattcaaa aagtacatat tccatgttga ttttaattcta agatgaacca ataaagacat	2100
aattcttgtg acttttggac agtagattta tcagtctgtg aagcgaagcc agcttcaaaa	2160
catatcccca agatttgtac ttatatatttc aaaagggcct ggccagttat ataaacctgt	2220
ttttgaatta taatgattaa ttaaaattgc aagtaggtgt tttttccagt gtagttagta	2280
aaatacttgt attttacagt gttgcataaa ctctagtgtc taactaactt tactctaaaa	2340
attactgttg aacatcttaa atatttttct atattttcta ctttcatagc catattttta	2400
ccttttcaac ttactggtga ccaagctttt aggtgataaa gaataaaaaga gggaaggga	2460
gagtaaggaa gctataagaa aaatagatct gattctttgt tcctttacct gttagactta	2520
caaaaagttt gtttttctaa taaaatttgt atcaactttg gggcatatta gggtgaggcc	2580
ttggctcctg cctgtagtcc cagctactta ggaggctgag agaggaggat cgcgtgaacc	2640
tggaagtttg aggctgtagt gagctatgat tgcaccagtg cactccagct tggatgacag	2700
agtaagaccc tacctctaataaaaaattttt aaaattgtaa aacattataa aattaatcag	2760
ttattttaat ctgaagccaa gaacatgtag aatgttatga ttagagttta tcacatatta	2820
atgtatactg gcaaattgtg ttactggagt atacccatag gaggaataaa ttcaaacctg	2880
ttttatttat ttgaacctat ttacggtatg cttagaatt gaatcagtat aaattctcaa	2940
atatgggaga aattttgttc ttgagaatta tctgagtcataaatatattttt caaaaacagc	3000
tctcactgac ttgaacctct tctgtaagct ctaacctttt acctgcttta catttccact	3060
tgaatgtcta gtaggcactct cttgaccaa aacagctttt gattcctgtt ctccaacctg	3120
ttcctctcct agttttctcc atctcagaaa tgttacttcc tctgcaaagt ctttcctga	3180
cttatctaaa ataataacct cctctgtttg ctgtgggaat ttgtatagaa tgggtgggaaa	3240
atttcaagtt tcatatttgg attagctctg acattttatt atctgaacac tggttaattgc	3300
ctcagtaaag aactgataa taagtacctt ttagagttat tttaatcttt aatgctttta	3360

tgtagtaggaa	gagtatagtg	tcctgttttg	cacagaaagg	cattctgtaa	ataataagtt	3420
gccttaattt	tcctgtaatg	ttcattatat	tgtagtgagg	aggtatttac	tcctattatt	3480
aaaaataaaa	atgtgtaaaa	tt				3502

<210> 21
 <211> 1971
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >CHEK2|ENSG00000183765|ENST00000382580|1971

<400> 21						
atctgcaggt	ttagcgccac	tctgctggct	gaggctgcgg	agagtgtgcg	gctccaggtg	60
ggctcacgcg	gtcgtgatgt	ctcgggagtc	ggatgttgag	gctcagcagt	ctcatggcag	120
cagtcctgt	tcacagcccc	atggcagcgt	taccagctcc	caaggctcct	cctcacagtc	180
ccagggcata	tccagctcct	ctaccagcac	gatgccaaac	tccagccagt	cctctcactc	240
cagctctggg	acactgagct	ccttagagac	agtgtccact	caggaactct	attctattcc	300
tgaggaccaa	gaacctgagg	accaagaacc	tgaggagcct	accctgccc	cctgggctcg	360
attatgggcc	cttcaggatg	gatttgccaa	tcttgagaca	gagtctggcc	atgttaccca	420
atctgatctt	gaactcctgc	tgtcatctga	tcctcctgcc	tcagcctccc	aaagtgtctg	480
gataagaggt	gtgaggcacc	atccccggcc	agtttgcagt	ctaaaatgtg	tgaatgacaa	540
ctactggttt	gggagggaca	aaagctgtga	atattgcttt	gatgaaccac	tgctgaaaag	600
aacagataaa	taccgaacat	acagcaagaa	acactttcgg	attttcaggg	aagtgggtcc	660
taaaaactct	tacattgcat	acatagaaga	tcacagtggc	aatggaacct	ttgtaaatac	720
agagcttgta	gggaaaggaa	aacgccgtcc	tttgaataac	aattctgaaa	ttgcactgtc	780
actaagcaga	aataaagttt	ttgtcttttt	tgatctgact	gtagatgatc	agtcagttta	840
tcctaaggca	ttaagagatg	aatacatcat	gtcaaaaact	cttggaagtg	gtgcctgtgg	900
agaggtaaag	ctggcttttcg	agaggaaaac	atgtaagaaa	gtagccataa	agatcatcag	960
caaaaggaag	tttgctattg	gttcagcaag	agaggcagac	ccagctctca	atgttgaaac	1020
agaaatagaa	attttgaaaa	agctaaatca	tccttgcatc	atcaagatta	aaaacttttt	1080

tgatgcagaa gattattata ttgttttggga attgatggaa gggggagagc tgtttgacaa	1140
agtgggtgggg aataaacgcc tgaaagaagc tacctgcaag ctctatTTTT accagatgct	1200
cttggctgtg cagtaccttc atgaaaacgg tattatacac cgtgacttaa agccagagaa	1260
tgttttactg tcatctcaag aagaggactg tcttataaag attactgatt ttgggcactc	1320
caagattttg ggagagacct ctctcatgag aaccttatgt ggaaccccca cctacttggc	1380
gcctgaagtt cttgtttctg ttgggactgc tgggtataac cgtgctgtgg actgctggag	1440
tttaggagtt attcttttta tctgccttag tgggtatcca ctttctctg agcataggac	1500
tcaagtgtca ctgaaggatc agatcaccag tggaaaatac aacttcattc ctgaagtctg	1560
ggcagaagtc tcagagaaag ctctggacct tgtcaagaag ttgttggttag tggatccaaa	1620
ggcacgtttt acgacagaag aagccttaag acaccctggg cttcaggatg aagacatgaa	1680
gagaaagttt caagatcttc tgtctgagga aaatgaatcc acagctctac cccaggttct	1740
agcccagcct tctactagtc gaaagcggcc ccgtgaaggg gaagccgagg gtgccgagac	1800
cacaaagcgc ccagctgtgt gtgctgctgt gttgtgaact ccgtggtttg aacacgaaag	1860
aaatgtacct tctttcactc tgtcatcttt cttttctttg agtctgtttt tttatagttt	1920
gtattttaat tatgggaata attgcttttt cacagtcact gatgtacaat t	1971

<210> 22

<211> 4464

<212> DNA

<213> Homo sapiens

<220>

<223> >DCLRE1A|ENSG00000198924|ENST00000361384|4464

<400> 22

ggactctgag ggcttttttg agctcgctat gcttaacctg gagatgatta aggccccgct	60
tcctggcctc ccagcctcta atgccaaaag ataagggaga ggctggcgtg tgacccccgtt	120
ttgagtcagg tggacagagg gctggccacc ttcggaacca tgggtgcaat acggagtcag	180
acctcaatac aagcccactc tttcacatat ttgaactttt ttcacatatc aacttttttt	240
gttcactgtg cagggattgt tcattgctgc tggaggaaga tcatggactg tcgcgggaaa	300
ctgaagtggg tgagtatcca ctagtcgtgg atgagggcag tgacttcgca gttttttgcg	360
aattacacat ctctttgatt atgttgtgac tagttttgtt agatagtcatt ttagtgtttg	420

ggatacctgt taagcccttt gtccaggac tgtggttgga tttatgaatt atttggacgg	480
ttgtccactt gaaagaactg acagtagctt cataacaatg ttacaaatct cgttctaaga	540
ttaagctgtt gaacctatat ttgccattag cgcttaattt ttgaagtatt atttttatga	600
atcaagccct ggaaaaggac aagatatttg aatgaaatag cacccataat ggagaacttc	660
acagttgcta ctctgtgat aggtttatct tagtttcatt gtggtataaa tggaatagca	720
ggtgttgca ggtacaaggt ttgtagcttg ccaatatgtt cattaccaac actgcagatt	780
cccattgagt tgggtgggggt tttgttacct ttgttttttt tctcagcaaa ataattctat	840
aactttttgt ttgtgacaag aaatggactt tcagtttact taagattaat acttcttgaa	900
tgataaaatc attttgccat gttagaagac atttccgaag aagacatttg ggaatacaaa	960
tctaaaagaa aaccaaacg agttgatcca aataatggct ctaaaaatat tctaaaatct	1020
gttgaaaaag caacagatgg aaaataccag tcaaacgga gtagaaacag aaaaagagcc	1080
gcagaagcta aagagtgaa ggacatgaa gtgccccttg gaaatgcagg ttgtcagact	1140
tctgttgctt ctagtcagaa ttcaagttgt ggagatggta ttcagcagac ccaagacaag	1200
gaaactactc caggaaaact ctgtagaact caaaaaagcc aacacgtgtc cccaaagata	1260
cgccagttt atgatggata ctgtccaaat tgccagatgc ctttttcctc attgataggg	1320
cagacacctc gatggcatgt ttttgaatgt ttggattctc caccacgtc tgaaacagag	1380
tgtcctgatg gtcttctgtg tacctcaacc attccttttc attacaagag atacactcac	1440
ttcctgctag ctcaaagcag ggctggtgat catcctttta gcagcccatc acctgcgtca	1500
ggtggcagtt tcagtgagac taagtcaggc gtcctttgta gccttgagga aagatggtct	1560
tcgtatcaga accaaactga taactcgggt tcaaatgata ccttattgat gacacagtat	1620
tttaaaaagt ctccgtctct gactgaagcc agtgaaaaga tttctactca tatccaaaca	1680
tccaacaag ctctacaatt tacagatttt gttgagaatg acaaactagt gggagttgct	1740
ttgcgtcttg caaacaactc agaacacata aatttgccat tgccagaaaa tgacttcagt	1800
gactgtgaaa tctcctattc tccacttcaa agtgatgaag acactcatga tatcgatgaa	1860
aaaccggatg attcacaaga acaactgttt ttaccgaaa gctcaaaaga tggcagcctc	1920
gaagaagatg atgacagctg tggttttttt aaaaaacgac atggtccctt actgaaggac	1980
caggatgaga gctgccccaa agtgaacagc ttcttaactc gggataagta tgatgaagga	2040

ttgtatagat tcaatagtct aaatgatttg tctcaaccta tttctcaaaa taatgagagt	2100
actttgcctt atgatctggc atgtactggg ggtgattttg tgttgtttcc acctgcattg	2160
gcagggaagc ttgctgcttc tgttcatcag gcaactaaag caaaacctga tgagccagaa	2220
tttcactcag ctcaatcaaa taaacagaaa caggtaattg aagaatcatc tgtttacaat	2280
caagtttctc ttccgttagt taagagttta atgttgaaac cttttgaaag tcaggtagaa	2340
gggtatcttt cttcccaacc aacccaaaat acaattagaa aattatcaag tgagaacttg	2400
aatgctaaga ataatactaa ctcagcatgt ttctgcagaa aggcattaga ggggtgtgcca	2460
gttggtaaaag ctacaatttt aaatacagaa aacttgtcta gtacacctgc tccgaagtat	2520
ttgaaaatat tgccttctgg tcttaagtat aatgcaagac atccttctac caaggtaatg	2580
aagcaaatgg atataggtgt gtattttggg ctacctcca aaagaaagga agagaaattg	2640
ctaggggaaa gtgcattaga agggataaac ttaaatccag ttccaagtcc taatcaaaag	2700
aggtcctcgc agtgcaagag gaaagcagaa aaatctttaa gtgatttaga atttgatgca	2760
agtactttac atgagagtca gctttctgtg gaactttcta gtgagaggtc acagcgtcaa	2820
aaaaagagat gtagaaagtc aaattcactg caggaaggag cgtgtcagaa gagatcagat	2880
caccttatta atacagaatc tgaagcagtc aatttaagta aagtcaaagt cttcacaaaa	2940
tcagctcatg gtgggctgca aaggggcaac aagaaaatcc cagagtcatc taatgtagga	3000
ggatcaagaa aaaagacatg tccattctat aagaaaatac ctggaaccgg ctttacagtt	3060
gatgcctttc agtatggcgt ggttgaagggt tgcacagcct attttctcac acattttcat	3120
tctgatcatt atgctggatt gtctaaacac ttcacatttc cagtttattg tagtgagata	3180
actggcaatt tgttgaagaa caagcttcat gtgcaagaac aatatattca cccattgcca	3240
ctggacactg aatgtattgt gaatgggtgc aaagttgttt tgcttgatgc caatcactgt	3300
ccagggtgctg tcatgatcct cttttatctt cctaattggtg ctgtcatatt acacacggga	3360
gacttcagag cagatcccag catggaacgt tctcttcttg cggaccagaa agtccatatg	3420
ctgtacttag ataccacata ttgtagccca gaatacacct ttccatctca gcaagagggt	3480
atccggtttg ccatcaacac tgcctttgag gctgtaactc taaaccaca tgctcttggt	3540
gtctgtggca cttactctat tggaaaagag aaagtcttcc tagccattgc tgatgtttta	3600
ggttcaaaag tgggcatgtc ccaggaaaaa tataaaactc tacagtcct caatatacca	3660

gaaattaatt cactcatcac taccgacatg tgcagttcat tggttcacct tctcccaatg	3720
atgcaaatta attttaaggg cttacagagt catttgaaga agtgtggtgg gaaatacaat	3780
cagattttgg catttcgacc tacaggatgg acacactcta acaagttcac tagaatagca	3840
gatgttattc cccagaccaa aggaaacatt tcaatatatg gaattcctta cagtgaacac	3900
agcagctacc tagaaatgaa gcgctttgtc cagtggctga agccccagaa aatcatacct	3960
actgtaaatg tgggcacctg gaaatctagg agcacaatgg agaaatattt tagagagtgg	4020
aaattggaag ctggatattg atgatacctc cgaggattca gtagtagtta agttccttgg	4080
atgtagcttg ttagtagtta aatctataga aatgtgaaat acactttgtg tggaaaaacc	4140
tcatgaagat tgttcagata ctttattttc tcatttatgt ttgaacaaca tgttcgtggt	4200
gctgaatgcc tctcagcatc atcaaggata actgaaactg ggtctccctg ggacccttaa	4260
tttcttgtcc cctgccctcc atgggcagtt atattctgca tcaagcctta gaagaggaag	4320
caaaggcaga ttcagggacc aaaaggatta atgataatta ataaagtagt ttgaagcatt	4380
atatatataa gtaattatgt gtctttaaaa ttatgagatg aaacttttat atgacgtgta	4440
tacttaaata aaattaatat aaaa	4464

<210> 23

<211> 3940

<212> DNA

<213> Homo sapiens

<220>

<223> >DCLRE1B|ENSG00000118655|ENST00000369563|3940

<400> 23

accaaacaga agaccgctgc cggctctccc gcgagcgagc gctaagcgta gtcccgactc	60
cgaccttagg atgcccttcc cccaacctct cactctccag gcgcagctgc gcggctccct	120
ctggtggagc tgcggcgggc aatcggaat cggctacttt ggcctgtctc ctccctgcag	180
cgcgcgcttt gaggccccgg ctccggcctcc gctcccgcgc gggtgggagt gtccagcgcc	240
ctccgcgatt tgggctccag cgggcagggt gacttccttt ttctgcccac tctggtaact	300
tattgctctg ctgggctctt tcccttaggg tctctggccc tgttcttgcc ccagcatgac	360
ttttatcggg acgccgttgt ggaagcctca cgcaggagcc ctgccccgt ggagaagatc	420

ccactggtga ctccaaccct accaccatga atggggctct gatcccccac acgcccacgc	480
cagtggactt ctggagcctg cgccgggctg gcaccgcacg tctcttcttc ttgtctcaca	540
tgcactcgga ccacaccgtg ggcctgtcta gcacctgggc ccggcccctc tactgctccc	600
caattacagc ccacctcttg catcgtcacc tacaggatc taagcaatgg atccaagccc	660
tggaggttgg tgagagccat gtattacccc tagatgaaat tggacaagag accatgaccg	720
taaccctcct cgatgccaat cactgtcctg gttctgtcat gtttctcttt gaaggatatt	780
ttggaacat cctctacaca ggtgattttc gatacacacc atccatgcta aaggagccag	840
ccctgacact ggggaaacag atccatactt tatacctaga caacaccaat tgcaatccag	900
ccctggttct tccttcccga caagaagctg cccaccagat tgtccagctc attcgaaaac	960
accacaaca taacataaag attggactct acagcctggg aaaggaatca ctgctggagc	1020
agctggccct ggagtttcag acctgggtgg tattgagtcc tcggcgccctg gagttggtac	1080
agctactggg cctggcagat gtgttcacag tggaggagaa ggctggccgc atccatgcag	1140
tagaccatat ggagatctgc cattccaaca tgctgcgttg gaaccagacc caccctacga	1200
ttgctatcct tcccacaagc cgaaaaatcc acagctccca ccctgatatc cacgtcatcc	1260
cttactctga ccattcctct tactccgagc ttcgtgcctt tgtcgcagca ctgaagcctt	1320
gccaggtggg gccattgta agtcggcggc cctgtggagg ctttcaggac agtctgagcc	1380
ccaggatctc cgtgcccctg attccggact ctgtacagca atacatgagt tcttcctcta	1440
gaaaaccaag ctttctctgg ctgttagaaa ggaggctaaa gaggccgaga acccaaggtg	1500
ttgtgtttga atcccctgag gaaagtgtg atcaatctca agctgacaga gactcaaaga	1560
aggccaagaa agagaaactt tctccctggc ctgcggacct tgaaaagcag ctttcccacc	1620
atcctttgcg gatcaagaag cagttgttcc cagatctcta tagcaaagaa tggaacaagg	1680
cagtgccttt ctgtgagtct caaaagaggg tgactatgtt gacggcccca ctgggatttt	1740
cagtgcactt aaggtctaca gatgaggagt ttattttctca aaaaaccagg gaggaattg	1800
gtttagggtc ccccttggtg cccatgggag atgatgatgg aggtccagaa gccacaggga	1860
atcagagtgc ctggatgggc catggttctc ccctgtccca cagcagcaag ggcaccctc	1920
ttctagctac tgaattcagg ggtctagcac tcaaatatct tctgactcca gtgaactttt	1980
tccaggcagg gtattcttcc aggagatttg accagcaagt ggaaaaatac cataaacct	2040

gctgaagaca ggagagtaca gaatgacaac attgagccca cactgcagtt ttgaagatag	2100
taactgatgg ctgggtgggaa agagtttggt tttggggcct acttttctat ctttacaaga	2160
ctcttatggg cccaccgtgg agcagcactt cccaaaactt gttcactggg gtcctcgtgc	2220
ctatggaatc cttcttttta taactaagtt taagaaatac tttttttata aaatctttgg	2280
agtatgcgtg agcaaattaa aagttctttg aagtcctaca gtaacttaat ctgtttaacc	2340
ttgtttaacc cagtatttct caaacttttg tgaacatgca atcatcttat gtgggtacag	2400
aaagaggtaa agagtctgaa tcaaaaagga ccaggttatt gctgttgctg ttttgtgggtg	2460
tcatgagcca ttctccatgt ccccttctcc ctcttctcag atcaaaatcc ctagggagtt	2520
ctatttttta aattatgaac tatggcgctg catgcttcaa tcctgaacgt cactgacttg	2580
ctgtgaccat ccaaataatt ttcctgtctc tgcctctggg agggaacagg aagcgatgaa	2640
gaggctttgg aacagtagtg aaaattctac ctctatgtcc ttcatgagga tgtgcagtat	2700
cccagtatca ctgggatcca tgtggaacag agccagctgg ggggttgggc agctctctcc	2760
aaggcagtac ctagagccca gctgaacaac aaggctttgg gtgtgaaggg actccccagc	2820
ctggagaccc tatttggctg aaacagttac aaaatatcaa atgtgttgct agatattcct	2880
ccaattgttc acatagctgg gatatttggt gctcccctca ccccttgat tatgtaggga	2940
gccagtgcac acagcctggt tgttttagta tccaaggaag agaccaagga gccagctggc	3000
gggaaggggt gggggtgtgc agtctgccct gtccttctgc tcataacctg acaaaatgcc	3060
aaactagtaa gcaggatagc tgataccacg gctatgaggg agtaggctct gagagggcac	3120
agacttgagg agctgggcgt ctggatcaaa actgctttgg gatggaacct cgagccctag	3180
cagtgaagaa gattccattt cttgtccagg ggatttaaaa gagttttctg ctttgagaga	3240
gaaatagaga gtttagaaag caattgctct tgggaaagct atacacagct ctgttttctc	3300
aatgaccttt gttgtaagtc tcccaacgtc ctattaggag ccacagcagg tgaggcattt	3360
ggtgcagcag gaaacatggg gactgcctag gctcgaatct gtggcaccct gagcaattac	3420
ttaaattgtg gagcctagtt cctcatctgt aagatggact tgagattcct acctctcatg	3480
attactatgg agattgaata attggtaaaa ttctcctagc tcagtgactg ccacaggatg	3540
ggtctttcag attttggttc tctttagctt ctggttcttg aaagaaatta atctgtatat	3600
aacataagaa actttgaaag tcaaaaaaac aaaaaatttt aattcctcgt agattaattg	3660

atttgctatc ttttagtttt tttttctatg catgtagatg tattaatat gtaaagtgtt	3720
ttcaaagttg aggtaatat gtatagaatt ttatagccta cattttaatt tcttgatc	3780
ttaagcattt cacttattag atattgttca aaaatgccat ttttaatat tgtataatac	3840
cctatcatgt gagtatacct taactaagcc attcccatat tcaacatttt gtgtactgtt	3900
tttctaatta catatattac aatgaacaac cttatgcata	3940

<210> 24
 <211> 2354
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >DCLRE1C|ENSG00000152457|ENST00000378278|2354

<400> 24	
gcgggttttg ggtcccgac tctgggatcg gcggcgctat gagttctttc gaggggcaga	60
tggccgagta tccaactatc tccatagacc gcttcgatag ggagaacctg agggcccgcg	120
cctacttcct gtcccactgc cacaagatc acatgaaagg attaagagcc cctaccttga	180
aaagaagggt ggagtgcagc ttgaagggtt atctatactg ttcacctgtg actaaggagt	240
tgttgtaac gagcccgaaa tacagatttt ggaagaaacg aattatatct attgaaatcg	300
agactcctac ccagatatct ttagtgatg aagcatcagg agagaaggaa gagattgttg	360
tgactctctt accagctggt cactgtccgg gatcagttat gtttttattt cagggaata	420
atggaactgt cctgtacaca ggagacttca gattggcgca aggagaagct gctagaatgg	480
agcttctgca ctccgggggc agagtcaaag acatccaaag tgtatatttg gatactacgt	540
tctgtgatcc aagattttac caaattcaa gtcgggagga gtgtttaagt ggagtcttag	600
agctgggtccg aagctggatc actcggagcc cgtaccatgt tgtgtggctg aactgcaaag	660
cggcttatgg ctatgaatat ctgttcacca acctagtgaga agaattagga gtccaggttc	720
atgtgaataa gctagacatg ttttaggaaca tgcctgagat cttcatcat ctcaaacag	780
accgcaacac tcagatccat gcatgccggc atcccaaggc agaggaatat tttcagtga	840
gcaaattacc ctgtggaatt acttccagaa atagaattcc actccacata atcagcatta	900
agccatccac catgtggttt ggagaaagga gcagaaaaac aaatgtaatt gtgaggactg	960
gagagagttc atacagagct tgtttttctt ttcactcctc ctacagtgag attaaagatt	1020

tcttgagcta cctctgtcct gtgaacgcat atccaaatgt cattccagtt ggcacaacta	1080
tggataaagt tgtcgaaatc ttaaagcctt tatgccggtc ttcccaaagt acggagccaa	1140
agtataaacc actgggaaaa ctgaagagag ctagaacagt tcaccgagac tcagaggagg	1200
aagatgacta tctctttgat gatcctctgc caataccttt aaggcacaaa gttccatacc	1260
cggaaacttt tcaccctgag gtattttcaa tgactgcagt atcagaaaag cagcctgaaa	1320
aactgagaca aaccccgagga tgctgcagag cagagtgtat gcagagctct cgtttcacaa	1380
actttgtaga ttgtgaagaa tccaacagtg aaagtgaaga agaagtagga atcccagctt	1440
cactgcaagg agatctgggc tctgtacttc acctgcaaaa ggctgatggg gatgtacccc	1500
agtgggaagt attctttaaa agaaatgatg aaatcacaga tgagagtttg gaaaacttcc	1560
cttcctccac agtggcaggg ggatctcagt caccaaagct tttcagtgac tctgatggag	1620
aatcaactca catctcctcc cagaattctt cccagtcaac acacataaca gaacaaggaa	1680
gtcaaggctg ggacagccaa tctgatactg ttttgttatc ttccaagag agaaacagtg	1740
gggatattac ttccttggac aaagctgact acagaccaac aatcaaagag aatattcctg	1800
cctctctcat ggaacaaaat gtaatttgcc caaaggatac ttactctgat ttgaaaagca	1860
gagataaaga tgtgacaata gttcctagta ctggagaacc aactactcta agcagtgaga	1920
cacatatacc cgaggaaaaa agtttgctaa atcttagcac aaatgcagat tcccagagct	1980
cttctgattt tgaagttccc tcaactccag aagctgagtt acctaaacga gagcatttac	2040
aatatttata tgagaagctg gcaactggtg agagtatagc agtcaaaaaa agaaaatgct	2100
cactcttaga tacctaagaa ttcaaagcgt ttcaacctag agcaaccact aaaaaacctg	2160
cacagagatg acagtcaata ttacaataga gaaaatacag tacttaaaaa tgttcaaata	2220
acctggttgg gtgtggtggc tcacacttgt aatcccagca ctttgaggtg ggcaatggct	2280
tgagcccagg agttcgacac cagcctggcc aacacagtga aatgtgtctc tacttacaaa	2340
aaaaaaaaaa aaaa	2354

<210> 25

<211> 2422

<212> DNA

<213> Homo sapiens

<220>

<223> >DYRK1A|ENSG00000157540|ENST00000398960|2422

<400> 25

gttatagttt tgccgctgga ctcttcctc ctttcccca ccccatcagg atgatatgag	60
acttgaaaga agacgatgca tacaggagga gagacttcag catgcaaacc ttcattctgtt	120
cggcttgcac cgtcattttc attccatgct gctggccttc agatggctgg acagatgccc	180
cattcacatc agtacagtga ccgtcgccag ccaaacataa gtgaccaaca ggtttctgcc	240
ttatcatatt ctgaccagat tcagcaacct ctaactaacc aggtgatgcc tgatattgtc	300
atgttacaga ggcggatgcc ccaaaccttc cgtgaccag caactgctcc cctgagaaaa	360
ctttctgttg acttgatcaa aacatacaag catattaatg aggtttacta tgcaaaaaag	420
aagcgaagac accaacaggg ccaggagagc gattctagtc ataagaagga acggaagggt	480
tacaatgatg gttatgatga tgataactat gattatattg taaaaaacgg agaaaagtgg	540
atggatcggt acgaaattga ctcccttgata ggcaaagggt cctttggaca ggttgtaaag	600
gcatatgatc gtgtggagca agaattgggt gccattaaaa taataaagaa caagaaggct	660
tttctgaatc aagcacagat agaagtgcga cttcttgagc tcatgaacaa acatgacact	720
gaaatgaaat actacatagt gcatttgaaa cgccacttta tgtttcgaaa ccatctctgt	780
ttagtttttg aaatgctgtc ctacaacctc tatgacttgc tgagaaacac caatttccga	840
ggggtctctt tgaacctaac acgaaagttt gcgcaacaga tgtgcactgc actgcttttc	900
cttgcgactc cagaacttag tatcattcac tgtgatctaa aacctgaaaa tacccttctt	960
tgtaaccca aacgcagtgc aatcaagata gttgactttg gcagttcttg tcagttgggg	1020
cagaggatat accagtatat tcagagtcgc ttttatcggg ctccagaggt gctactggga	1080
atgccttatg accttgccat tgatatgtgg tccctcgggt gtattttggt tgaaatgcac	1140
actggagaac ctctgttcag tgggtgccaat gaggtagatc agatgaataa aatagtggaa	1200
gttctgggta ttccacctgc tcatattctt gaccaagcac caaaagcaag aaagttcttt	1260
gagaagttgc cagatggcac ttggaactta aagaagacca aagatggaaa acgggagtac	1320
aaaccaccag gaaccgtaa acttcataac attcttgagg tggaaacagg aggacctggt	1380
gggcgacgtg ctggggagtc aggtcatagc gtcgtgact acttgaagtt caaagacctc	1440
attttaagga tgcttgatta tgacccaaa actcgaattc aaccttatta tgctctgcag	1500

cacagtttct tcaagaaaac agctgatgaa ggtacaaata caagtaatag tgtatctaca	1560
agccccgcca tggagcagtc tcagtcttcg ggcaccacct ccagtacatc gtcaagctca	1620
ggtggctcat cggggacaag caacagtggg agagcccggg cggatccgac gcaccagcat	1680
cggcacagtg gtgggcactt cacagctgcc gtgcaggcca tggactgcga gacacacagt	1740
ccccagggtgc gtcagcaatt tcctgctcct cttggttggg caggcactga agctcctaca	1800
caggtcactg ttgaaactca tcctgttcaa gaaacaacct ttcattgtagc ccctcaacag	1860
aatgcattgc atcatcacca tggtaacagt tcccatcacc atcaccacca ccaccacat	1920
caccaccacc atggacaaca agccttgggt aaccggacca ggccaagggt ctacaattct	1980
ccaacgaata gtcctctac ccaagattct atggaggttg gccacagtca ccactccatg	2040
acatccctgt cttcctcaac gacttcttcc tcgacatctt cctcctctac tggtaaccaa	2100
ggcaatcagg cctaccagaa tcgcccagtg gctgctaata ccttggactt tggacagaat	2160
ggagctatgg acgttaattt gaccgtctac tccaatcccc gccaaagagac tggcatagct	2220
ggacatccaa cataccaatt ttctgctaata acaggtcctg cacattacat gactgaagga	2280
catctgacaa tgaggcaagg ggctgataga gaagagtccc ccatgacagg agtttgtgtg	2340
caacagagtc ctgtagctag ctcgtgacta cattgaaact tgagtttggt tcttgtgtgt	2400
ttttatagaa gtggtgtttt tt	2422

<210> 26

<211> 9821

<212> DNA

<213> Homo sapiens

<220>

<223> >EGFR|ENSG00000146648|ENST00000275493|9821

<400> 26

gccggagtcc cgagctagcc ccggcggccg ccgccgcca gaccggacga caggccacct	60
cgtcggcgtc cgcccagtc cccgcctcgc cgccaacgcc acaaccaccg cgcacggccc	120
cctgactccg tccagtattg atcgggagag ccggagcgag ctcttcgggg agcagcgatg	180
cgaccctccg ggacggccgg ggcagcgctc ctggcgctgc tggctgcgct ctgcccggcg	240
agtcgggctc tggaggaaaa gaaagtttgc caaggcacga gtaacaagct cacgcagttg	300
ggcacttttg aagatcattt tctcagcctc cagaggatgt tcaataactg tgaggtggtc	360

cttggaatt tggaaattac ctatgtgcag aggaattatg atctttcctt cttaaagacc	420
atccaggagg tggctggtta tgtcctcatt gccctcaaca cagtggagcg aattcctttg	480
gaaaacctgc agatcatcag aggaaatatg tactacgaaa attcctatgc cttagcagtc	540
ttatctaact atgatgcaaa taaaaccgga ctgaaggagc tgcccatgag aaatttacag	600
gaaatcctgc atggcgccgt gcggttcagc aacaaccctg ccctgtgcaa cgtggagagc	660
atccagtggc gggacatagt cagcagtgac tttctcagca acatgtcgat ggacttccag	720
aaccacctgg gcagctgcc aagtgatgat ccaagctgtc ccaatgggag ctgctggggt	780
gcaggagagg agaactgcc aaaactgacc aaaatcatct gtgccagca gtgctccggg	840
cgctgccgtg gcaagtcccc cagtgactgc tgccacaacc agtgtgctgc aggctgcaca	900
ggccccggg agagcgactg cctggctctgc cgcaaattcc gagacgaagc cacgtgcaag	960
gacacctgcc cccactcat gctctacaac cccaccacgt accagatgga tgtgaacccc	1020
gagggcaa at acagcttttg tgccacctgc gtgaagaagt gtccccgtaa ttatgtggtg	1080
acagatcacg gctcgtgcgt ccgagcctgt ggggccgaca gctatgagat ggaggaagac	1140
ggcgtccgca agtgtaagaa gtgcgaaggg ccttgccgca aagtgtgtaa cggaataggt	1200
attggtgaat ttaaagactc actctcata aatgctacga atattaaaca cttcaaaaac	1260
tgcacctcca tcagtggcga tctccacatc ctgccggtgg catttagggg tgactccttc	1320
acacatactc ctctcttgga tccacaggaa ctggatatct tgaaaaccgt aaaggaaatc	1380
acagggtttt tgctgattca ggcttggcct gaaaacagga cggacctcca tgcctttgag	1440
aacctagaaa tcatacgcg caggaccaag caacatggtc agttttctct tgcagtcgtc	1500
agcctgaaca taacatcctt gggattacgc tccctcaagg agataagtga tggagatgtg	1560
ataatttcag gaaacaaaaa tttgtgctat gcaaatacaa taaactggaa aaaactgttt	1620
gggacctccg gtcagaaaac caaaattata agcaacagag gtgaaaacag ctgcaaggcc	1680
acaggccagg tctgcatgc cttgtgctcc cccgagggtc gctggggccc ggagcccagg	1740
gactgcgtct cttgccggaa tgtcagccga ggcagggaat gcgtggacaa gtgcaacctt	1800
ctggagggtg agccaaggga gtttgtggag aactctgagt gcatacagtg ccaccagag	1860
tgcctgcctc aggccatgaa catcacctgc acaggacggg gaccagacaa ctgtatccag	1920
tgtgcccact acattgacgg cccccactgc gtcaagacct gcccggcagg agtcatggga	1980

gaaaacaaca ccctggtctg gaagtacgca gacgccggcc atgtgtgcca cctgtgccat	2040
ccaaactgca cctacggatg cactgggcca ggtcttgaag gctgtccaac gaatgggcct	2100
aagatcccggt ccatcgccac tgggatgggtg ggggccctcc tcttgctgct ggtggtggcc	2160
ctggggatcg gcctcttcat gcgaaggcgc cacatcgttc ggaagcgcac gctgcggagg	2220
ctgctgcagg agaggagct tgtggagcct cttacacca gtggagaagc tcccaaccaa	2280
gctctcttga ggatcttgaa ggaaactgaa ttcaaaaaga tcaaagtgtgct gggctccggt	2340
gcgttcggca cggtgtataa gggactctgg atcccagaag gtgagaaagt taaaattccc	2400
gtcgctatca aggaattaag agaagcaaca tctccgaaag ccaacaagga aatcctcgat	2460
gaagcctacg tgatggccag cgtggacaac cccacgtgt gccgcctgct gggcatctgc	2520
ctcacctcca ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat	2580
gtccgggaac acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc	2640
gcaaagggca tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg	2700
aacgtactgg tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg	2760
ctgggtgcgg aagagaaaga ataccatgca gaaggaggca aagtcctat caagtggatg	2820
gcattggaat caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg	2880
gtgactgttt gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc	2940
gagatctcct ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc	3000
gatgtctaca tgatcatggt caagtgcctg atgatagacg cagatagtcg cccaaagttc	3060
cgtgagttga tcatcgaatt ctccaaaatg gcccgagacc ccagcgcta cttgtcatt	3120
cagggggatg aaagaatgca tttgccagt cctacagact ccaacttcta ccgtgccctg	3180
atggatgaag aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag	3240
ggcttcttca gcagccccctc cacgtcacgg actccccctc tgagctctct gagtgaacc	3300
agcaacaatt ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag	3360
gaagacagct tcttgtagcg atacagctca gacccacag gcgccttgac tgaggacagc	3420
atagacgaca ctttctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc	3480
gctggctctg tgcagaatcc tgtctatcac aatcagcctc tgaacccgc gccagcaga	3540
gaccacact accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact	3600

gtccagccca cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc	3660
agccaccaaa ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc	3720
aagccaaatg gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg	3780
ccacaaagca gtgaatttat tggagcatga ccacggagga tagtatgagc ctaaaaaatc	3840
cagactcttt cgataccag gaccaagcca cagcagggtcc tccatcccaa cagccatgcc	3900
cgcattagct cttagacca cagactgggt ttgcaacgtt tacaccgact agccaggaag	3960
tacttccacc tcgggcacat tttgggaagt tgcattcctt tgtcttcaaa ctgtgaagca	4020
tttacagaaa cgcatccagc aagaatattg tccctttgag cagaaattta tctttcaaag	4080
aggtatatat gaaaaaaaa aaaagtatat gtgaggattt ttattgattg gggatcttgg	4140
agtttttcat tgtcgctatt gatttttact tcaatgggct cttccaacaa ggaagaagct	4200
tgctggtagc acttgctacc ctgagttcat ccaggcccaa ctgtgagcaa ggagcacaag	4260
ccacaagtct tccagaggat gcttgattcc agtggttctg cttcaaggct tccactgcaa	4320
aacactaaag atccaagaag gccttcatgg cccagcagg ccggatcggg actgtatcaa	4380
gtcatggcag gtacagtagg ataagccact ctgtcccttc ctgggcaaag aagaaacgga	4440
ggggatggaa ttcttcctta gacttacttt tgtaaaaatg tccccacggg acttactccc	4500
cactgatgga ccagtgggtt ccagtcatga gcgttagact gacttgtttg tcttccattc	4560
cattgttttg aaactcagta tgctgcccct gtcttgctgt catgaaatca gcaagagagg	4620
atgacacatc aaataataac tcggattcca gccacattg gattcatcag catttgacc	4680
aatagcccac agctgagaat gtggaatacc taaggatagc accgcttttg ttctcgaaa	4740
aacgtatctc ctaatttgag gctcagatga aatgcatcag gtcctttggg gcatagatca	4800
gaagactaca aaaatgaagc tgctctgaaa tctcctttag ccatcacccc aaccccccaa	4860
aattagtttg tgttacttat ggaagatagt tttctccttt tacttcactt caaaagcttt	4920
ttactcaaag agtatatgtt ccctccagggt cagctgcccc caaacccct ctttacgctt	4980
tgtcacacaa aaagtgtctc tgccttgagt catctattca agcacttaca gctctggcca	5040
caacagggca ttttacagggt gcgaatgaca gtagcattat gagtagtgtg gaattcagggt	5100
agtaaatatg aaactagggt ttgaaattga taatgctttc acaacatttg cagatgtttt	5160
agaaggaaaa aagttccttc ctaaaataat ttctctacaa ttggaagatt ggaagattca	5220

gctagttagg agcccacctt ttttccta	ctgtgtgtgc cctgtaacct gactgg	ttaa	5280
cagcagtcct ttgtaaacag tgttttaa	ac tctcctagtc aatatccacc ccatccaatt		5340
tatcaaggaa gaaatggttc agaaaatatt	ttcagcctac agttatgttc agtcacacac		5400
acatacaaaa tgttcctttt gcttttaa	ag taatttttga ctcccagatc agtcagagcc		5460
cctacagcat tgtaagaaa gtatttgatt	tttgtctcaa tgaaaataaa actatattca		5520
tttccactct attatgctct caaatacccc	taagcatcta tactagcctg gtatgggtat		5580
gaaagataca aagataaata aaacatagtc	cctgattcta agaaattcac aatttagcaa		5640
aggaaatgga ctcatagatg ctaaccttaa	aacaacgtga caaatgccag acaggaccca		5700
tcagccaggc actgtgagag cacagagcag	ggaggttggg tcctgcctga ggagacctgg		5760
aaggaggaggc tcacaggagg atgaccagg	t ctcagtcagc ggggagggtgg aaagtgcagg		5820
tgcatcaggg gcaccctgac cgaggaaaca	gctgccagag gcctccactg ctaaagtcca		5880
cataaggctg aggtcagtca ccctaaacaa	cctgctccct ctaagccagg ggatgagctt		5940
ggagcatccc acaagttccc taaaagttgc	agcccccagg gggattttga gctatcatct		6000
ctgcacatgc ttagtgagaa gactacacaa	catttctaag aatctgagat tttatattgt		6060
cagttaacca ctttcattat tcattcacct	caggacatgc agaaatattt cagtcagaac		6120
tgggaaacag aaggacctac attctgctgt	cacttatgtg tcaagaagca gatgatcgat		6180
gaggcaggtc agttgtaagt gtagcacatt	gtagcattaa attctagtat ttttgtagtt		6240
tgaaacagta acttaataaa agagcaaaag	ctattctagc tttcttcttc atattttaat		6300
tttccaccat aaagtttagt tgctaaattc	tattaatttt aagattgtgc ttcccaaat		6360
agttctcact tcatctgtcc agggaggcac	agttctgtct ggtagaagcc gcaaagccct		6420
tagcctcttc acggatctgg cgactgtgat	gggcaggtca ggagaggagc tgcccaaagt		6480
cccatgattt tcacctaaaca gccctgatca	gtcagtactc aaagcttgga ctccatccct		6540
gaaggtcttc ctgattgata gcctggcctt	aataccctac agaaagcctg tccattggct		6600
gtttcttcct cagtcagttc ctggaagacc	ttaccccatg accccagctt cagatgtggt		6660
ctttggaaac agaggctgaa ggaaagtaag	gagctgagag ctcacattca taggtgccgc		6720
cagccttcgt gcatcttctt gcatcatctc	taaggagctc ctctaattac accatgcccg		6780
tcaccccatg agggatcaga gaagggatga	gtcttctaaa ctctatattc gctgtgagtc		6840

caggttgtaa gggggagcac tgtggatgca tcctattgca ctccagctga tgacacaaaa	6900
gcttaggtgt ttgctgaaag ttcttgatgt tgtgacttac caccctgcc tcacaactgc	6960
agacataagg ggactatgga ttgcttagca ggaaaggcac tggttctcaa gggcggctgc	7020
ccttgggaat cttctggtcc caaccagaaa gactgtggct tgattttctc aggtgcagcc	7080
cagccgtagg gccttttcag agcaccctcct ggttattgca acattcatca aagtttctag	7140
aacctctggc ctaaaggaag ggcctgggtg gatctacttg gcactcgctg gggggccacc	7200
ccccagtgcc actctcacta ggcctctgat tgcacttggt taggatgaag ctggtgggtg	7260
atgggaactc agcacctccc ctccaggcaga aaagaatcat ctgtggagct tcaaaagaag	7320
gggcctggag tctctgcaga ccaattcaac ccaaatctcg ggggctcttt catgattcta	7380
atgggcaacc aggggtgaaa cccttatttc tagggctctc agttgtacaa gactgtgggt	7440
ctgtaccaga gccccgtca gagtagaata aaaggctggg tagggtagag attcccatgt	7500
gcagtggaga gaacaatctg cagtcactga taagcctgag acttggctca tttcaaaagc	7560
gttcaattca tcctcaccag cagttcagct ggaaaggggc aaatacccc acctgagctt	7620
tgaaaacgcc ctgggacctt ctgcattctc taagtaagtt atagaaacca gtctcttccc	7680
tcctttgtga gtgagctgct attccacgta ggcaacacct gttgaaattg ccctcaatgt	7740
ctactctgca tttctttctt gtgataagca cacactttta ttgcaacata atgatctgct	7800
cacatttcct tgcctggggg ctgtaaaacc ttacagaaca gaaatccttg cctctttcac	7860
cagccacacc tgccatacca ggggtacagc tttgtactat tgaagacaca gacaggattt	7920
ttaaagttaa atctatTTTT gtaactttgt tgcgggatat agttctcttt atgtagcact	7980
gaactttgta caatatattt ttagaaactc atttttctac taaaacaaac acagtttact	8040
ttagagagac tgcaatagaa tcaaaatttg aaactgaaat ctttgtttaa aagggttaag	8100
ttgaggcaag aggaaagccc tttctctctc ttataaaaag gcacaacctc attggggagc	8160
taagctaggt cattgtcatg gtgaagaaga gaagcatcgt ttttatattt aggaaatttt	8220
aaaagatgat ggaaagcaca tttagcttgg tctgaggcag gttctgttgg ggcagtgtta	8280
atggaaaggg ctactgttg ttactactag aaaaatccag ttgcatgcca tactctcatc	8340
atctgccagt gtaaccctgt acatgtaaga aaagcaataa catagcactt tgttggttta	8400
tatatataat gtgacttcaa tgcaaatTTT atttttatat ttacaattga tatgcattta	8460

ccagtataaa ctagacatgt ctggagagcc taataatggt cagcacactt tggtttagttc	8520
accaacagtc ttaccaagcc tgggccccagc caccctagag aagttattca gccctggctg	8580
cagtgcacatc acctgaggag cttttaaaag cttgaagccc agctacacct cagaccgatt	8640
aaacgcaa at ctctggggct gaaaccaag cattcgtagt ttttaaagct cctgaggtca	8700
ttccaatgtg cggccaaagt tgagaactac tggcctaggg attagccaca aggacatgga	8760
cttggaggca aattctgcag gtgtatgtga ttctcaggcc tagagagcta agacacaaag	8820
acctccacat ctgtcgtga gagtcaagaa cctgaacaga gtttccatga aggttctcca	8880
agcactagaa gggagagtgt ctaaacaatg gttgaaaagc aaaggaaata taaaacagac	8940
acctctttcc atttcctaag gtttctctct ttattaaggg tggactagta ataaaatata	9000
atattcttgc tgcttatgca gctgacattg ttgccctccc taaagcaacc aagtagcctt	9060
tatttcccac agtgaaagaa aacgctggcc tatcagttac attacaaaag gcagatttca	9120
agaggattga gtaagtagtt ggatggcttt cataaaaaca agaattcaag aagaggattc	9180
atgctttaag aaacatttgt tatacattcc tcacaaatta tacctgggat aaaaactatg	9240
tagcaggcag tgtgttttcc ttccatgtct ctctgcacta cctgcagtgt gtcctctgag	9300
gctgcaagtc tgtcctatct gaattcccag cagaagcact aagaagctcc accctatcac	9360
ctagcagata aaactatggg gaaaacttaa atctgtgcat acatttctgg atgcatttac	9420
ttatctttaa aaaaaaagga atcctatgac ctgatttggc cacaaaaata atcttgctgt	9480
acaatacaat ctcttgga aa ttaagagatc ctatggattt gatgactggg attagagggtg	9540
acaatgtaac cgattaacaa cagacagcaa taacttcgtt ttagaaacat tcaagcaata	9600
gctttatagc ttcaacatat ggtacgtttt aaccttgaaa gttttgcaat gatgaaagca	9660
gtatttgtac aaatgaaaag cagaattctc ttttatatgg tttatactgt tgatcagaaa	9720
tggtgattgt gcattgagta ttaaaaaatt agatgtatat tattcattgt tctttactcc	9780
tgagtacctt ataataataa taatgtattc tttgttaaca a	9821

<210> 27

<211> 4545

<212> DNA

<213> Homo sapiens

<220>

<223> >ERBB2|ENSG00000141736|ENST00000269571|4545

<400> 27

ggagaaacca ggggagcccc ccgggcagcc gcgcgcccct tcccacgggg ccctttactg	60
cgccgcgcgc ccggccccc cccctcgcag caccgcgcgc ccgcgcctt cccagccggg	120
tccagccgga gccatggggc cggagccgca gtgagcacca tggagctggc ggccttgtgc	180
cgctgggggc tcctcctcgc cctcttgccc cccggagccg cgagcaccca agtgtgcacc	240
ggcacagaca tgaagctgcg gctccctgcc agtcccgaga cccacctgga catgctccgc	300
cacctctacc agggctgcc ggtggtgcag ggaaacctgg aactcaccta cctgcccacc	360
aatgccagcc tgtccttctt gcaggatatc caggaggtgc agggctacgt gctcatcgct	420
cacaaccaag tgaggcaggt cccactgcag aggctgcgga ttgtgcgagg caccagctc	480
tttgaggaca actatgccct ggccgtgcta gacaatggag acccgctgaa caataccacc	540
cctgtcacag gggcctcccc aggaggcctg cgggagctgc agcttcgaag cctcacagag	600
atcttgaaag gaggggtctt gatccagcgg aacccccagc tctgctacca ggacacgatt	660
ttgtggaagg acatcttcca caagaacaac cagctggctc tcacactgat agacaccaac	720
cgctctcggg cctgccaccc ctgttctccg atgtgtaagg gctcccgtg ctggggagag	780
agttctgagg attgtcagag cctgacgcgc actgtctgtg ccggtggctg tgcccgtgc	840
aaggggccac tgcccactga ctgctgccat gagcagtgtg ctgccggctg cacgggcccc	900
aagcactctg actgcctggc ctgcctccac ttcaaccaca gtggcatctg tgagctgcac	960
tgcccagccc tggtcaccta caacacagac acgtttgagt ccatgccc aa tcccaggggc	1020
cggatatacat tcggcgccag ctgtgtgact gcctgtccct acaactacct ttctacggac	1080
gtgggatcct gcaccctcgt ctgccccctg cacaaccaag aggtgacagc agaggatgga	1140
acacagcggg gtgagaagtg cagcaagccc tgtgcccagag tgtgctatgg tctgggcatg	1200
gagcacttgc gagaggtag ggcagttacc agtgccaata tccaggagtt tgctggctgc	1260
aagaagatct ttgggagcct ggcatttctg ccggagagct ttgatgggga cccagcctcc	1320
aacactgccc cgctccagcc agagcagctc caagtgtttg agactctgga agagatcaca	1380
ggttacctat acatctcagc atggccggac agcctgcctg acctcagcgt cttccagaac	1440
ctgcaagtaa tccggggacg aattctgcac aatggcgcct actcgtgac cctgcaaggg	1500

ctgggcatca gctggctggg gctgcgctca ctgagggaac tgggcagtgg actggccctc	1560
atccaccata acaccacct ctgcttcgtg cacacggtgc cctgggacca gctcttttcgg	1620
aaccgcacc aagctctgct ccacactgcc aaccggccag aggacgagtg tgtgggcgag	1680
ggcctggcct gccaccagct gtgcgcccga gggcactgct ggggtccagg gcccaccag	1740
tgtgtcaact gcagccagtt ctttcggggc caggagtgcg tggaggaatg ccgagtactg	1800
caggggctcc ccaggagta tgtgaatgcc aggcactggt tgccgtgcca ccctgagtgt	1860
cagccccaga atggctcagt gacctgtttt ggaccggagg ctgaccagtg tgtggcctgt	1920
gccactata aggacctcc cttctgcgtg gcccgtgcc ccagcgggtg gaaacctgac	1980
ctctcctaca tgcccatctg gaagtttcca gatgaggagg gcgcatgcca gccttgcccc	2040
atcaactgca cccactcctg tgtggacctg gatgacaagg gctgccccgc cgagcagaga	2100
gccagccctc tgacgtccat catctctgcg gtggttgga ttctgctggt cgtggtcttg	2160
ggggtggtct ttgggatcct catcaagcga cggcagcaga agatccgga gtacacgatg	2220
cggagactgc tgcaggaaac ggagctggtg gagccgtga cacctagcgg agcgatgccc	2280
aaccaggcgc agatgcggat cctgaaagag acggagctga ggaaggtgaa ggtgcttgga	2340
tctggcgctt ttggcacagt ctacaagggc atctggatcc ctgatgggga gaatgtgaaa	2400
attccagtgg ccatcaaagt gttgaggga aacacatccc ccaaagccaa caaagaaatc	2460
ttagacgaag catacgtgat ggctggtgtg ggctcccat atgtctcccg ctttctgggc	2520
atctgcctga catccacggt gcagctggtg acacagctta tgccctatgg ctgcctctta	2580
gaccatgtcc gggaaaaccg cggacgcctg ggctcccagg acctgctgaa ctggtgtatg	2640
cagattgcca aggggatgag ctacctggag gatgtgcggc tcgtacacag ggacttggcc	2700
gctcggaacg tgctggtcaa gagtcccaac catgtcaaaa ttacagactt cgggctggct	2760
cggctgctgg acattgacga gacagagtac catgcagatg ggggcaaggt gccatcaag	2820
tggatggcgc tggagtccat tctccgccgg cggttcacc accagagtga tgtgtggagt	2880
tatggtgtga ctgtgtggga gctgatgact tttggggcca aaccttacga tgggatccca	2940
gcccgggaga tccctgacct gctggaaaag ggggagcggc tgccccagcc ccccatctgc	3000
accattgatg tctacatgat catggtcaaa tgttggatga ttgactctga atgtcggcca	3060
agattccggg agttggtgtc tgaattctcc cgcattggcca gggaccccca gcgctttgtg	3120

gtcatccaga atgaggactt gggcccagcc agtcccttgg acagcacctt ctaccgctca	3180
ctgctggagg acgatgacat gggggacctg gtggatgctg aggagtatct ggtaccccag	3240
cagggcttct tctgtccaga ccctgccccg ggcgctgggg gcatggtcca ccacaggcac	3300
cgcagctcat ctaccaggag tggcggtggg gacctgacac tagggctgga gccctctgaa	3360
gaggaggccc ccaggtctcc actggcacc tccgaagggg ctggctccga tgtatttgat	3420
ggtgacctgg gaatgggggc agccaagggg ctgcaaagcc tccccacaca tgaccccagc	3480
cctctacagc ggtacagtga ggaccccaca gtaccctgc cctctgagac tgatggctac	3540
gttgcccccc tgacctgcag cccccagcct gaatatgtga accagccaga tgttcggccc	3600
cagccccctt cgccccgaga gggccctctg cctgctgccc gacctgctgg tgccactctg	3660
gaaaggccca agactctctc cccagggaag aatggggctg tcaaagacgt ttttgccttt	3720
gggggtgccg tggagaacct cgagtacttg acaccccagg gaggagctgc ccctcagccc	3780
caccctcctc ctgccttcag cccagccttc gacaacctct attactggga ccaggacca	3840
ccagagcggg gggctccacc cagcaccttc aaaggacac ctacggcaga gaaccagag	3900
tacctgggtc tggacgtgcc agtgtgaacc agaaggccaa gtccgcagaa gccctgatgt	3960
gtcctcaggg agcaggaag gcctgacttc tgctggcatc aagaggtggg agggccctcc	4020
gaccacttcc aggggaacct gccatgccag gaacctgtcc taaggaaacct tccttcctgc	4080
ttgagttccc agatggctgg aaggggtcca gcctcgttgg aagaggaaca gcaactggga	4140
gtctttgtgg attctgaggc cctgccaat gagactctag ggtccagtgg atgccacagc	4200
ccagcttggc cctttccttc cagatcctgg gtactgaaag ccttagggaa gctggcctga	4260
gaggggaagc ggccctaagg gagtgtctaa gaacaaaagc gaccattca gagactgtcc	4320
ctgaaaccta gtactgcccc ccatgaggaa ggaacagcaa tgggtgtcagt atccaggctt	4380
tgtacagagt gcttttctgt ttagttttta ctttttttgt tttgtttttt taaagatgaa	4440
ataaagacct agggggagaa tgggtgttgt atggggaggc aagtgtgggg ggtccttctc	4500
cacaccact ttgtccattt gcaaataat tttggaaaac agcta	4545

<210> 28

<211> 5919

<212> DNA

<213> Homo sapiens

<220>

<223> >ERBB3|ENSG00000065361|ENST00000267101|5919

<400> 28

attacccttc cctggatctg ggggctttcg gaatctcgac ctccccttgg cctatctcct	60
gcagaaaaat tagggtgagc cccatcctcg atctgctccg ccaagttgcg ggaccgcggg	120
gcgtggcacg ctcggggcag gcggtccgag gctccgcaat ccctactcca gcctcgcgcg	180
ggagggggcg cggccgtgac tcacccctt ccctctgcgt tcctccctcc ctctctctct	240
ctctctcaca cacacacacc cctcccctgc catccctccc cggactccgg ctccggctcc	300
gattgcaatt tgcaacctcc gctgccgtcg ccgcagcagc caccaattcg ccagcggttc	360
aggtggctct tgccctgatg tcctagccta ggggcccccg ggccggactt ggctgggctc	420
ccttcaccct ctgcggagtc atgagggcga acgacgctct gcaggtgctg ggcttgcttt	480
tcagcctggc ccggggctcc gaggtgggca actctcaggc agtgtgtcct gggactctga	540
atggcctgag tgtgaccggc gatgctgaga accaatacca gacactgtac aagctctacg	600
agaggtgtga ggtggtgatg gggaaccttg agattgtgct cacgggacac aatgccgacc	660
tctccttcct gcagtggatt cgagaagtga caggctatgt cctcgtggcc atgaatgaat	720
tctctactct accattgccc aacctccgcg tgggtgcgagg gaccaggtc tacgatggga	780
agtttgccat cttcgtcatg ttgaactata acaccaactc cagccacgct ctgcgccagc	840
tccgcttgac tcagctcacc gagattctgt caggggggtgt ttatattgag aagaacgata	900
agctttgtca catggacaca attgactgga gggacatcgt gagggaccga gatgctgaga	960
tagtggtgaa ggacaatggc agaagctgtc cccctgtca tgaggtttgc aaggggcgat	1020
gctggggctc tggatcagaa gactgccaga cattgaccaa gaccatctgt gctcctcagt	1080
gtaatggtca ctgctttggg cccaaccca accagtgtg ccatgatgag tgtgccgggg	1140
gctgctcagg ccctcaggac acagactgct ttgcctgccg gcacttcaat gacagtggag	1200
cctgtgtacc tcgctgtcca cagcctcttg tctacaacaa gctaactttc cagctggaac	1260
ccaatcccca caccaagtat cagtatggag gagtttgtgt agccagctgt ccccataact	1320
tttggtgga tcaaacatcc tgtgtcaggg cctgtcctcc tgacaagatg gaagtagata	1380
aaaatgggct caagatgtgt gagccttgtg ggggactatg tcccaaagcc tgtgagggaa	1440
caggctctgg gagccgcttc cagactgtgg actcgagcaa cattgatgga tttgtgaact	1500

gcaccaagat cctgggcaac ctggactttc tgatcaccgg cctcaatgga gaccctggc	1560
acaagatccc tgccctggac ccagagaagc tcaatgtctt ccggacagta cgggagatca	1620
caggttacct gaacatccag tcctggccgc cccacatgca caacttcagt gttttttcca	1680
at ttgacaac cattggaggc agaagcctct acaaccgggg cttctcattg ttgatcatga	1740
agaacttgaa tgtcacatct ctgggcttcc gatccctgaa ggaaattagt gctgggcgta	1800
tctatataag tgccaatagg cagctctgct accaccactc tttgaactgg accaaggtgc	1860
ttcggggggc tacggaagag cgactagaca tcaagcataa tcggccgcgc agagactgcg	1920
tggcagaggg caaagtgtgt gaccactgt gtcctcttg gggtgctgg ggcccaggcc	1980
ctggtcagtg cttgtcctgt cgaaattata gccgaggagg tgtctgtgtg acccactgca	2040
actttctgaa tggggagcct cgagaatttg cccatgaggc cgaatgcttc tcctgccacc	2100
cggaatgcca acccatggag ggcaactgcca catgcaatgg ctgggctct gatacttgtg	2160
ctcaatgtgc ccattttcga gatgggcccc actgtgtgag cagctgcccc catggagtcc	2220
taggtgccaa gggcccaatc tacaagtacc cagatgttca gaatgaatgt cggccctgcc	2280
atgagaactg caccagggg tgtaaaggac cagagcttca agactgttta ggacaaacac	2340
tggtgctgat cggcaaaacc catctgacaa tggctttgac agtgatagca ggattggtag	2400
tgattttcat gatgctgggc ggcaactttc tctactggcg tgggcgccgg attcagaata	2460
aaagggctat gaggcgatac ttggaacggg gtgagagcat agagcctctg gacccagtg	2520
agaaggctaa caaagtcttg gccagaatct tcaaagagac agagctaagg aagcttaaag	2580
tgcttggtc ggggtgtctt ggaactgtgc acaaaggagt gtggatccct gagggatgaat	2640
caatcaagat tccagtctgc attaaagtca ttgaggacaa gagtggacgg cagagttttc	2700
aagctgtgac agatcatatg ctggccattg gcagcctgga ccatgcccac attgtaaggc	2760
tgctgggact atgcccaggg tcactctctgc agcttgtcac tcaatatttg cctctgggtt	2820
ctctgctgga tcatgtgaga caacaccggg gggcactggg gccacagctg ctgctcaact	2880
ggggagtaca aattgccaa ggaatgtact accttgagga acatggtatg gtgcatagaa	2940
acctggctgc ccgaaacgtg ctactcaagt caccagtc ggttcaggtg gcagattttg	3000
gtgtggctga cctgctgcct cctgatgata agcagctgct atacagtgag gccaaagactc	3060
caattaagtg gatggccctt gagagtatcc actttgggaa atacacacac cagagtgatg	3120

tctggagcta tgggtgtgaca gtttgggagt tgatgacctt cggggcagag ccctatgcag	3180
ggctacgatt ggctgaagta ccagacctgc tagagaaggg ggagcggttg gcacagcccc	3240
agatctgcac aattgatgtc tacatgggtga tgggtcaagtg ttggatgatt gatgagaaca	3300
ttcgcccaac ctttaaagaa ctagccaatg agttcaccag gatggcccga gacccaccac	3360
ggtatctggt cataaagaga gagagtgggc ctggaatagc ccctgggcca gagccccatg	3420
gtctgacaaa caagaagcta gaggaagtag agctggagcc agaactagac ctagacctag	3480
acttggaaagc agaggaggac aacctggcaa ccaccacact gggctccgcc ctgagcctac	3540
cagttggaac acttaatcgg ccacgtggga gccagagcct ttttaagtcca tcatctggat	3600
acatgccccat gaaccagggt aatcttgggg agtcttgcca ggagtctgca gtttctggga	3660
gcagtgaacg gtgccccgt ccagtctctc tacaccaat gccacgggga tgcctggcat	3720
cagagtcac agaggggcat gtaacaggct ctgaggctga gctccaggag aaagtgtcaa	3780
tgtgtaggag ccggagcagg agccggagcc cacggccacg cggagatagc gcctaccatt	3840
cccagcgcca cagtctgctg actcctgtta cccactctc cccacccggg ttagaggaag	3900
aggatgtcaa cggttatgtc atgccagata cacacctcaa aggtactccc tcctccggg	3960
aaggcacctt ttcttcagt ggtctcagtt ctgtcctggg tactgaagaa gaagatgaag	4020
atgaggagta tgaatacatg aaccggagga gaaggcacag tccacctcat ccccctaggc	4080
caagttccct tgaggagctg gggttatgagt acatggatgt ggggtcagac ctgagtcct	4140
ctctgggcag cacacagagt tgcccactcc accctgtacc catcatgccc actgcaggca	4200
caactccaga tgaagactat gaatatatga atcggcaacg agatggaggt ggtcctgggg	4260
gtgattatgc agccatgggg gcctgcccag catctgagca agggatatgaa gagatgagag	4320
cttttcaggg gcctggacat caggcccccc atgtccatta tgcccgcta aaaactctac	4380
gtagcttaga ggctacagac tctgcctttg ataaccctga ttactggcat agcaggcttt	4440
tccccaaggc taatgcccag agaacgtaac tcctgctccc tgtggcactc agggagcatt	4500
taatggcagc tagtgccttt agagggtacc gtcttctccc tattccctct ctctcccagg	4560
tcccagcccc ttttccccag tcccagacaa ttccattcaa tctttggagg cttttaaaca	4620
ttttgacaca aaattcttat ggtatgtagc cagctgtgca ctttcttctc tttcccaacc	4680
ccaggaaagg ttttccttat tttgtgtgct ttcccagtc cattcctcag cttcttcaca	4740

ggcactcctg gagatatgaa ggattactct ccatatccct tcctctcagg ctcttgacta	4800
cttggaacta ggctcttatg tgtgcctttg tttcccatca gactgtcaag aagaggaaag	4860
ggaggaaacc tagcagagga aagtgttaatt ttggtttatg actcttaacc ccctagaaag	4920
acagaagctt aaaatctgtg aagaaagagg ttaggagtag atattgatta ctatcataat	4980
tcagcactta actatgagcc aggcatacata ctaaacttca cctacattat ctcacttagt	5040
cctttatcat ccttaaaaca attctgtgac atacatatta tctcatttta cacaaaggga	5100
agtcgggcat ggtggctcat gcctgtaatc tcagcacttt gggaggctga ggcagaagga	5160
ttacctgagg caaggagttt gagaccagct tagccaacat agtaagaccc ccctctcttt	5220
aaaaaaaaa aaaaaaaaaa aaaaaaaact ttagaactgg gtgcagtggc tcatgcctgt	5280
aatcccagcc agcactttgg gaggctgaga tgggaagatc acttgagccc agaattagag	5340
ataagcctat ggaaacatag caagacactg tctctacagg ggaaaaaaaa aaaagaaact	5400
gagccttaaa gagatgaaat aaattaagca gtagatccag gatgcaaaat cctcccaatt	5460
cctgtgcatg tgctcttatt gtaagggtgcc aagaaaaact gatttaagtt acagcccttg	5520
tttaaggggc actgtttctt gtttttgcac tgaatcaagt ctaaccccaa cagccacatc	5580
ctcctatacc tagacatctc atctcaggaa gtggtggtgg gggtagtcag aaggaaaaat	5640
aactggacat ctttgtgtaa accataatcc acatgtgccg taaatgatct tcactcctta	5700
tccgagggca aattcacaag gatccccaag atccactttt agaagccatt ctcatccagc	5760
agtgagaagc ttccaggtag gacagaaaaa agatccagct tcagctgcac acctctgtcc	5820
ccttggatgg ggaactaagg gaaaacgtct gttgtatcac tgaagttttt tgttttgttt	5880
ttatacgtgt ctgaataaaa atgccaaagt tttttttca	5919

<210> 29
 <211> 4153
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >ERCC2|ENSG00000104884|ENST00000391945|4153

ggatgtccac gacccggcct ctcgctgaat attcatgagg gaggcgggtc gaccccgctg	60
---	----

cacagtccgg ccggcgccat gaagctcaac gtggacgggc tcctggtcta cttcccgtac	120
gactacatct accccgagca gttctcctac atgcgggagc tcaaacgcac gctggacgcc	180
aagggtcatg gagtccctgga gatgccctca ggcaccggga agacagtatc cctgttggcc	240
ctgatcatgg cataccagag agcatatccg ctggaggtga ccaaactcat ctactgctca	300
agaactgtgc cagagattga gaaggtgatt gaagagcttc gaaagttgct caacttctat	360
gagaagcagg agggcgagaa gctgccgttt ctgggactgg ctctgagctc ccgcaaaaac	420
ttgtgtattc accctgaggt gacaccctg cgctttggga aggacgtcga tgggaaatgc	480
cacagcctca cagcctccta tgtgcgggcg cagtaccagc atgacaccag cctgccccac	540
tgccgattct atgaggaatt tgatgcccac gggcgtgagg tgcccctccc cgctggcatc	600
tacaacctgg atgacctgaa ggccctgggg cggcgccagg gctggtgccc atacttcctt	660
gctcgatact caatcctgca tgccaatgtg gtggtttata gctaccacta cctcctggac	720
cccaagattg cagacctggt gtccaaggaa ctggcccga aggccgtcgt ggtcttcgac	780
gaggcccaca acattgacaa cgtctgcatc gactccatga gcgtcaacct caccgcgagg	840
acccttgacc ggtgccaggg caacctggag accctgcaga agacggtgct caggatcaaa	900
gagacagacg agcagcgctt gcgggacgag taccggcgctc tgggtggaggg gctgcgggag	960
gccagcgccg cccgggagac ggacggccac ctggccaacc ccgtgctgcc cgacgaagtg	1020
ctgcaggagg cagtgcctgg ctccatccgc acggccgagc atttcctggg cttcctgagg	1080
cgggtgctgg agtacgtgaa gtggcggctg cgtgtgcagc atgtggtgca ggagagcccg	1140
cccgcttcc tgagcggcct ggcccagcgc gtgtgcatcc agcgcaagcc cctcagattc	1200
tgtgctgaac gcctccggtc cctgctgcat actctggaga tcaccgacct tgctgacttc	1260
tccccgtca ccctccttgc taactttgcc acccttgtca gcacctacgc caaaggcttc	1320
accatcatca tcgagccctt tgacgacaga accccgacca ttgccaaccc catcctgcac	1380
ttcagctgca tggacgcctc gctggccatc aaaccgtat ttgagcgttt ccagtctgtc	1440
atcatcacat ctgggacact gtccccgtg gacatctacc ccaagatcct ggacttccac	1500
cccgtcacca tggcaacctt caccatgacg ctggcacggg tctgcctctg ccctatgatc	1560
atcggccgtg gcaatgacca ggtggccatc agctccaaat ttgagacccg ggaggatatt	1620
gctgtgatcc ggaactatgg gaacctcctg ctggagatgt ccgctgtggt ccctgatggc	1680

atcgtggcct tcttcaccag ctaccagtac atggagagca ccgtggcctc ctggtatgag	1740
caggggatcc ttgagaacat ccagaggaac aagctgctct ttattgagac ccaggatggt	1800
gccgaaacca gtgtcgccct ggagaagtac caggaggcct gcgagaatgg ccgcggggcc	1860
atcctgctgt cagtggcccg gggcaaagtg tccgaggga tgcactttgt gcaccactac	1920
gggcggggccg tcatcatgtt tggcgtcccc tacgtctaca cacagagccg catttctcaag	1980
gcgcggctgg aatacctgcg ggaccagttc cagattcgtg agaatgactt tcttaccttc	2040
gatgccatgc gccacgcggc ccagtgtgtg ggtcggggca tcaggggcaa gacggactac	2100
ggcctcatgg tctttgccga caagcggttt gcccgtaggg acaagcgggg gaagctgccc	2160
cgctggatcc aggagcacct cacagatgcc aacctcaacc tgaccgtgga cgagggtgtc	2220
caggtggcca agtacttcct gcggcagatg gcacagccct tccaccggga ggatcagctg	2280
ggcctgtccc tgctcagcct ggagcagcta gaatcagagg agacgctgaa gaggatagag	2340
cagattgctc agcagctctg agtggggcgg gtggggccat aaacggttcc tggtagctcc	2400
tgagtcttgc ctggccctgg ttcccagcgg cggtaggtgt agaaggtctt atgaagtcag	2460
gtgacatttc tctactgtcac gtccacagcc tttaatcgca ggagaaggca gctatccacc	2520
aggtaccagc aggcaagggg gggccaggag atgatagacc ccctctcacc ccaccagccc	2580
atccctcctg cactgttccc aagaagctgc ggccagcgca gttcactgga cgttagtggg	2640
ggctcaggtc ctgggtgctg gcgctgaggg tccgaggggc cttgtccagg tgccagctgg	2700
gaaactggag acaggcacag gggatggttg gaggggaggt gaggggcaag tgttacatcc	2760
accctccagc tccagggtcca gcctggcagc caactatcca acccctcgc ctctacccca	2820
agtaaaacc atctttgctc ctgtcctggg tctcttctga caccaacca ccacctgctg	2880
actccaccac tggcccgacc gccctgcctt tgcctcatct ctgcacctac ccacatggcc	2940
tcctctatct ttgaccaga tgtccccctc acctgagtcc caggagccct tggagcatcc	3000
acgttcagtg tgttgagtga catggctctc ttcattctgc aaagagggca gcagggagga	3060
aatgagtga tccaggagtg gccccctcc acgagggacc tttccagcac agggtttgat	3120
ctgtgtgtat cacaggggag atgggagcca tggaaggttc ttgagcaaga tgggggtggg	3180
ggtgggggtg gctgtggatt ctgctgatgt cagggtctct tgccgacctg atcaacactc	3240
acccggctgc tcctgccgc gcctcgcctc ggagccggga gaccagcttc tcaattcctc	3300

gcctgataga ctcacggatc ttggagagtg agctgctgcg gcgaagggcc tggggagggg	3360
ccgggggatgc tgctcagtcc cccaggccca tgcagtcccc acgcccaccc caggagccga	3420
agcacagccc atcctcacct gttctgctgc accagctgtg cctgtgttgg gggcacctag	3480
ggggaaagat gggggacact gagatgggga catagagaag ccaccaggag ggaagacaga	3540
caggggctag gggctcaccg agaggggagg gtaggtcctc cttgtggagg atttctttgt	3600
acagctcttc cgcttggtga tacttggtct gtttcaggta ggctgaggcc tgcggagaag	3660
ggagtgtctg ggaacttttc ggaccccagt tcccaacact gcatccagtc ccgccttttt	3720
tttttttttt gtagagacag gggctcttgg atgttgccca ggctggtctt gaactcctag	3780
cctcaagtga tcctcctgcc ttggcctccc aaagtgttgg ggttacaagc gtgaaccacc	3840
acgtctggcc ccactttcta atgtgctgac ctggccgtat ccttcccctg cctaaagcct	3900
tgccctgggg cagagtgggt ttggagctcg gccttacttc ctgccctggg gatggaggac	3960
aggcagcctc ctgtttctga gcctgtctgc tgtactttgg gctggtggca tggcctcggt	4020
gcctcagtct tcccacacgc aaagtggagc gacagctccc acctcatagg gacgtataa	4080
gaactagcga tgtgggggacc gggctaagag gggagccggc cacagtgagc tcaataaatg	4140
tctgctgctg cca	4153

<210> 30

<211> 2750

<212> DNA

<213> Homo sapiens

<220>

<223> >ERCC3|ENSG00000163161|ENST00000285398|2750

<400> 30

gggagcttcc ggattgagcc ggaagtcccc ccagagcgga tgccgcggcg ggcctgtggg	60
agcgggggtca tcttctctct gctgctgtag ctgccatggg caaaagagac cgagcggacc	120
gcgacaagaa gaaatccagg aagcggcact atgaggatga agaggatgat gaagaggacg	180
ccccggggaa cgaccctcag gaagcgggtc cctcggcggc ggggaagcag gtggatgagt	240
caggcaccaa agtggatgaa tatggagcca aggactacag gctgcaaatg ccgctgaagg	300
acgaccacac ctccaggccc ctctgggtgg ctcccgatgg ccatatcttc ttggaagcct	360
tctctccagt ttacaaatat gcccaagact tcttggtggc tattgcagag ccagtgtgcc	420

gaccaaccca tgtgcatgag tacaaactaa ctgcctactc cttgtatgca gctgtcagcg	480
ttgggctgca aaccagtgc atcaccgagt acctcaggaa gctcagcaag actggagtcc	540
ctgatggaat tatgcagttt attaagttgt gtactgtcag ctatggaaaa gtcaagctgg	600
tcttgaagca caacagatac ttcgttgaaa gttgccaccc tgatgtaatc cagcatcttc	660
tccaggaccc cgtgatccga gaatgccgct taagaaactc tgaaggggag gccactgagc	720
tcatacacaga gactttcaca agcaaactctg ccattttctaa gactgctgaa agcagtgggtg	780
ggccctccac ttcccagtg acagatccac agggtaaatc tgacatcccc atggacctgt	840
ttgacttcta tgagcaaatg gacaaggatg aagaagaaga agaagagaca cagacagtgt	900
cttttgaagt caagcaggaa atgattgagg aactccagaa acgttgcatc cacctggagt	960
accctctgtt ggcaagaatat gacttccgga atgattctgt caaccctgat atcaacattg	1020
acctaaagcc cacagctgtc ctgagaccct atcaggagaa gagcttgca aagatgtttg	1080
gaaacgggcg tgcacgttcg ggggtcattg ttcttccctg cgggtgctgga aagtccctgg	1140
ttggtgtgac tgctgcatgc actgtcagaa aacgctgtct ggtgctgggc aactcagctg	1200
tttctgtgga gcagtggaaa gccagttca agatgtgggtc caccattgac gacagccaga	1260
tctgccggtt cacctccgat gccaaggaca agcccatcgg ctgctccgtt gccattagca	1320
cctactccat gctggggccac accaccaaaa ggtcctggga ggccgagcga gtcattggagt	1380
ggctcaagac ccaggagtgg ggcctcatga tcctggatga agtgcacacc ataccagcca	1440
agatgttccg aagggtgctc accatcgtgc aggccactg taagctgggt ttgactgcga	1500
ccctcgtccg cgaagatgac aaaattgtgg atttaaattt tctgattggg cctaagctct	1560
acgaagccaa ctggatggag ctgcagaata atggctacat cgccaaagtc cagtgtgctg	1620
aggctctgggt ccctatgtct cctgaatttt accgggaata tgtggcaatc aaaaccaaga	1680
aacgaatctt gctgtacacc atgaaccca acaaatttag agcttgccag tttctgatca	1740
agtttcatga aaggaggaat gacaagatta ttgtctttgc tgacaatgtg tttgccctaa	1800
aggaatatgc cattcgactg aacaaaccct atatctacgg acctacgtct caggggggaaa	1860
ggatgcaaat tctccagaat ttcaagcaca accccaaaat taacaccatc ttcatatcca	1920
aggtaggtga cacttcgttt gatctgccgg aagcaaatgt cctcattcag atctcatccc	1980
atggtggctc caggcgtcag gaagcccaa ggctagggcg ggtgcttcga gctaaaaaag	2040

ggatggttgc agaaggtac aatgcctttt tctactcact ggtatcccag gacacacagg	2100
aaatggctta ctcaaccaag cggcagagat tcttggtaga tcaaggttat agcttcaagg	2160
tgatcacgaa actcgctggc atggaggagg aagacttggc gttttcgaca aaagaagagc	2220
aacagcagct cttacagaaa gtcctggcag ccactgacct ggatgccgag gaggaggtgg	2280
tggctgggga atttggctcc agatccagcc aggcattctcg gcgctttggc accatgagtt	2340
ctatgtctgg ggccgacgac actgtgtaca tggagtacca ctcacgcgg agcaaggcgc	2400
ccagcaaaca tgtacacccg ctcttcaagc gctttaggaa atgatgctta ggcagggtac	2460
ttcgttcaag accggcgctt ggcacccttg ttggaaaggg attttcagca taacattttc	2520
cttcacctc tttgacctc cctccagcgt tggccaaatt gtgctgagga agatgcatca	2580
agggcttggc tgtgccttca taggtcatct agggttttat aaaggaggag gagacaatat	2640
tttttcaaac tttttgggga gtgggggtcat ttctgtatat aaaaaatggt aatatttaag	2700
gtgtatttat gttaccgttc tgaataaaca gaatggacca ttgaaccagt	2750

<210> 31

<211> 6758

<212> DNA

<213> Homo sapiens

<220>

<223> >ERCC4|ENSG00000175595|ENST00000311895|6758

<400> 31

agagcttcca tggagtcagg gcagccggct cgacggattg ccatggcgcc gctgctggag	60
tacgagcgac agctggtgct ggaactgctc gacactgacg ggctagtagt gtgcgcccgc	120
gggctcggcg cggaccggct cctctaccac tttctccagc tgcactgcca cccagcctgc	180
ctggtgctgg tgctcaacac gcagccggcc gaggaggagt attttatcaa tcagctgaag	240
atagaaggag ttgaacacct ccctcgccgt gtaacaaatg aaatcacaag caacagtcgc	300
tatgaagttt acacacaagg tgggtgtata tttgcgaca gtaggatact tgtggttgac	360
ttcttgactg atagaatacc ttcagattta attactggca tcttggtgta tagagcccac	420
agaataatcg agtcttgtca agaagcattc atcttgcgcc tctttcgcca gaaaaacaaa	480
cgtggtttta ttaaagcttt cacagacaat gctgttgcct ttgatactgg tttttgtcat	540

gtggaaagag tgatgagaaa tctttttgtg aggaaactgt atctgtggcc aaggttccat	600
gtagcagtaa actcattttt agaacagcac aaacctgaag ttgtagaaat ccatgtttct	660
atgacaccta ccatgcttgc tatacagact gctatactgg acattttaaa tgcattgtcta	720
aaggaactaa aatgccataa cccatcgctt gaagtggaag atttatcttt agaaaatgct	780
attggaaaac cttttgacaa gacaatccgc cattatctgg atcctttgtg gcaccagctt	840
ggagccaaga ctaaattcctt agttcaggat ttgaagatat tacgaacttt gctgcagtat	900
ctctctcagt atgattgtgt cacatttctt aatcttctgg aatctctgag agcaacggaa	960
aaagcttttg gtcagaattc aggttggttg tttcttgact ccagcacctc gatgtttata	1020
aatgctcgag caagggttta tcattctcca gatgccaaaa tgagtaaaaa agaaaaata	1080
tctgaaaaaa tggaaattaa agaaggggaa gaaacaaaaa aggaactggt cctagaaagc	1140
aacccaaagt gggaggcact gactgaagta ttaaaagaaa ttgaggcaga aaataaggag	1200
agtgaagctc ttggtggtcc aggtcaagta ctgatttgtg caagtgatga ccgaacatgt	1260
tcccagctga gagactatat cactcttgga gcggaggcct tcttattgag gctctacagg	1320
aaaacctttg agaaggatag caaagctgaa gaagtctgga tgaaatttag gaaggaagac	1380
agttcaaaga gaattaggaa atctcacaaa agacctaaag acccccaaaa caaagaacgg	1440
gcttctacca aagaaagaac cctcaaaaag aaaaaacgga agttgacctt aactcaaagt	1500
gtaggaaaac ctgaagaact ggaagaggaa ggagatgtcg aggaaggata tcgtcgagaa	1560
ataagcagta gccagaaaag ctgcccggaa gaaattaagc atgaagaatt tgatgtaaat	1620
ttgtcatcgg atgctgcttt cggaatcctg aaagaacccc tcactatcat ccatccgctt	1680
ctgggttgca gcgacccta tgctctgaca aggggtactac atgaagtgga gccaaagatac	1740
gtggttcttt atgacgcaga gctaactttt gttcggcagc ttgaaattta cagggcgagt	1800
aggcctggga aacctctgag ggtttacttt cttatatacg gaggttcaac tgaggaacaa	1860
cgctatctca ctgctttgcg gaaagaaaag gaagcttttg aaaaactcat aagggaaaaa	1920
gcaagcatgg ttgtccctga agaaagagaa ggcagagatg aaacaaactt agacctagta	1980
agaggcacag catctgcaga tgtttccact gacactcgga aagccggtgg ccaggaacag	2040
aatggtacac agcaaagcat agttgtggat atgcgtgaat ttcgaagtga gcttccatct	2100
ctgatccatc gtcggggcat tgacattgaa cccgtgactt tagaggttgg agattacatc	2160

ctcactccag aaatgtgcgt ggagcgcaag agtatcagtg atttaatcgg ctctttaaat	2220
aacggccgcc tctacagcca gtgcatctcc atgtcccgtc actacaagcg tcccgtgctt	2280
ctgattgagt ttgaccctag caagcctttc tctctcactt cccgaggtgc cttgtttcag	2340
gagatctcca gcaatgacat tagttccaaa ctcactcttc ttacacttca cttccccaga	2400
ctacggattc tctggtgccc ctctcctcat gcaacggcgg agttgtttga ggagctgaaa	2460
caaagcaagc cacagcctga tgcggcgaca gcaactggcca ttacagcaga ttctgaaacc	2520
cttcccagat cagagaagta taatcctggg cccaagact tcttgttaaa aatgccaggg	2580
gtgaatgcca aaaactgccg ctcttctgatg caccacgtta agaacatcgc agaattagca	2640
gccctgtcac aagacgagct cacgagtatt ctggggaatg ctgcaaatgc caaacagctt	2700
tatgatttca ttcacacctc ttttgcagaa gtcgtatcaa aaggaaaagg gaaaaagtga	2760
acagtgatgg ctgttttctt atcccatgcc tgtacttttc agcggctcct tgccagacat	2820
cataggtcat tattaattat tggtttgcta tttcattctt ttccaatgct cttaatgatt	2880
gtacggtgga ccagaagcca ggattcctct ctgaactctg cagttaggca tcacttgaac	2940
ttgcctgtgc ctgctctttt tcctccctgc accgtctatg ccgggcttag catgtttctt	3000
tttaaatgag gtttgtcagg atcaggtaaa gttcctacaa gtgattacag aaggtagaaa	3060
ctttacctga tcctaacaga tctcatttag aaaggaatat gctaagcctg gcatggacgg	3120
tgcagggagg gaaaagagca ggcacaagaa agctaccatt tttaacagtc cttgttatct	3180
agtgaacat aaataacagt cttaatgca cttataccca tgtcctgtgg ctctccaaat	3240
ctggcttttg ctgttggtgc tgctggacgc ttgaactgat gtttgtgtag gaaatcatgt	3300
tctgaccctt tgtctacaaa ggagccttct ggaacactga gaagaaacat ctctttgcca	3360
ttcctgacca gttctctcta ccacattttc ttcagctcca tacttctgcc tgtctgctct	3420
aaggaaattt catggagcct tcctactact aattcaagac agtctcctca aaaactgggt	3480
gactagtctt ctaatgacct taacatatgt agcatatact ataatttcat tgttccaaat	3540
tagtattttt aaagcaaaat gaattacctg tttgcaaaag ttaatgatga aggagctctt	3600
agaatttctc atttttgcac atattcagtc tcctaataatc agagatccct aagtccagct	3660
ggctagttac agagtttttt cagacttcct cgtttctcag ctcttatatc ctaagacacc	3720
agcatcatat cctctagaaa tacaacctaa ttggcagtga gccgagatcg caccactgca	3780

cccctgcctg ggcgacagag tgagactttg tctctattac aaaaagaaaa gaaaagaaat	3840
acaacctaag ctcacctgcc tgtgattcct ctttctcac catcctgtgc cagggtggct	3900
acttctctct gtgaggactc aaatacaagc caatgagtgg ccactaagc ttttaagatt	3960
tgattttcct gccttgagat aaaaaggagt gtaggtaaat gaaagatcaa tgtatggaat	4020
atataaaaaat acgaaagaaa tatatacggt taaaaatcca taaagaaaaa aatctcattc	4080
taaacctgat taagttggct ttttacgtaa gtgtacaaat aggatattca cagcatcttt	4140
gtgcagtttt taaactttta ttttaaaaca ttattaagtt ggcttttggt cacatgttga	4200
gtaatgggta gtaaattttc tacctcagga gctgatatag acatcagttc tgctagccat	4260
atcacatatt ttaatgtttc atcaacatca gctgtttttt tgtttgctac actatttgaa	4320
ctaatagaca gtggatcatg taacacaaaa ttgtcttcaa ctttaacaaa attgtcattg	4380
ttattttttt ttgagacaga gtctccctct gttgccagg ctggagtgca gtggcgcaat	4440
ctcggctcac tgcaacctcg cctgctgggt tcaagcagtt ctcccacctc agcctcccaa	4500
gtagctggga ttataggtgt gcaccaccag acccggttaa tttttgtatt tttagtagag	4560
acggggtttc accatgttgg ccaggctgggt ctcaaactac tgacctcagg tgatccaccc	4620
accttggcct cccaaagtgc tgagataaca ggcgtgagcc accgcacca cccaacaaaa	4680
ttatttttta ctgtcgtttc tgaactgaac ctctcacatt ggcatcttga tttattggta	4740
cttaacagta tatatggatt ttgaactcta cacaagggtta taaccagaat gaattgaggg	4800
gtatcaagaa ccagattgggt tacagtagaa ctctgtaaga tgatgtggag agtaaggaaa	4860
aggagaagaa ataaaaatta gtttaagatg gaataaagat ttgggctgct aattttttcc	4920
caggattcaa aatatacccg ctagaatgga aaacaaaaat ttgaatgatt acaacattat	4980
cgtaaacatt caattttgta tttattgttt tattgaatat agttcagagt atattaaaaat	5040
agacctacca cttcaaatga taatgattat tataatgtct cttcaccta ttctagcact	5100
tctgcttgca gtatgtgggt tctatttttt tcctcttgta taattccact tgcttttaat	5160
tgttgtttca ttatataaaa ggaacatctt cccatagcat attctatgaa aggggtttca	5220
ttccaagttg agttttcaaa aaaaaggctc tcctaaagct accattttca accgtccttg	5280
ttatctagta caacataaat aacagtctta aaaattgcac taataccagt gccccctgg	5340
ctctccaaat ctgttctttg ctcttgatc tgctggacgc ttgaagacag gtgcactgtc	5400

tcgtatgtat ttgaattatg aacagtaatt tctaataaat tctaaaatgg tcattgtaag	5460
tgaaagcctc tcgctaccac ttcctcttcc aactacataa atatatttca atgtatttcc	5520
agttttggaa agttttcaat acatacatca agtggtttact tagattttta taaaaat	5580
ttttacaatc taataatctt tggtaaagga actagagatg catgcagttg caaaattaat	5640
gtatttattt tccagcataa ttttattaac ttcacttttt ttctctctag taaatatcca	5700
gtgtacttat gaactcatgt ttggctcttt taaaaccttt tctaaaagct agatcagcat	5760
ttttctattt tacaagtttt ttgtataaaa aggtgaacat atgagatatt gtgagaaatc	5820
attttaagtt tcattttaa at catggtgcct cttttgatac tttcttaaaa ttgtgcaaga	5880
agaaatcatt tttagtagtg gtcataaata ttatcctttt ggcagtaagc tattactaat	5940
tcagcctgaa gctcggtaga caatatgtct acatgtgttt gagtacatcc tggatacagt	6000
ttcccagctc atgagggact gaaaatagtc ttattcatca acccactagg atgtgaaggg	6060
ttaagtctag atttggtcgc attgaaaccc cacaatatag aattaataaa tggccttcag	6120
taggaaaacc tacactaaag caaatccgaa gaagtggggc agggggaaaag aggcattact	6180
ggtctttcct tttgttttgc aagcataatt tgattttcct ttggctcaga aaactcattt	6240
ggggaaattc tcttttgtgt tcagtttaac ctagaaaggt cctcttgaaa aaccaacatt	6300
ttaggaaagt tcttttttca ggatggagta tattaaaatt aagccaggct ttgacgtgaa	6360
ttatcacttt tctttattat tttgttttct attttggttt atagctattt ctggttcagt	6420
tctgaacttc agcacttaat catccttatc aaccaggctt ttggtagcct aaaccgctat	6480
gctgttgttt ttttaattta aagatgtata agccaaaatt tggatgggag tgagacataa	6540
ctgatttata tgaattttta cagagttgta tttgtgtgtg ttttaataaaa tatatattta	6600
ttcagtactt tcctcagtat tttatgggca aagtaaaaat aacaatgcat agtgaaaggg	6660
catatattac cagcagtaat aattcaaaat cctgaaaatg tttcattttt tttgtttttg	6720
ttatgcagaa taaacaaggc agaaatgctc tttgaacc	6758

<210> 32

<211> 5076

<212> DNA

<213> Homo sapiens

<220>

<223> >ERCC5|ENSG00000134899|ENST00000355739|5076

<400> 32

tttgccaacg ttaaggacgc gcagctgtga cacagcccca ggaagtccag atgacatgtg	60
cccaaggtgg ttggggcaca gattggtttt atacatttta gggagacagg agacatcaat	120
caacatatgt aagtacactg gttccttcca gaaaggtggg gacaactcgg aagcaggaag	180
ggcttctagg tcacaggtag atgagagaca aaaggctgca tacgagtttc tgataagcct	240
ttccaaagga gacaatcaga atatgcatct atctcagtga gcagaaggat gactgactag	300
aatgggaggc aggttttgcc ctgagcagtt cccagcttga cttttccctt ttgcttagta	360
attttgggac cctaacattt tcacaggctt taaattttat tattcttttag ttactacgtg	420
ctagcatata aataaatagt aaaaaaccaa gaaggcatcc accttttggg tgtctcttca	480
cgtgtaaaac aacactttgt gttaagtatc ttcacacacg gcggcgcaaa ggtagaaacc	540
gatactaaaa aagcgtgtag aaaatagttc ccagcctggg caacacaggg agacctcatt	600
tctacaaaaa taattcgcca agcatagtgg tgcgcacctg cgggtcccagc tacttgagag	660
gctgagatgg gaaagttgct tgagctcggg cagcaggagt tccaggctgc agtgagctaa	720
gaatgcgcca ctggactcca gtatgggcca cagcgtgaga ccctgtctca aacaaaaaca	780
aaagcccgtt actccaccaa gaaggcgctt ttgcacattg ttttaatgct taacgccttc	840
aggatgccag cgtgacggaa gcaagtaacc accaaggcat caccactggc gctaaacttc	900
tcacttccgg agtgctgcaa gcgcagaaaa tatacgtcat gtgcggaggc ggagcttccg	960
ccctgcgcgt cgtattagac ggaaaccgag cgggcccatt tttcatgggt ttgcggaccc	1020
accagcgaag gcgggagggt tcgcaggac atcttctggc tgtttccgtc gcctgcgtgg	1080
cccttgcacc ccggtcttcc attagcggcg cagacgtttg ggcctaagcg ctgggcgagg	1140
cgaggccctg cccctccccg ccaacggcca ttctctggac ctgtctttct tccgggaggc	1200
ggtgacagct gctgagacgt gttgcagcca gagtctctcc gctttaatgc gctccatta	1260
gtgccgtccc ccactggaaa accgtggctt ctgtattatt tgccatcttt gttgtgtagg	1320
agcagggagg gcttcctccc ggggtcctag gcggcggtgc agtccgtcgt agaagaatta	1380
gagtagaagt tgtcggggtc cgctcttagg acgcagccgc ctcagggggg tccaggggct	1440
ctggaagctg ctggagtgtc ccgggcggca ggtcagcccc gaagcgctgg aagggaagat	1500
cctggctgtt gatattagca tttggttaaa ccaagcactt aaaggagtcc gggatcgcca	1560

tggaactca atagaaaatc ctcatcttct cactttgttt catcggctct gcaaactctt	1620
attttttcga attcgtccta tttttgtgtt tgatggggat gctccactat tgaagaaaca	1680
gactttggtg aagagaaggc agagaaagga cttagcgtcc agtgactcca ggaaaacgac	1740
agagaagctt ctgaaaacat ttttgaaaag acaagccatc aaaactgcct tcagaagcaa	1800
aagagatgaa gcactaccca gtcttaccca agttcgaaga gaaaacgacc tctatgtttt	1860
gcctccttta caagaggaag aaaaacacag ttcagaagag gaagatgaaa aagaatggca	1920
agaaagaatg aatcaaaaac aagcattaca ggaagagttc tttcataatc ctcaagcgat	1980
agatattgag tctgaggact tcagcagcct gccccctgaa gtaaagcatg aaatcttgac	2040
tgatatgaaa gagttcacca agcgcagaag aacattattt gaagcaatgc cagaggagtc	2100
tgatgacttt tcacagtacc aactcaaagg cttgcttaaa aagaactatc tgaaccagca	2160
tatagaacat gtccaaaagg aaatgaatca gcaacattca ggacacatcc gaaggcagta	2220
tgaagatgaa gggggctttc tgaaggaggt agagtcaagg agagtgggtct ctgaagacac	2280
ttcacattac atcttgataa aaggtattca agctaagaca gttgcagaag tggattcaga	2340
gtctcttcct tcttcagca aaatgcacgg catgtctttt gacgtgaagt catctccatg	2400
tgaaaaactg aagacagaga aagagcctga tgctaccctt ctttctcaa gaactttact	2460
agctatgcaa gctgccctgc tgggaagtag ctgagaagag gagctggaga gtgaaaatcg	2520
aaggcaggcc cgtggggagga acgcacctgc tgctgtagac gaaggctcca tatcaccccg	2580
gactctttca gccattaaga gagctcttga cgatgacgaa gatgtaaaag tgtgtgtctgg	2640
ggatgatgtg cagacgggag ggccaggagc agaagaaatg cgtataaaca gctccaccga	2700
gaacagtgat gaaggactta aagtgagaga tggaaaagga ataccgttta ctgcaacact	2760
tgcgatcatct agtgtgaact ctgcagagga gcacgtagcc agcactaatg aggggagaga	2820
gccacagac tcagttcaa aagaacaaat gtcacttggt cacgtgggga ctgaagcctt	2880
tccgataagt gatgagtcta tgattaagga cagaaaagat cggctgcctc tggagagtgc	2940
agtggttaga catagtgacg cacctgggct cccgaatgga agggaaactga caccggcatc	3000
tccaacttgt acaaattctg tgtcaaagaa tgaaacacat gctgaagtgc ttgagcagca	3060
gaacgaactt tgcccatatg agagtaaatt cgattcttct cttctttcaa gtgatgatga	3120
aacaaaatgt aaaccgaatt ctgcttctga agtcattggc cctgtcagtt tgcaagaaac	3180

aagtagcata gtaagtgtcc cttcagaggc agtagataat gtggaaaatg tgggtgtcatt	3240
taatgctaaa gagcatgaga attttctgga aaccatccaa gaacagcaga ccaactgaatc	3300
tgcaaggccag gatttaattt ccattccaaa ggccgtggaa ccaatggaaa ttgactcgga	3360
agaaagtgaa tctgatggaa gtttcattga agtgcaaagt gtgattagtg atgaggaact	3420
tcaagcagaa ttccctgaaa cttccaaacc tccctcagaa caaggcgaag aggaactggt	3480
aggaactagg gagggagaag cccctgctga gtccgagagc ctcctgaggg acaactctga	3540
gagggacgac gtggatggtg agccacagga agctgagaaa gatgcggaag attcgctcca	3600
tgaatggcaa gatattaatt tggaggagt tggaaactctg gagagcaacc tcttagcaca	3660
gcagaattca ctgaaagctc aaaaacagca gcaagaacgg atcgctgcta ctgtcaccgg	3720
acagatgttc ctggaaagcc aggaactcct gcgcctgttc ggcatccctt acatccaggc	3780
tcccatggaa gcagaggcgc agtgcgccat cctggacctg actgatcaga cttccggaac	3840
catcactgat gacagtgata tctggctgtt tggagcgcgg catgtctata gaaacttttt	3900
taataaaaaac aagttttag aatattatca atatgtggac tttcacaatc aattgggatt	3960
ggaccggaat aagttaataa atttggctta tttgcttgga agtgattata ccgaaggaat	4020
accaactgtg ggttgtgtaa ccgccatgga aattctcaat gaattccctg ggcattgcct	4080
ggaacctctc ctaaaattct cagaatggtg gcatgaagct caaaaaaatc caaagataag	4140
acctaattcct catgacacca aagtgaaaaa aaaattacgg acattgcaac tcacccttg	4200
ctttcctaac ccagctgttg ccgaggccta cctcaaacc gtggtggatg actcgaaggg	4260
atcctttctg tgggggaaac ctgatctcga caaaattaga gaattttgtc agcgttattt	4320
cggctggaac agaacgaaga cagatgaatc tctgtttcct gtattaaagc aactcgatgc	4380
ccagcagaca cagctccgaa ttgattcctt ctttagatta gcacaacagg agaaagaaga	4440
tgctaaacgt attaagagcc agagactaaa cagagctgtg acatgtatgc taaggaaaga	4500
gaaagaagca gcagccagcg aaatagaagc agtttctgtt gccatggaga aagaatttga	4560
gctacttgat aaggcaaaag gaaaaacca gaagagaggc ataacaaata ccttagaaga	4620
gtcatcaagc ctgaaaagaa agaggctttc agattctaaa ggaaagaata catgcggtgg	4680
atttttgggg gagacctgcc tctcagaatc atctgatgga tcttcaagtg aagatgctga	4740
aagttcatct ttaatgaatg tacaaaggag aacagctgcg aaagagccaa aaaccagtgc	4800

ttcagattcg cagaactcag tgaaggaagc tcccgtgaag aatggaggtg cgaccaccag	4860
cagctctagt gatagtgatg acgatggagg gaaagagaag atggtcctcg tgaccgccag	4920
atctgtgttt gggaagaaaa gaaggaaact aagacgtgcg aggggaagaa aaaggaaaac	4980
ctaattaaaa aatatgtatc ctctataatt agttatgaca gccatttgta atgaatttgt	5040
cgcaaagacg taataaaatt aactggtggc acggtc	5076

<210> 33
 <211> 2869
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FAM175A|ENSG00000163322|ENST00000321945|2869

<400> 33	
gccggcgag gccccgcctc ctcgctgggc cccgcctcgc tgccaccaca gggctttgcc	60
tccgcgcgcc ccgccctcgt cctcttgtgt agcctgaggc ggcggtagca tggaggggga	120
gagtacgtcg gcgggtgctct cgggctttgt gctcggcgca ctcgctttcc agcacctcaa	180
cacggactcg gacacggaag gttttcttct tggggaagta aaaggtgaag ccaagaacag	240
cattactgat tcccaaattg atgatgttga agttgtttat acaattgaca ttcagaaata	300
tattccatgc tatcagcttt ttagctttta taattcttca ggcgaagtaa atgagcaagc	360
actgaagaaa atattatcaa atgtcaaaaa gaatgtggta ggttgggtaca aattccgtcg	420
tcattcagat cagatcatga cgtttagaga gaggctgctt cacaaaaact tgcaggagca	480
tttttcaaac caagaccttg tttttctgct attaacacca agtataataa cagaaagctg	540
ctctactcat cgactggaac attccttata taaacctcaa aaaggacttt ttcacagggt	600
acctttagt gttgccaatc tgggcatgtc tgaacaactg gggtataaaa ctgtatcagg	660
ttcctgtatg tccactgggt ttagccgagc agtacaacaa cacagctcta aattttttga	720
agaagatgga tccttaaagg aggtacataa gataaatgaa atgtatgctt cattacaaga	780
ggaattaaag agtatatgca aaaaagtgga agacagtga caagcagtag ataaactagt	840
aaaggatgta aacagattaa aacgagaaat tgagaaaagg agaggagcac agattcaggc	900
agcaagagag aagaacatcc aaaaagaccc tcaggagaac atttttcttt gtcaggcatt	960

acggaccttt tttccaaatt ctgaatttct tcattcatgt gttatgtctt taaaaaatag	1020
acatgttttct aaaagtagct gtaactacaa ccaccatctc gatgtagtag acaatctgac	1080
cttaatggta gaacacactg acattcctga agctagtcca gctagtacac cacaaatcat	1140
taagcataaa gccttagact tagatgacag atggcaattc aagagatctc ggttgttaga	1200
tacacaagac aaacgatcta aagcagatac tggtagtagt aaccaagata aagcatccaa	1260
aatgagcagc ccagaaacag atgaagaaat tgaaaagatg aagggttttg gtgaatattc	1320
acggctctcct acattttgat ccttttaacc ttacaaggag atttttttat ttggctgatg	1380
ggtaaagcca aacatttcta ttgtttttac tatgttgagc tacttgcagt aagttcattt	1440
gtttttacta tgttcacctg tttgcagtaa tacacagata actcttagtg cattttacttc	1500
acaaagtact ttttcaaaca tcagatgctt ttattttcaa accttttttt cacctttcac	1560
taagttgttg aggggaaggc ttacacagac acattcttta gaattggaaa agtgagacca	1620
ggcacagtgg ctcacacctg taatcccagc acttagggaa gacaagtcag gaggattgat	1680
tgaagttagg agttagagac cagcctgggc aacgtattga gaccatgtct attaaaaaat	1740
aaaatggaaa agcaagaata gccttatttt caaaatatgg aaagaaattt atatgaaaat	1800
ttatctgagt cattaaaatt ctccttaagt gatacttttt tagaagtaca ttatggctag	1860
agttgccaga taaaatgctg gatatcatgc aataaatttg caaaacatca tctaaaattt	1920
aaatttgtcc tacatttgta cttgctaaat ctggcagcct taattattac acaactataa	1980
cacaacacaa acagtaatct ctcccctagt ttagagtcct ttgcatgaaa tttagcaaat	2040
ctatcatttt gaatacagta caagcctggc aatccagtgt atttgggata ggccagtgga	2100
actcagaact gatacatata tacatgtatg tctatcatat tcaatgtcta ttaggtgtaa	2160
atcatctctt catgtaaaag acactactgg tttgagggta gactcgtcta ggaatatttt	2220
gtccttctgt gacagcacca ctatccatcc caatgtaaca tttctacaaa gggaagatat	2280
cttaaccagg catacttata tttgctagtt ggaaaatcat tctgaacagt ttttaattcca	2340
ttctattgaa tggattattg tatgagctat ttggggatca gatataaatg acaatcacac	2400
agaaccactt ttattcaaca tagataaaat gtgatgagtg ttaacatttt aaataagcaa	2460
tctgcttaat ggacattggc atgaagggga agtgtcattc tttacctcac tcctcaaaca	2520
cttttctatc caaaggttgg tttgagtcag tattggcatc atacaaccac ttactgtaaa	2580

ataagtttat tagtggtaac gtgatagaat ttattcccga tatctgatgt tacaaacttt	2640
agggtccttg agatatgcag gatccttgta tgtaactggc ataaaccctg taaagagata	2700
attaaagttc ttgacaaaac ataatgagca aaatttcagc tagctcaatt cagttgttga	2760
ggtattcttt cttaccatt atttgagctc tcttaattgc ttttgtgatt tctttctgtt	2820
tcttcccaca aagacctgta aaataaaaag ctttcttatt cataaaaaa	2869

<210> 34
 <211> 5451
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FANCA|ENSG00000187741|ENST00000389301|5451

<400> 34	
agccgccgcc ggggctgtag gcgccaaggc catgtccgac tcgtgggtcc cgaactccgc	60
ctcgggccag gaccagggg gccgccggag ggcctgggcc gagctgctgg cgggaagggt	120
caagagggaa aaatataatc ctgaaagggc acagaaatta aaggaatcag ctgtgcgcct	180
cctgcgaagc catcaggacc tgaatgccct tttgcttgag gtagaaggtc cactgtgtaa	240
aaaattgtct ctgagcaaag tgattgactg tgacagttct gaggcctatg ctaatcattc	300
tagttcattt ataggctctg ctttgcagga tcaagcctca aggctggggg ttcccgtggg	360
tattctctca gccgggatgg ttgcctctag cgtgggacag atctgcacgg ctccagcgga	420
gaccagtcac cctgtgctgc tgactgtgga gcagagaaag aagctgtctt ccctgttaga	480
gtttgctcag tatttattgg cacacagtat gttctccgt ctttccttct gtcaagaatt	540
atggaaaata cagagttctt tgttgcttga agcgggtgtgg catcttcacg tacaaggcat	600
tgtgagcctg caagagctgc tggaaagcca tcccgacatg catgctgtgg gatcgtggct	660
cttcaggaat ctgtgctgcc tttgtgaaca gatggaagca tcctgccagc atgctgacgt	720
cgccagggcc atgctttctg attttgttca aatgtttgtt ttgaggggat ttcagaaaaa	780
ctcagatctg agaagaactg tggagcctga aaaaatgccg caggtcacgg ttgatgtact	840
gcagagaatg ctgatttttg cacttgacgc tttggctgct ggagtacagg aggagtcctc	900
cactcacaag atcgtgaggt gctggttcgg agtgttcagt ggacacacgc ttggcagtgt	960
aatttccaca gatcctctga agaggttctt cagtcatacc ctgactcaga tactcactca	1020

cagccctgtg ctgaaagcat ctgatgctgt tcagatgcag agagagtgga gctttgcgcg	1080
gacacaccct ctgctcacct cactgtaccg caggctcttt gtgatgctga gtgcagagga	1140
gttggttggc catttgcaag aagttctgga aacgcaggag gttcactggc agagagtgct	1200
ctcctttgtg tctgccctgg ttgtctgctt tccagaagcg cagcagctgc ttgaagactg	1260
ggtggcgcgt ttgatggccc aggcattcga gagctgccag ctggacagca tggtcactgc	1320
gttcctgggt gtgcgccagg cagcactgga gggccctct gcgttcctgt catatgcaga	1380
ctggttcaag gcctccttg ggagcacacg aggctaccat ggctgcagca agaaggccct	1440
ggtcttcctg tttacgttct tgtcagaact cgtgcctttt gagtctcccc ggtacctgca	1500
ggtgcacatt ctccaccac ccctggttcc cggcaagtac cgctccctcc tcacagacta	1560
catctcattg gccaaagacac ggctggccga cctcaagggt tctatagaaa acatgggact	1620
ctacgaggat ttgtcatcag ctggggacat tactgagccc cacagccaag ctcttcagga	1680
tgttgaaaag gccatcatgg tgtttgagca tacggggaac atcccagtca ccgtcatgga	1740
ggccagcata ttcaggaggc cttactacgt gtcccacttc ctccccgcc tgctcacacc	1800
tcgagtgtc cccaaagtcc ctgactcccg tgtggcgttt atagagtctc tgaagagagc	1860
agataaaatc ccccatctc tgtactccac ctactgccag gcctgctctg ctgctgaaga	1920
gaagccagaa gatgcagccc tgggagtgag ggcagaaccc aactctgctg aggagcccct	1980
gggacagctc acagctgcac tgggagagct gagagcctcc atgacagacc ccagccagcg	2040
tgatgttata tcggcacagg tggcagtgat ttctgaaaga ctgagggtg tcctgggcca	2100
caatgaggat gacagcagcg ttgagatatc aaagattcag ctcagcatca acacgccgag	2160
actggagcca cgggaacaca tggctgtgga cctcctgctg acgtctttct gtcagaacct	2220
gatggctgcc tccagtgtcg ctccccgga gaggcagggt ccctgggctg ccctcttcgt	2280
gaggaccatg tgtggacgtg tgctccctgc agtgctcacc cggtctgcc agctgctccg	2340
tcaccagggc ccgagcctga gtgccccaca tgtgctgggg ttggctgccc tggccgtgca	2400
cctgggtgag tccaggtctg cgctcccaga ggtggatgtg ggtcctcctg cacctggtgc	2460
tggccttcct gtccctgcgc tctttgacag cctcctgacc tgtaggacga gggattcctt	2520
gttcttctgc ctgaaatctt gtacagcagc aatttcttac tctctctgca agttttcttc	2580
ccagtcacga gatactttgt gcagctgctt atctccaggc cttattaaaa agtttcagtt	2640

cctcatgttc agattgttct	cagaggcccg acagcctctt	tctgaggagg acgtagccag	2700
cctttcctgg agacccttgc	accttccttc tgcagactgg	cagagagctg ccctctctct	2760
ctggacacac agaaccttcc	gagaggtgtt gaaagaggaa	gatgttcact taacttacca	2820
agactggtta cacctggagc	tggaaattca acctgaagct	gatgctcttt cagatactga	2880
acggcaggac ttccaccagt	gggcgatcca tgagcacttt	ctccctgagt cctcggcttc	2940
agggggctgt gacggagacc	tgcaggctgc gtgtaccatt	cttgtcaacg cactgatgga	3000
tttccaccaa agctcaagga	gttatgacca ctcaaaaaat	tctgatttgg tctttggtgg	3060
ccgcacagga aatgaggata	ttatttccag attgcaggag	atggtagctg acctggagct	3120
gcagcaagac ctcatagtgc	ctctcggcca cacccttcc	caggagcact tcctctttga	3180
gattttccgc agacggctcc	aggctctgac aagcgggtgg	agcgtggctg ccagccttca	3240
gagacagagg gagctgctaa	tgtacaaacg gatcctcctc	cgctgcctt cgtctgtcct	3300
ctgcggcagc agcttccagg	cagaacagcc catcactgcc	agatgcgagc agttcttcca	3360
cttgggtcaac tctgagatga	gaaacttctg ctcccacgga	ggtgccctga cacaggacat	3420
cactgcccac ttcttcaggg	gcctcctgaa cgcctgtctg	cggagcagag acccctccct	3480
gatggctcac ttcatactgg	ccaagtcca gacgaaatgc	cccttaattt tgacctctgc	3540
tctgggtgtgg tggccgagcc	tggagcctgt gctgctctgc	cgggtggagga gacactgcc	3600
gagcccgtg ccccggaac	tgcagaagct acaagaaggc	cggcagtttg ccagcgattt	3660
cctctcccct gaggtgcct	ccccagcacc caaccggac	tggctctcag ctgctgcact	3720
gcactttgcg attcaacaag	tcaggaaga aaacatcagg	aagcagctaa agaagctgga	3780
ctgcgagaga gaggagctat	tggttttcct tttcttcttc	tccttgatgg gcctgctgtc	3840
gtcacatctg acctcaaata	gcaccacaga cctgccaaag	gctttccacg tttgtgcagc	3900
aatcctcgag tgtttagaga	agaggaagat atcctggctg	gcactctttc agttgacaga	3960
gagtgcctc aggctggggc	ggctcctcct ccgtgtggcc	ccggatcagc acaccaggct	4020
gctgcctttc gctttttaca	gtcttctctc ctacttccat	gaagacgcgg ccatcaggga	4080
agaggccttc ctgcatgttg	ctgtggacat gtacttgaag	ctggtccagc tcttcgtggc	4140
tggggataca agcacagttt	cacctccagc tggcaggagc	ctggagctca agggtcaggg	4200
caaccccgctg gaactgataa	caaaagctcg tctttttctg	ctgcagttaa tacctcggtg	4260

cccgaaaaag agcttctcac acgtggcaga gctgctggct gatcgtgggg actgcgaccc	4320
agaggtgagc gccgccctcc agagcagaca gcaggctgcc cctgacgctg acctgtccca	4380
ggagcctcat ctcttctgac gggacctgcc actgcacacc agcccagctc ccgtgtaaat	4440
aatttattac aagcataaca tggagctctt gttgcactaa aaagtggatt acaaatctcc	4500
tcgactgctt tagtggggaa aggaatcaat tatttatgaa ctgtccggcc ccgagtcact	4560
cagcgtttgc gggaaaataa accactggtc ccagagcaga ggaaggctac ttgagccgga	4620
caccaagccc gcctccagca ccaagggcgg gcagcacctt ccgaccctcc catgcgggtg	4680
cacacgaagg gtgaggctga cacagccact gcggagtcca ggctgctaga ggtgctcatc	4740
ctcactgccg tcctcagggtg ggttcgggct tcaccgcctg gccctctgtg gtcacagagg	4800
ggctcgggtg gccaggtggt ggttccgcct ccaggggcag ggccctgtcc tgggtctgtg	4860
tcagcgggtg caccatggac atgtgtacat tgaggttgtg ggcccttctca aaccgccggc	4920
cacactggtc acaggcaaag tccagctcag tctcagcctt gtgtttggtc atgtggtact	4980
tgagggatgc ccgctgcctg cactggaacc cacagacctc acacctgggg gacagaggca	5040
gataagaagg tgcgagggcc acagccctgg gagggggtcc tgactcacac ttactgaaa	5100
ggcttggtc ccgaatgtcg catttggtg acgagaaggt gcttccgctg cttgaaggtt	5160
tgtccacatt cgtcacagat atagttccgc acctctgaga ggggagagtc cagtgagtcc	5220
aggcccctga tgctccaacc tcccgggggg acgacgatga caatgtgaaa ccatcacagc	5280
tgggaagaca tttctgcaca tggttcacca tgcagtgggc ccaagcaagg ggcctatgag	5340
ggcctcgttt attaagatct ttaaactgct ttatacactg tcacgtggct tcatcagctg	5400
tgtgcatttc aggatggttt ttaaagaaac ctcaaaaagc tatttcctta a	5451

<210> 35

<211> 3008

<212> DNA

<213> Homo sapiens

<220>

<223> >FANCB|ENSG00000181544|ENST00000398334|3008

<400> 35

agcgcgctgg cgggaggttt ggagcagatg gataccgtat cgacgtgggg cctccggtat	60
---	----

gttgccgctg cgttgagttt cataagtaca agccagacct ttgatttacc taccctatct	120
tctttactga tgaagctgaa actactttgt gatgcctatt gtcccagatc tttctgttaa	180
ttgcctaaca acaaaccagt ttagaaacat caactggaat caatttgcat ttgaaaatta	240
gatcagcaca aagacttgat attgtggaat gactagcaaa caagcaatgt catctaacga	300
acaagaaagg ctcttggtgtt ataatgggga agtccttggt ttccagttgt ctaaaggaaa	360
ttttgcagat aaagagccta caaaaacacc catattacat gtcagaagaa tgggtatttga	420
cagaggaaca aaagtatttg ttcagaagtc cactggattt ttaccataa aggaagaaaa	480
ctctcattta aaaatcatgt gttgcaactg tgtgtcagat ttcagaactg gaattaacct	540
cccttacatt gtgatagaaa aaaataaaaa gaataatgtt tttgaatatt ttttactaat	600
ccttcacagt actaataaat ttgaaatgcg tttgagtttt aaactaggct atgagatgaa	660
ggatggccta agggtcctta atggcccttt aattttatgg aggcattgtca aagcattctt	720
ctttatctct tctcaaaactg gcaaagttgt tagtgtgtca ggtaactttt cctctattca	780
gtgggcaggg gagattgaaa atttaggtat ggttttattg ggactaaagg aatgtttgtt	840
atctgaggaa gaatgtactc aagagccttc aaaatcagat tatgcaattt ggaataccaa	900
attttgtgta tattctcttg aaagtcaaga agtattaagt gatatatata ttattcctcc	960
tgcttacagc agtgtggtga cttatgtaca tatttgtgca actgagatca tcaaaaacca	1020
gttaagaata tctctcattg cccttactcg aaagaatcag ctgatttcat ttcagaatgg	1080
aactcctaaa aatgtgtgcc agcttcatt tggagatcct tgtgcagttc aacttatgga	1140
ttcaggtgga ggaaacctct ttttcgttgt atcctttata tccaataatg cttgtgtgt	1200
atggaaagag agctttcagg ttgctgctaa atgggaaaaa cttagcttag tactgataga	1260
tgactttatt ggaagtggaa ctgaacaagt actcctactt ttttaaggact ctttgaactc	1320
agactgcctg acttcattta aaataacgga tcttgaaaaa ataaactatt cgagtgaacc	1380
atcagattgc aatgaagatg acttatttga agacaaacaa gagaatcggt acctggtggt	1440
tccacctcta gaaacaggac tgaaagtttg tttttcttct tttcgggaat tacggcagca	1500
tctgttgctt aaggaaaaaa ttatttcaaa atcttataaa gctttaataa acctagttca	1560
aggaaaagat gataatacgt caagtgcaga ggagaaggaa tgtcttggtc ctctttgtgg	1620
tgaagaagaa aattctgtcc atatcttaga tgaaaagtta tcagacaatt ttcaagattc	1680

agaacagcta gtagagaaga tatggtatcg tgtaatagat gatagcttgg ttgttggagt	1740
gaaaactaca tcttctttga agctgtccct gaatgatgtg actttatcat tgттаатgga	1800
tcaagcccat gactccagat ttcggcttct aaagtgtcaa aatagggtga ttaagttgag	1860
tacaaatcct ttcccagcac catacttgat gccatgtgaa ataggattgg aagcaaaaag	1920
ggtcacgttg acccctgata gcaagaaaga ggaaagcttt gtttgtgaac acccatctaa	1980
gaaagagtgt gtacagataa ttactgctgt aacatctctt tcaccacttt taacattcag	2040
taaattttgt tgcactgtac tgctacaaat tatggagaga gaaagtggta actgtcctaa	2100
agatcgttat gttgtgtgtg gcagagtttt ttttaagtcta gaagatcttt caactgggaa	2160
gtacctactg acattttcaa agaagaaacc tatagagcac atggaagatc tttttgcact	2220
tcttgacgca ttccataaat cttgttttca aatcacatca cccggctatg ccctgaattc	2280
aatgaagggtg tggctcttag aacatatgaa atgtgaaata atcaaagaat ttccagaagt	2340
gtacttttgt gaaagaccgg gaagtttcta tgggacactc ttcacttgga aacagagaac	2400
accattcgaa gggattttaa taatctattc caggaatcaa acagttatgt tccagtgcct	2460
tcataatctc atcagaattc tccctataaa ctgtttcctc aaaaatctaa aatcaggaag	2520
tgagaatttc ctaattgata atatggcatt tactttggag aaggaactag tcacccttag	2580
ttctctttct tctgccatag ctaaacatga aagcaatttt atgcagaggt gtgaagtgag	2640
caaaggaaag agtagtgtcg tcgcggtcgc tttatcagac agaagggaaa atatccatcc	2700
ctacagaaaa gaacttcaga gagaaaagaa gaaaatgttg caaacgaacc taaaagttag	2760
tggtgccctt tacagagaaa taactttgaa agtagctgag gttcagttga aatcagactt	2820
tgctgcacag aaactgagta atttataatt ataatttcaa tttgatcata ttttaaaata	2880
tgttttcacc accatataag attttggttc tacttgctat atgcctcctt tgtaaaaata	2940
aacaccgagg cttactgtag tagaaacttt agttttgatg atgcacaaaa taaacagcag	3000
tggttctc	3008

<210> 36

<211> 4585

<212> DNA

<213> Homo sapiens

<220>

<223> >FANCC|ENSG00000158169|ENST00000289081|4585

<400> 36

actgctgaca cgtgtgcgcg cgcgcgggctc cactgccggg cgaccgcggg aaaattccaa	60
aaaaactcaa aaagccaata cgaggcaaag ccaaattttc aagccacaga tcccgggcgg	120
tggcttcctt tccgccactg cccaaactgc tgaagcagct cccgcgagga ccacccgatt	180
taatgtgtgc cgaccatttc cttcagtgcg ggacaggctg ctgtgaaggg acatcacctt	240
ttcgcttttt ccaagatggc tcaagattca gtagatcttt cttgtgatta tcagttttgg	300
atgcagaagc tttctgtatg ggatcaggct tccactttgg aaaccagca agacacctgt	360
cttcacgtgg ctcaattcca ggagttccta aggaagatgt atgaagcctt gaaagagatg	420
gattctaata cagtcattga aagattcccc acaattggct aactgttggc aaaagcttgt	480
tggaatcctt ttatttttagc atatgatgaa agccaaaaaa ttctaatatg gtgcttatgt	540
tgtctaatta acaaagaacc acagaattct ggacaatcaa aacttaactc ctggatacag	600
ggtgtattat ctcatatact ttcagcactc agatttgata aagaagttgc tcttttact	660
caaggctctg ggtatgcacc tatagattac tatcctgggt tgcttaaaaa tatggtttta	720
tcattagcgt ctgaactcag agagaatcat cttaatggat ttaacactca aaggcgaatg	780
gctcccagc gagtggcgct cctgtcacga gtttgtgtcc cacttattac cctgacagat	840
gttgaccccc tgggtggaggc tctcctcatc tgtcatggac gtgaacctca ggaaatcctc	900
cagccagagt tctttgaggc tgtaaacgag gccattttgc tgaagaagat ttctctcccc	960
atgtcagctg tagtctgcct ctggcttcgg caccttccca gccttgaaaa agcaatgctg	1020
catctttttg aaaagctaata ctccagtgc agaaattgtc tgagaaggat cgaatgcttt	1080
ataaaagatt catcgctgcc tcaagcagcc tgccaccctg ccatattccg ggttggtgat	1140
gagatgttca ggtgtgcact cctggaaacc gatggggccc tggaatcat agccactatt	1200
caggtgttta cgagtgctt tgtagaagct ctggagaaag caagcaagca gctgcggttt	1260
gcactcaaga cctactttcc ttacacttct ccatctcttg ccatggtgct gctgcaagac	1320
cctcaagata tccctcgggg aactggctc cagacactga agcatatttc tgaactgctc	1380
agagaagcag ttgaagacca gactcatggg tcctgcggag gtccctttga gagctggttc	1440
ctgttcattc acttcggagg atgggctgag atggtggcag agcaattact gatgtcggca	1500
gccgaacccc ccacggccct gctgtggctc ttggccttct actacggccc ccgtgatggg	1560

aggcagcaga gagcacagac tatggtccag gtgaaggccg tgctgggcca cctcctggca	1620
atgtccagaa gcagcagcct ctcagcccag gacctgcaga cggtagcagg acagggcaca	1680
gacacagacc tcagagctcc tgcacaacag ctgatcaggc accttctcct caacttcctg	1740
ctctgggctc ctggaggcca cacgatcgcc tgggatgtca tcaccctgat ggctcacact	1800
gctgagataa ctcacgagat cattggcttt cttgaccaga cttgttacag atggaatcgt	1860
cttggcattg aaagccctag atcagaaaaa ctggcccag agctccttaa agagctgcga	1920
actcaagtct agaaggcacg caggccgtgt gggtgcccgg cgtgagggat caggctcgcc	1980
agggccacag gacaggtgat gacctgtggc cacgcatttg tggagtaagt gccctcgctg	2040
ggctgtgaga atgagctgta cacatcttgg gacaatctgc tagtatctat ttacaaaat	2100
gcagagccag gtccctcagc ccagactcag tcagacatgt tcactaatga ctcaagtgag	2160
ccttcggtac tcctggtgcc cgcccggcca gaccgtcagc ttgataatta ctaaagcaaa	2220
ggcctgggtg ggagaacagg tttctagttt ttaccaagt caagctgcac atctattatt	2280
taaaaattca aagtcttaga accaagaatt tggatcatgaa ccattaaaga atttagagag	2340
aacttagctc ttttagact ctttttagga gtcagggatc tgggataaag ccacactgtc	2400
ttgctgtatg gagaaattct tcaaggggag tcagggtccc tcaggcttcc cttgtgtctc	2460
cctggacctg cctgacaggc cacaggagca gacagcacac ccaagcccgg gcctccggca	2520
cactctttcc actctgtatt tgctaaatga tgctaactgc taccaaaagg cccttgggac	2580
atcagaggag ccggcaggcg aaggtagagg atgtgttcca gaaacattag aaggcaggat	2640
taattcagtt agttagttct cttgttaaatt ggaaatggga attggaaatt cctgataaag	2700
aattggcctg gctgggtgca gtggctcaca cctgtgatcc cagcactttg ggaggccaag	2760
gcagggggat tacttcagcc caggagtcc agactgcctg gctaacatgg caatacccta	2820
tctctactaa aaatacaaaa attatcgggg tgcaatggca tgcatctgta atcccagcta	2880
ttcaagaggc tgaggcatga ggatctcttg aaccgggag gtgggagttg tagtgagccg	2940
agatcatgac actgcactcc agcctgggca acagagcgag accatctctt aaaaaaggc	3000
attgttagtg taatctcaag gttaacattt atttcatgtc agtacagggt gctttttcct	3060
ttcagggaca ttctggaatt gtattggttg tacattcttt tgtgtctatt ctgtttgtca	3120
agtgagtcaa gacttgcttt tgtccatttt gatttgtgtg tattagtctg agtcttggct	3180

ccgttttgag gtatgagcaa agttttgctg gattagaagt taaccttttag ggaaattcct	3240
tattttggtg tgtggcaatg ctaatagatc cactgaagat ctggaaaatt ccaggaactt	3300
ttcacctgag cttttcttct gagaaatgct gcagtcagaa ggggtgtgctg gtaaagtatt	3360
ttggtggcag ctgccatcat ggtcattgcc ttcatataac atgcttcgtg ctcatgggtca	3420
ttgccttcat ataacatgct tcgtgccatc atgatccttg ctttcatata acaaacatgc	3480
ttcgtcagag gtgttgggggt tgaaaaagga gctgcatgct tcactggagt tgagggcctc	3540
tctcctgttc tgactttaag ccagaacttg tggctggggc atggaagctg tgactcctct	3600
gtggacatgg tggcagcagg gaacccttag agagaggggc cactgggacc aggcctcctg	3660
ttgtggaggg actcctggga cagtcctcca ccctgtcctg tggctcctgtg tacagggttg	3720
gcctcttcct cctcccctgc caggcctctg cccatgcccc ttccttcctt ctctgggac	3780
tggatgaagct aggcatctgg aagacttctt cctagcctgg aagccctgac ctcgcccat	3840
ctgcagaatc tcccagttcc ttcacagctg ccgagtcctc tcacgggtgc ggtggaggcg	3900
gccttgccgg tggtgctttc tgggcagcca ggggttcctg ggtgggagga ctgtccctct	3960
ggggacgtgg cactgaagtg cctgctggct tcatgtggcc ctttgccctt tcccagcctg	4020
agagatgctc aaaggtgggg agctggggga gccaccctc ggccattccc tccacctcca	4080
agacaggtgg cggccgggca ggcactctta agcccacctc cccctcttgt tgccttcgat	4140
ttcggcaaag cctgggcagg tgccaccggg aaggaatggc atccgagatg ctgggcgggg	4200
acgcggcgtg gccgaggggg ccttgacggc gttggcgggg cctgggcaca ggggcagccg	4260
cagggaggca gggatggcaa ggcgtgaagc caccctggaa ggaactggac caaggtcttc	4320
agaggtgcga cagggctctgg aatctgacct tactctagca ggagtttttg tagactctcc	4380
ctgatagttt agtttttgat aaagcatgct ggtaaaacca ctaccctcag agagagccaa	4440
aaatacagaa gaggcggaga gcgcccctcc aaccaggctg ttattcccct ggactccgtg	4500
acatctgtgg aatttttttag ctctttaaaa tctgtaattt gttgtctatt ttttcattct	4560
aaataaaaact tcagtttgca cctaa	4585

<210> 37

<211> 5219

<212> DNA

<213> Homo sapiens

<220>

<223> >FANCD2|ENSG00000144554|ENST00000287647|5219

<400> 37

gtcagagcgg cgtcgggcct ggcgggaaag tcgaaaacta cgggcggcga cggcttctcg	60
gaagtaattt aagtgcacaa gacattggtc aaaatggttt ccaaaagaag actgtcaaaa	120
tctgaggata aagagagcct gacagaagat gcctccaaaa ccaggaagca accactttcc	180
aaaaagacaa agaaatctca tattgctaata gaagttgaag aaaatgacag catctttgta	240
aagcttctta agatatcagg aattattctt aaaacgggag agagtcagaa tcaactagct	300
gtggatcaaa tagctttcca aaagaagctc tttcagacct tgaggagaca cccttcctat	360
cccaaaataa tagaagaatt tgtagtggtc ctggagtctt acattgagga tgaagacagt	420
ttcaggaact gccttttgtc ttgtgagcgt ctgcaggatg aggaagccag tatgggtgca	480
tcttattcta agagtctcat caaactgctt ctggggattg acatactgca gcctgccatt	540
atcaaaacct tatttgagaa gttgccagaa ttttttttg aaaacaagaa cagtgatgaa	600
atcaacatac ctcgactcat tgtcagtcaa ctaaaatggc ttgacagagt tgtggatggc	660
aaggacctca ccaccaagat catgcagctg atcagtattg ctccagagaa cctgcagcat	720
gacatcatca ccagcctacc tgagatccta ggggattccc agcacgctga tgtggggaaa	780
gaactcagtg acctactgat agagaatact tcaactcactg tccaatcct ggatgtcctt	840
tcaagcctcc gacttgacct aaacttccta ttgaaggttc gccagttggt gatggataag	900
ttgtcgtcta ttagattgga ggatttacct gtgataataa agttcattct tcattccgta	960
acagccatgg atacacttga ggtaatttct gagcttcggg agaagttgga tctgcagcat	1020
tgtgttttgc catcacggtt acaggcttcc caagtaaagt tgaaaagtaa aggacgagca	1080
agttcctcag gaaatcaaga aagcagcggc cagagctgta ttattctcct ctttgatgta	1140
ataaagtcag ctattagata tgagaaaacc atttcagaag cctggattaa ggcaattgaa	1200
aacactgcct cagtatctga acacaagggtg tttgacctgg tgatgctttt catcatctat	1260
agcaccaata ctcagacaaa gaagtacatt gacagggtgc taagaaataa gattcgcata	1320
ggctgcattc aagaacagct gctccagagt acattctctg ttcattactt agttcttaag	1380
gatatgtgtt catccattct gtcgctggct cagagtttgc ttcactctct agaccagagt	1440

ataatttcat ttggcagtct cctatacaaa tatgcattta agttttttga cacgtactgc	1500
cagcaggaag tggttggtgc cttagtgacc catatctgca gtgggaatga agctgaagtt	1560
gatactgcct tagatgtcct tctagagttg gtagtgttaa acccatctgc tatgatgatg	1620
aatgctgtct ttgtaaaggg catttttagat tatctggata acatatcccc tcagcaaata	1680
cgaaaactct tctatgttct cagcacactg gcatttagca aacagaatga agccagcagc	1740
cacatccagg atgacatgca cttggtgata agaaagcagc tctctagcac cgtattcaag	1800
tacaagctca ttgggattat tgggtgctgtg accatggctg gcatcatggc ggcagacaga	1860
agtgaatcac ctagtttgac ccaagagaga gccaacctga gcgatgagca gtgcacacag	1920
gtgacctcct tgttgcagtt ggttcattcc tgcagtgagc agtctcctca ggcctctgca	1980
ctttactatg atgaatttgc caacctgac caacatgaaa agctggatcc aaaagccctg	2040
gaatgggttg ggcataccat ctgtaatgat ttccaggatg ccttcgtagt ggactcctgt	2100
gttgttccgg aaggtgactt tccatttcct gtgaaagcac tgtacggact ggaagaatac	2160
gacactcagg atgggattgc cataaacctc ctgccgctgc tgttttctca ggactttgca	2220
aaagatgggg gtccggtgac ctcacaggaa tcaggccaaa aattgggtgtc tccgctgtgc	2280
ctggctccgt atttccggtt actgagactt tgtgtggaga gacagcataa cggaacttg	2340
gaggagattg atggtctact agattgtcct atattcctaa ctgacctgga gcctggagag	2400
aagttggagt ccatgtctgc taaagagcgt tcattcatgt gttctctcat atttcttact	2460
ctcaactggg tccgagagat tgtaaatgcc ttctgccagg aaacatcacc tgagatgaag	2520
gggaagggtc tctctcggtt aaagcacatt gtagaattgc aaataatcct ggaaaagtac	2580
ttggcagtca cccagacta tgtccctcct cttggaaact ttgatgtgga aacttttagat	2640
ataacacctc atactgttac tgctatttca gcaaaaatca gaaagaaagg aaaaatagaa	2700
aggaaacaaa aaacagatgg cagcaagaca tcctcctctg acacactttc agaagagaaa	2760
aattcagaat gtgaccctac gccatctcat agaggccagc taaacaagga gttcacaggg	2820
aaggaagaaa agacatcatt gttactacat aattcccatg cttttttccg agagctggac	2880
attgaggtct tctctattct acattgtgga cttgtgacga agttcatctt agatactgaa	2940
atgcacactg aagctacaga agttgtgcaa cttgggcccc ctgagctgct tttcttgctg	3000
gaagatctct cccagaagct ggagagtatg ctgacacctc ctattgccag gagagtcccc	3060

tttctcaaga acaaaggaag ccggaatatt ggattctcac atctccaaca gagatctgcc	3120
caagaaattg ttcatgtgtg ttttcaactg ctgaccccaa tgtgtaacca cctggagaac	3180
attcacaact attttcagtg ttttagctgct gagaatcacg gtgtagttga tggaccagga	3240
gtgaaagttc aggagtacca cataatgtct tcctgctatc agaggctgct gcagatTTTT	3300
catgggcttt ttgcttggag tggatTTTTct caacctgaaa atcagaattt actgtattca	3360
gccctccatg tccttagtag ccgactgaaa cagggagaac acagccagcc tttggaggaa	3420
ctactcagcc agagcgtcca ttacttgcag aatttccatc aaagcattcc cagtttccag	3480
tgtgctcttt atctcatcag actttt gatg gttatTTTTg agaaatcaac agcttctgct	3540
cagaacaaag aaaaaattgc ttcccttgcc agacaattcc tctgtcgggt gtggccaagt	3600
ggggataaag agaagagcaa catctcta at gaccagctcc atgctctgct ctgtatctac	3660
ctggagcaca cagagagcat tctgaaggcc atagaggaga ttgctggtgt tggtgtccca	3720
gaactgatca actctcctaa agatgcatct tcctccacat tccctacact gaccaggcat	3780
acttttgttg ttttcttccg tgtgatgatg gctgaactag agaagacggt gaaaaaatt	3840
gagcctggca cagcagcaga ctcgcagcag attcatgaag agaaactcct ctactggaac	3900
atggctgttc gagacttcag tatcctcatc aacttgataa aggtatttga tagtcatcct	3960
gttctgcatg tatgtttgaa gtatgggcgt ctctttgtgg aagcatttct gaagcaatgt	4020
atgccgctcc tagacttcag ttttagaaaa caccgggaag atgttctgag cttactggaa	4080
accttccagt tggacacaag gctgcttcat cacctgtgtg ggcatccaa gattcaccag	4140
gacacgagac tcaccaaca tgtgcctctg ctcaaaaaga ccctggaact tttagtttgc	4200
agagtcaaag ctatgctcac tctcaacaat tgtagagagg ctttctggct gggcaatcta	4260
aaaaaccggg acttgcaggg tgaagagatt aagtcccaa attcccagga gagcacagca	4320
gatgagagtg aggatgacat gtcattcccag gcctccaaga gcaaagccac tgaggatatct	4380
ctacaaaacc caccagagtc tggcactgat ggttgcattt tgtaattgt tctaagttgg	4440
tggagcagaa ctttgcctac ttatgtttat tgtcaaatgc ttctatgcc atttccattc	4500
cctccataac agcttctgtg cttatataat ttttgggacc cagaagaaac aacgacacaa	4560
tcttagaatc actcctgagt atctcgagtt gtggcatttg ttatagagtt gacaattttc	4620
tgcatatatag cctctcattt tccatgaatt catatctgaa accatttttag aaggggagaag	4680

tcatcgaagt attttctgag tgttgagaag aatgagttaa accattttaa cacatttgaa	4740
acatacaaaa atagaaatgt gaaagcattt ggtgaaagcc aaagcacaga gtcagaagct	4800
gccaccttag agaactgaaa taaaaataga agttcttacg cttttttgtg gtacagatgc	4860
tttcgacaat ttaaagaaag ctaaataaaa atgtagacat ggctggcgca gtggctcatg	4920
cttgtaatcc tagcactttt tgaggccaag gtaggaggat tgcttgagtc cgggagctca	4980
aggcaaagct gcacaacata acaagaccct atctccacaa aaaaaatgaa aaataaacct	5040
gggtgcggtg gctcacacct gtaatcccag cactttggga ggccgatgtg ggcagatcac	5100
aaggtcagga gttcaagacc agcctggcca acatagttaa accccatctc tactgaaaat	5160
acaaaaatta gctgggtgtg gtggcacgtg cctgttatct cagctacttg ggaggctga	5219

<210> 38

<211> 2554

<212> DNA

<213> Homo sapiens

<220>

<223> >FANCE|ENSG00000112039|ENST00000229769|2554

<400> 38

aacggctgcg gcttcgggcg gccgggtttc tccggtctcc caacgccgag gagagcttgt	60
aacaggcgct ggagctggcc cgccaccgcc gcgtcaggga cggcgctgga gtcctccgtt	120
cccctcagcc tctgagctga ggccccacac cagagtaggg ggcggcgcgg caccctgccc	180
ccggcatggc gacaccggac gcggggctcc ctggggctga gggcgtggag ccggcgccct	240
gggcgcagct ggaggcccc gccgcctcc tgctgcaggc gctgcaggcg gggcctgagg	300
gggcgcggcg cggcctgggg gtgctccggg cgctgggcag ccgcggctgg gagcccttcg	360
actggggtcg cttgctcgag gccctgtgcc gggaggagcc ggtcgtgcag gggcctgacg	420
gccgtctgga gctgaaacca ctgttgctgc gattgccccg gatatgccag aggaacctga	480
tgtccctgct gatggccgtt cggccatcgc tgccgaaaag tgggctcctc tctgtgctgc	540
agattgcccga gcaggaccta gccctgacc cagatgcctg gctccgtgcc ctgggggaat	600
tgctgcgaag ggatttgggg gtggggacct ccatggaggg agcttctcca ctgtctgaaa	660
gatgccagag acagctccaa agtctatgta gggggctggg cctggggggc aggaggttga	720
aatcccccca ggctccagac cctgaagaag aggagaacag ggactcccag cagcctggga	780

aacgcagaaa ggactcagag gaagaggctg ccagtcctga ggggaagagg gtccccaaaa	840
gattacggtg ttgggaagag gaagaagatc atgagaagga gagacccgaa cataagtcac	900
tggaatccct ggcatgatgga ggaagtgcac ctctatttaa ggaccagcct gtcattggcag	960
ttaagactgg cgaggacggt tcgaatctgg atgatgctaa aggtctggct gagagtttgg	1020
agttgcccac agctatccag gaccagcttc ccaggctgca gcagctgctg aagaccttgg	1080
aggaggggtt agagggttgg gaggatgccc cccagttga gctacagctt cttcacgaat	1140
gtagtcccag ccagatggac ttgctgtgtg cccagctgca gctccctcag ctctcagacc	1200
tcggtctcct gcggctctgc acctggctgc tggccctttc acctgatctc agcctcagca	1260
atgctactgt gctgaccaga agcctctttc ttggacggat cctctccttg acttcctcag	1320
cctcccgctt gcttacaact gccctgacct ccttctgtgc caaatataca taccctgtct	1380
gcagcgccct ccttgaccct gtgctccagg cccagggcac aggtcctgct caaacagagt	1440
tactgtgttg ccttgtgaag atggagtccc tggagccaga tgcacagggt ctaatgctgg	1500
gacagatctt ggagctgccc tggaaggagg aaactttctt ggtgttgagc tcactcctag	1560
agcggcagggt ggagatgacc cctgagaagt tcagtgtctt aatggagaag ctctgtaaaa	1620
aggggctggc agccaccacc tccatggcct atgccaagct catgctgaca gtgatgacca	1680
agtatcaggc taacatcact gagaccaga ggctgggcct ggctatggcc ctagaaccta	1740
acaccacctt cctgaggaag tccctgaagg ccgccttgaa acatttgggc ccctgaccat	1800
ccaccaaggg accaccctct tgggtgctca tcaccagctt cctgaagggc atttctttct	1860
tcaccacctt gtcttgagcc ctagcctgag gataaaggct gagcctggcc atcccagatt	1920
gcaactgcct ccctgttaca ggctgcatag ggaatattgg cttcccagcc agcagggagg	1980
ctcctgggct aagggtgctc agctatattt tctttttcat ttcttttggt tttgagagaa	2040
ggtattgctc tgtcatccag gctggagtgc agtgatgcga ttgatcatgg ctactgttaa	2100
cctctgcctc ccaggctcga tcctcccatg tcagcctccc aagtagctgg gactacatgc	2160
acatgacact atgcccagct aattttttat tttgtagata atatgctggg cagggtggt	2220
cttgaactcc tgagctcaag tgatcttcct gccttggcac ccaaagtgtt gagatttaca	2280
ggtgtgagcc accacaccag gccagagcta tatttccaaa ggctgctggc cccaggcaca	2340
ctcctcatca attctcaggc tgcagggaca aactttccga tagggctcag taggatcaag	2400

ccgacccaga gtgggcatgg gatgctccag gaggtctggg gtaccagtgg gctcagatga	2460
agctcttgtg ctcatgtata ttatgacctt ctgtgttttt gtttctgact tgaataattt	2520
atcaatggtg ttgaataaaa gcaagttcaa taat	2554

<210> 39
 <211> 3309
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FANCF|ENSG00000183161|ENST00000327470|3309

<400> 39	
tgaaagcggg agtagggcct tcgcgcacct catggaatcc cttctgcagc acctggatcg	60
cttttccgag cttctggcgg tctcaagcac tacctacgtc agcacctggg accccgccac	120
cgtgcgccgg gccttgcagt gggcgcgcta cctgcgccac atccatcggc gctttggtcg	180
gcatggcccc attcgcacgg ctctggagcg gcggctgcac aaccagtgga ggcaagaggg	240
cggctttggg cgggggtccag ttccgggatt agcgaacttc caggccctcg gtcactgtga	300
cgtcctgctc tctctgcgcc tgctggagaa ccggggcctc ggggatgcag ctcgttacca	360
cctggtgcag caactctttc ccggcccggg cgtccgggac gccgatgagg agacactcca	420
agagagcctg gcccgccttg cccgccggcg gtctgcggtg cacatgctgc gcttcaatgg	480
ctatagagag aacccaaatc tccaggagga ctctctgatg aagaccagg cggagctgct	540
gctggagcgt ctgcaggagg tggggaaggc cgaagcggag cgtcccgcc ggtttctcag	600
cagcctgtgg gagcgcttgc ctcaagaaca cttcctgaag gtgatagcgg tggcgctgtt	660
gcagccgcct ttgtctcgtc ggccccaaga agagttggaa cccggcatcc acaaatcacc	720
tggagagggg agccaagtgc tagtcactg gcttctgggg aattcggaag tctttgctgc	780
cttttgtcgc gccctcccag ccgggctttt gacttttagtg actagccgcc acccagcgt	840
gtctcctgtc tatctgggtc tgctaacaga ctgggggtcaa cgtttgcact atgaccttca	900
gaaaggcatt tgggttggaa ctgagtccca agatgtgccc tgggaggagt tgcacaatag	960
gtttcaaagc ctctgtcagg cccctccacc tctgaaagat aaagttctaa ctgccctgga	1020
gacctgtaaa gcgcaggatg gagattttga agtacctggt cttagcatct ggacagacct	1080

cttattagct cttcgtagt gtgcatttag gaaaagacaa gttttgggtc tcagcgcagg	1140
cctcagttct gtataggcaa tgctgtgtta ttacttgaat atagaatata tagtttacaa	1200
aatgaaaatt acaatgttct caccaaata atgccttcgt gtgtccaaag tataattatt	1260
ttagatgcta attttgaata gtttattaaa cagttataaa tatgcaaagt agctggcatg	1320
tagtgtcacg gattttctgg atagaggaag tgattggaag tactccactt aaagccatgg	1380
aattagcaat agtttgcttt ttaatagaag gcccatgtgt aagaatgttg aaaatatgtg	1440
taccgtttaa agaaaaagca gctttaaagt gacaaacaaa ataccctttt tcttttagta	1500
tggtttatth ttctaggttt tctgtccctc cctcagtagt gaagagtttt ctttattcct	1560
ggcagtgta ggaatattgg tttgaaaagc tgttggccta tctggagttt ggccttgta	1620
acctagtatt ctaaccagtt aaccagcctt agtatgcatt aaaattgtat tgttcagaaa	1680
gtttgtttct cattttctgc aaattcttac tttgaaaatg aatcaccaca tagtatgtcc	1740
ctttaagca ttgacgcaca gacaaatgtt taaagcacag taaatacgaa tatatgcctt	1800
tggatattaa attaattgctt gatgataaaa gaatcaaact tttttttttt tgagatggag	1860
tctcgctctg tcacccagac tggagtgcag tgggtgtgatc actgctcagt gcaacctctg	1920
cctcccagga tcaagcaatt ctgactcagc ctccaagta gctgggatta caggcgcagg	1980
ccaccatgcc cggctaattt tttgtatttt tagtagagac ggggtttcac catgctggcc	2040
aggctggtct caaacacctg accttgatgat ccgtccgcct tggcctccca aagtgctggg	2100
attacaggcg tgagccaccg cgcctggcca aaacaacatt ttaagtagaa gatccaggtt	2160
ttagtgcagc ttctgccgtt aactagggtta ataaatcaca accttggggc cacagttgcc	2220
ttatatgtaa atgaagtgtt tagaataaaa tagttaaatt tccttatttt tcccttggtg	2280
gctgccctgt ggaaacagtt tagaatattt gttttgtgtg taggaacctt gttgtgttag	2340
ttacctggg tgttccacag ctgatagtga ttgccttgaa taaattcaag ggcaatttat	2400
tcatttttac tagggagata gacctttaca gcaatcaaga ttttttgtc catatccagg	2460
ttagctggta agaggatttt tttggagaaa aaaatgatat ttagaaaagt aatttcta	2520
tccggaatgg aataaaaaa atatgagtag tgtaatcttg tagaaaaaga gttgtataat	2580
ctttagaat ttctattct gtggtacaac ccaggggtaa actattattc cagtagtcag	2640
tacacttttc tagataaatc ttgagtga accagcaatt tctttttcct tgtggtctga	2700

ttcctttttc taatccatga aggccatctt gtagattaca tttatcatta atgcaagaat	2760
aaagacaatt cctcctgtca gttgcgtgaa ttttttttaa gaaacaaccc agtgaagagt	2820
tctaccatag caaggcctaa tgttagcttt agcttttagaa aataacagtt tgtgaactta	2880
cttccttata tttgcagctg tatctcacac tatgatttac aataaaattg taaagattga	2940
caatagactt aagaaataac attttaaaat ctattttata cttaccattt attattctgt	3000
tatttttagtc tccatatgtt cattacatac ataatcttat ttaatcttca caccaaaact	3060
gtattcttat gaatatacgc taaaagatta agtaaaatgc ccaagggtat aaacaaaagc	3120
aacatgaaag tggaagccgt atctgtcatt tattttatatt ccagaagcct agcacagtgt	3180
ccagcatatg gtagatactt gtagtgtttg aataaatgaa accagcatta gagctttata	3240
tactttctct taaggacttg aaaagattag gaatctacgc atacactgag agagaaaaaa	3300
gtgagagga	3309

<210> 40

<211> 2631

<212> DNA

<213> Homo sapiens

<220>

<223> >FANCG|ENSG00000221829|ENST00000378643|2631

<400> 40

aggaggacct ggggggtgtgg cagcgaggaa gggccgagcc acggactgtg gggccgaaac	60
tcgctcccg cccacccttc tcgaggctgt ggcctccgcg agagccgagc gggccgcacc	120
gccggccgtg cgactgcccc agtcagacac gaccccggt tctagcccg ctaagcctgt	180
ttggggttgc tgactcgttt cctccccgag tttcccgagg gaactaactc ttcaagagga	240
ccaaccgcag cccagagctt cgcagacccg gccaaaccaga ggcgagggtt agagcccggc	300
gggcccgcggg gagagagcgt cccatctgtc ctggaaagcc tgggcgggtg gattgggacc	360
ccgagagaag caggggagct cggcgggggtg cagaagtgcc caggcccctc cccgctgggg	420
ttgggagctt gggcaggcca gcttcaccct tcctaagtcc gcttctggtc tccgggcccc	480
gcctcggcca ccatgtcccg ccagaccacc tctgtgggct ccagctgcct ggacctgtgg	540
agggaaaaga atgaccggct cgttcgacag gccaaaggtg ctcagaactc cggtctgact	600
ctgaggcgac agcagttggc tcaggatgca ctggaagggc tcagagggt cctccatagt	660

ctgcaagggc tccctgcagc tgttcctgtt cttcccttgg agctgactgt cacctgcaac	720
ttcattatcc tgagggcaag cttggcccag ggtttcacag aggatcaggc ccaggatatac	780
cagcggagcc tagagagagt gctggagaca caggagcagc agggggcccag gttggaacag	840
gggctcaggg agctgtggga ctctgtcctt cgtgcttcct gccttctgcc ggagctgctg	900
tctgccctgc accgcctggt tggcctgcag gctgccctct ggttgagtgc tgaccgtctt	960
ggggacctgg ctttgttact agagaccctg aatggcagcc agagtggagc ctctaaggat	1020
ctgctgttac ttctgaaaac ttggagtccc ccagctgagg aattagatgc tccattgacc	1080
ctgcaggatg cccagggatt gaaggatgtc ctctgacag catttgctta ccgccaaggt	1140
ctccaggagc tgatcacagg gaaccagac aaggcactaa gcagccttca tgaagcggcc	1200
tcaggcctgt gtccacggcc tgtgttggtc caggtgtaca cagcactggg gtcctgtcac	1260
cgtaagatgg gaaatccaca gagagcactg ttgtacttgg ttgcagccct gaaagaggga	1320
tcagcctggg gtcctccact tctggaggcc tctaggctct atcagcaact gggggacaca	1380
acagcagagc tggagagtct ggagctgcta gttgaggcct tgaatgtccc atgcagttcc	1440
aaagccccgc agtttctcat tgaggtagaa ttactactgc caccacctga cctagcctca	1500
ccccttcatt gtggcactca gagccagacc aagcacatac tagcaagcag gtgcctacag	1560
acggggaggg caggagacgc tgcagagcat tacttggacc tgctggccct gttgctggat	1620
agctcggagc caaggttctc cccaccccc tcccctccag ggccctgtat gcctgagggtg	1680
tttttgaggg cagcggtagc actgatccag gcaggcagag cccaagatgc cttgactcta	1740
tgtgaggagt tgctcagccg cacatcatct ctgctacca agatgtccc gctgtgggaa	1800
gatgccagaa aaggaaccaa ggaactgcca tactgcccac tctgggtctc tgccaccac	1860
ctgcttcagg gccaggcctg ggttcaactg ggtgccc aaa aagtggcaat tagtgaattt	1920
agcagggtgcc tcgagctgct cttccgggcc acacctgagg aaaaagaaca aggggcagct	1980
ttcaactgtg agcagggatg taagtcagat gcggcactgc agcagcttcg ggcagccgcc	2040
ctaattagtc gtggactgga atgggtagcc agcggccagg ataccaaagc cttacaggac	2100
ttcctcctca gtgtgcagat gtgcccagggt aatcgagaca cttactttca cctgcttcag	2160
actctgaaga ggctagatcg gagggatgag gccactgcac tctggtggag gctggaggcc	2220
caaactaagg ggtcacatga agatgctctg tggctctctc ccctgtacct agaaagctat	2280

ttgagctgga tccgtccctc tgatcgtgac gccttccttg aagaatttcg gacatctctg	2340
ccaaagtctt gtgacctgta gctgccacgt tttgaagagc ttgagctggg tccccagtg	2400
gctgtctctc tgtggggagg gctttctgct tcaccatcat taggaatgtg accattccta	2460
tataattcct ggactggtga gattggtggt aggccctgtga aatttgcct agttactacc	2520
attctcgttt tggaggaaac aatctctgcc accaccaagt cattgacttt gctcgaggca	2580
ccttttttcc tgtttctcct tttctgttgt cgagtaaaat ttcataattta t	2631

<210> 41
 <211> 4743
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FANCI|ENSG00000140525|ENST00000310775|4743

<400> 41	
gttgttacgg gtaacggaag tgtggcggcg ttgggttgag cgggcttttt ggaagtttgt	60
ggcggagttc tgtgatatga gcaacaatgg accagaagat tttatctcta gcagcagaaa	120
aaacagcaga caaactgcaa gaatttcttc aaaccctgag agaaggtgat ttgactaatc	180
tccttcagaa tcaagcagt aaaggaaaag ttgctggagc actcctgaga gccatcttca	240
aaggttcccc ctgctctgag gaagctggaa cacttaggag acgtaagata tacacttggt	300
gtatccagtt ggtggaatcg ggggatttgc agaaagaaat agcgtctgag atcataggat	360
tactgatgct ggaggctcac cattttccag gaccattatt ggttgaatta gccaatgagt	420
ttattagtgc tgtcagagaa ggcagcctag tgaatggaaa atctttggag ttactaccta	480
tcattctcac tgccctggct acgaaaaagg aaaatctggc ttatggaaaa ggtgtactga	540
gtggggaaga atgtaagaaa cagttgatta acaccctgtg ttctggcagg tgggatcagc	600
aatatgtaat ccaactcacc tccatgttca aggatgtccc tctgactgca gaagaggtgg	660
aatttgtggt ggaaaaagca ttgagcatgt tctccaagat gaatcttcaa gaaataccac	720
ctttgggtcta tcagcttctg gttctctcct ccaagggaag cagaaagagt gttttggaag	780
gaatcatagc cttcttcagt gcactagata agcagcacia tgaggaacag agtggtgacg	840
agctattgga tgttgtcact gtgccatcag gtgaacttcg tcatgtggaa ggcaccatta	900

ttctacacat tgtgtttgcc atcaaattgg actatgaact aggcagagaa ctcgtgaaac	960
acttaaaggt aggacagcaa ggagattcca ataataactt aagtccttc agcattgctc	1020
ttcttctgtc tgtaacaaga atacaaagat ttcaggacca ggtgcttgat cttttaaaga	1080
cttcggttgt aaagagcttt aaggatcttc aactcctcca aggctcaaaa tttcttcaga	1140
atctagttcc tcatagatct tatgtttcaa ccatgatctt ggaagtagtg aagaatagcg	1200
ttcatagctg ggacatggt actcagggcc tcgtagaact tggtttcatt ttgatggatt	1260
catatgggcc aaagaagggt cttgatggaa aaactattga aaccagccca agtctttcta	1320
gaatgcaaaa ccagcatgca tgtaagctcg gagctaatat cctgttgga acttttaaga	1380
tccatgagat gatcagacaa gaaattttgg agcaggtcct caacagggtt gttaccagag	1440
catcttctcc catcagtcatt ttcttagacc tgctttcaaa tatcgtcatg tatgcaccct	1500
tagttcttca aagttgttct tctaaagtca cagaagcttt tgactatttg tcctttctgc	1560
cccttcagac tgtacaaagg ctgcttaagg cagtgcagcc ccttctcaaa gtcagcatgt	1620
caatgagaga ctgcttgata cttgtccttc ggaaagctat gtttgccaac cagcttgatg	1680
cccgaaaatc tgcagttgct gggtttttgc tgctcctgaa gaactttaaa gttttaggca	1740
gcctgtcatc ctctcagtc agtcagtcct tcagtgctcag tcaggttcat gtggatgttc	1800
acagccatta caattctgtc gccaatgaaa ctttttgcct tgagatcatg gatagtttga	1860
ggagatgctt aagccagcaa gctgatgttc gactcatgct ttatgagggg ttttatgatg	1920
ttcttcgaag gaactctcag ctggctaatt cagtcatgca aactctgctc tcacagttaa	1980
aacagttcta tgagccaaaa cctgatctgc tgcctcctct gaaattagaa gcttgtattc	2040
tgaccaagg agataagatc tctctacaag aaccactgga ttatctgctg tgttgtattc	2100
agcattgttt ggcctggtat aagaatacag tcataccctt acagcaggga gaggaggaag	2160
aggaggagga agaggcattc tacgaagacc tagatgatat attggagtcc attactaata	2220
gaatgattaa gagtgagctg gaagactttg aactggataa atcagcagat ttttctcaga	2280
gcaccagtat tggcataaaa aataatatct gtgcttttct tgtgatggga gtttgtgagg	2340
ttttaataga atacaatttc tccataagta gtttcagtaa gaatagggtt gaggacattc	2400
tgagcttatt tatgtgttac aaaaaactct ctgacattct taatgaaaaa gcgggtaaag	2460
ccaaaactaa aatggccaac aagacaagt atagtctttt gtccatgaaa tttgtgtcca	2520

gtcttctcac tgctcttttc agggatagta tccaaagcca ccaagaaagc ctttctgttc	2580
tcagggtccag caatgagttt atgcgctatg cagtgaatgt agctctgcag aagggtacagc	2640
agctaaagga aacagggcat gtgagtggcc ctgatggcca aaaccagaa aagatctttc	2700
agaacctctg tgacataact cgagtcttgc tatggagata cacttcaatt cctacttcag	2760
tggaagagtc gggaaagaaa gagaaaggaa agagcatctc actgctgtgc ttggagggtt	2820
tacagaaaat attcagtgtc gtgcaacagt tctatcagcc caagattcag cagtttctca	2880
gagctctgga tgtcacagat aaggaaggag aagagagaga agatgcagat gtcagtgtca	2940
ctcagagaac agcattccag atccggcaat ttcagaggtc cttgttgaat ttacttagca	3000
gtcaagagga agattttaat agcaaagaag ccctcctgct agtcacggtt cttaccagtt	3060
tgtccaagtt actggagccc tcctctctc agtttgtgca gatgttatcc tggacatcaa	3120
agatttgcaa ggaaaacagc cgggaggatg cttgttttg caagagcttg atgaacttgc	3180
tcttcagcct gcatgtttcg tataagagtc ctgtcattct gctgcgtgac ttgtcccagg	3240
atatccacgg gcatctggga gatatagacc aggatgtaga ggtggagaaa acaaaccact	3300
ttgcaatagt gaatttgaga acggctgccc ccactgtctg ttacttggtt ctgagtcagg	3360
ccgagaaggt tctagaagaa gtggactggc taatcaccaa gcttaaggga caagtgagcc	3420
aagaaacctt atcagaagag gcctcttctc aggcaacctt accaaatcag cctgttgaga	3480
aagctatcat catgcaactg ggaactctgc ttacattttt ccacgagctg gtgcagacag	3540
ctctgccatc aggagctgt gtggacacct tgttaaagga cttgtgcaaa atgtacacca	3600
cacttacagc cttgtcaga tattatctcc aggtgtgtca gagctccgga ggaattccaa	3660
aaaatatgga aaagctggtg aagctgtctg gttctcatct gacccccctg tgttattctt	3720
tcattttcta cgtacagaat aagagtaaga gcctgaacta tacgggagag aaaaaggaga	3780
aacctgctgc cgttgccaca gccatggcca gagttcttcg ggaaaccaag ccaatcccta	3840
acctcatctt tgccatagaa cagtatgaaa aatttctcat ccacctttct aagaagtcca	3900
aggtgaacct gatgcagcac atgaagctca gcacctcacg agacttcaag atcaaaggaa	3960
acatcctaga catggttctt cgagaggatg gtgaagatga aaatgaagag ggcactgcat	4020
cagagcatgg gggacagAAC aaagaaccag ccaagaagaa aaggaaaaaa taaatgaaat	4080
gcctgagtta atgtgaactt tggggcttct gcttcatttt tacccaacaa gcaacaatgc	4140

cccttgtcct gtagtcaca ccgatgttgg catcttggtt ctgaaccac tgaattcaac	4200
tgcaccttca gttagaagga atcttcttgg caggtcctgc tactgaaaaa tggctggcct	4260
taggcaagcc cttttgcaaa aagcacagct gaaagcctga gtttgggagc ctgcaccacc	4320
ccgatgaagc tccacgggag caaatacaga gcctccaggc agtgctatgg tccaggctgg	4380
cttcgttttt ccaaggagcc tttggtgagt tcaattatct ggtaaataatc cagcgcttca	4440
cctgaaagat agtgcaaatt ggtaggatg ccacctcaag aactgtaact gagagctcag	4500
aagtgagcaa aggagcttaa tgctaaggctc aaaaggagag tgaaagggtg agaacaattg	4560
ccacgaacgg taatgttaca tgtaggagg gtctgttttc tttttatata agtgtgtctt	4620
agatatattt taaatagaaa ataagctttc tgatttactt gtttgggtatt taaagcacag	4680
tttggttttc tgtcacctat agagtgaag aatgcactct atagaataaa ttatctttaa	4740
aca	4743

<210> 42
 <211> 1698
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FANCL|ENSG00000115392|ENST00000402135|1698

<400> 42	
gtctagagct tttctgtgtt tctccggact tcgagccatg gcggtgacgg aagcgagcct	60
gttgcgccag tgccccctgc ttctgccccca gaaccggtcg aaaaccgtgt atgagggatt	120
catctcggct cagggaagag acttccacct taggatagtg ttgcctgaag atttacaact	180
gaagaatgca agattattat gtagttggca gctgagaaca atacttagtg gataccatcg	240
aatagtacaa cagagaatgc agcactctcc tgatctaatt agctttatga tggagttgaa	300
gatgcttttg gaagttgcct taaagaatag acaagagctg tatgcactac ctcctcctcc	360
ccagttctac tcaagcctta ttgaagagat aggaactctt gggtgggata aacttgtgta	420
tgcggatacc tgcttcagta ccatcaagtt aaaagcagaa gatgcttctg gtagagagca	480
tttaatcact ctcaagttga aggcaaagta tcctgcagaa tcaccagatt attttgtgga	540
ttttcctgtt ccattttgtg cctcctggac acctcaggta aattctcctc agagctcctt	600
aataagcatt tatagtcagt ttttggcagc aatagaatca ctaaaggcat tctgggatgt	660

tatggatgaa atcgatgaga agacctgggt acttgagcca gaaaaacctc cacggagtgc	720
aacagcacgc agaattgcat taggtaataa tgtttccata aatatagagg tagaccccag	780
gcatcctact atgcttcctg agtgcttctt tcttgagct gaccatgtgg taaaaccct	840
gggaattaag ctgagcagga acatacattt gtgggatcca gaaaatagtg tgttacaaaa	900
tttgaaagat gttttagaaa ttgattttcc agctcgtgct atcctggaaa aatctgattt	960
tactatggat tgtggaattt gttatgctta tcaacttgac ggtaccattc ctgatcaagt	1020
gtgtgataat tctcagtgtg gacaaccttt ccatcaaata tgcttatatg agtggctgag	1080
aggactacta actagtagac agagttttta catcatattt ggtgaatgtc catattgtag	1140
taagccaatt accttaaaaa tgtctggaag gaaacactga aataagaata caacatttcg	1200
gtgaagagct ggaaacttaa aaaattatca aaaggaattt tggtatcatc ttcagagaaa	1260
aaataaagca agaaatacta acatcaaaag gacaggtatg atgatgcat aataataaac	1320
atctgcgttt gtctcttcac taagagtaaa ctgggaaatt gtaggcaaaa gtccagttga	1380
actttctaag tctgtgatcc ccgtgctgac tgtggaagtg tatttatacc aagatggaga	1440
tcttgacttc ttgaatatat ctggactgggt aaaatcttga tgaggctcat aaaatgagtt	1500
tggaattgt gtatagctga ttttttgtgg gaaactgttt acttcattca aaggttcttg	1560
agactcttga tatttctgtc ttctccttgt gctttcctat ggaaaaaata catatatagt	1620
ttagtttgtt agacgtgagt tatccaagta tttattttgt gtagtgtgta agaatgctaa	1680
ataaaatgtt atacaaga	1698

<210> 43
 <211> 7111
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FANCM|ENSG00000187790|ENST00000267430|7111

<400> 43	
ggaaaccgat ggggatcgga accgtagcgg ttgagctgct gctgctacgg atatctgaca	60
gaagccttcg gtggttgtcg gcctaataag cggacggcaa agaacgcttt ttcagacgtg	120
gggctcaagt atctcccgat catctgggac tccgggttgc agctccggaa ctgagcgacc	180

tcagagccct ggcagctcca aggcgccttt gccagcagca gcggaggctc agctggagtc	240
ggacgatgat gtgttgcttg tcgcggcgta cgaggctgag cggcagttgt gtctagagaa	300
tggcgggttc tgcacctccg cgggcgcctt gtggatttac cctaccaatt gcccagtgcg	360
ggactaccag ctgcacattt cccgggctgc tctgttttgc aatacgctgg tgtgtctgcc	420
taccggactg ggaaagacct ttattgccgc cgtgggtcatg tacaatttct accgctggtt	480
cccttcagga aagggtgtct tcatggcccc aacgaaacct ttggtgacac agcagatcga	540
ggcttgctac cagggtgatgg gtatcccgca atcccacatg gccgaaatga cagggtctac	600
acaagcttcc accaggaagg aaatatggtg cagtaagaga gtgctttttc ttacacctca	660
ggtcatggta aatgaccttt ctagaggagc ttgtcccgtc gctgaaataa agtgtttagt	720
tattgatgaa gtcataaag ctctcggaaa ctatgcttat tgccaggttg taagagaact	780
agtcaaatat acaaatactt ttagaatctt ggctctaagt gccacaccag gtagtgatat	840
aaaggctgtg caacaagtta ttactaacct gctaattggg cagatagagc ttcgttctga	900
agatttctca gatattttga catattctca tgaaagaaaa gttgaaaagc ttattgttcc	960
gcttggtgaa gaacttgcag ccatccaaaa gacctatata cagatttttg aatcatttgc	1020
tcgttctttg attcagagga atgttttgat gagaagggat atcccaaata taacaaaata	1080
tcagataatt ctggcaagag atcagtttag gaaaaaccca tctccgaata ttgtgggaat	1140
acaacaaggc ataatcgagg gagagtttgc tatttgtatt agtttatata atggttatga	1200
attattgcag caaatgggaa tgagatcatt atatttcttc ctttgtggaa ttatggatgg	1260
aactaaaggg atgacacggt caaaaaatga acttggccga aatgaagact tcatgaaact	1320
ctataatcat ctagagtgta tgtttgcacg tacacgtagt acttcagcaa atggtatttc	1380
tgctatccaa caaggagata aaaataaaaa atttgtttat agtcatccaa agttaaagaa	1440
attagaagaa gttgtaattg aacacttcaa gtcatggaat gctgaaaaca ctactgaaaa	1500
gaaacgtgat gagacccgag ttatgatctt ctcttcattt cgagatagtg ttcaagaaat	1560
tgcagaaatg ctttcacagc atcagccaat tattagagta atgacttttg tcggccatgc	1620
ctcagggaaa agcacgaagg gttttaccca gaaggagcaa ctggaggtag tgaaacagtt	1680
tcgtgacggt ggttacaaca cgctggtttc tacctgtgtg ggtgaagaag gtttggatat	1740
aggagaagtt gatcttataa tatgttttga ttcccagaag agcccaattc gtcttgtaca	1800

acgaatgggt agaactggcc gtaaactgca aggcaggata gttattatcc tttctgaagg	1860
acgagaggaa cgtatttata atcagagtca gtccaacaaa agaagtatat ataaagctat	1920
ttcaagtaac aggcagggtcc ttcattttta ccaaagaagt ccacgaatgg ttcctgatgg	1980
aatcaacca aaattacaca aaatgttcat cacacatggt gtctatgaac cagagaagcc	2040
ttctcggaac ttgcagcgaa agtcatctat cttttcctat agggatggaa tgaggcaaag	2100
tagcctaaag aaagattggt tcttatcaga agaagaattt aaattatgga acagacttta	2160
tagattaagg gacagtgatg aaattaaaga gataacattg cctcaagttc agttttcttc	2220
tttacaaaat gaggaaaaca aaccagctca agaatcaacc actggaattc atcaactctc	2280
tctctctgaa tggagactgt ggcaagatca tcctttgcct acacatcaag ttgatcactc	2340
agatcgatgc cgccatttta taggccttat gcaaatgata gagggaatga gacacgaaga	2400
gggagaatgc agctatgaat tggaagttga atcttattta caaatggaag atgttacctc	2460
aacatttatt gctcccagga atgaatctaa taatcttgcc agtgacacct ttatcactca	2520
caagaaatcg tcatttataa agaacataaa tcaaggcagt tcatcctcag tgatagaatc	2580
tgatgaagaa tgtgctgaaa ttgttaaaca aactcatatc aaacctacta aaattgtttc	2640
tttaaagaaa aaagtgtcta aagaaataaa aaaagatcag cttaaaaaag aaaataatca	2700
cggtattata gattctgtag ataatgacag aaattccact gttgaaaata tttttcaaga	2760
agacctacca aatgataaaa ggacatcaga tacagatgaa attgctgcca catgtactat	2820
taatgaaaat gttattaaag aaccgtgtgt gttattaaca gagtgtcagt ttacaaataa	2880
atccactagt tcacttgctg gaaatgtttt agattctggt tataacagtt tcaatgatga	2940
aaaatctgtt tcatctaact tatttcttcc attcgaagaa gagctttata ttgttagaac	3000
agatgaccaa ttttataatt gtcactcatt gacaaaagag gtactagcta atgtagagag	3060
atttttatct tattctcctc cgcctctcag tggactctca gacttggaat atgaaattgc	3120
taagggtact gcacttgaga atttgctttt cttaccctgt gcagagcatt tacgaagtga	3180
taaatgcacc tgtttgctgt cacattcagc tgtgaattct caacagaatt tagaattgaa	3240
ttcacttaaa tgtataaatt atccatctga aaaaagttgc ctttatgata tacctaataa	3300
taatatttct gatgagccaa gtctctgtga ctgtgatgta cataaacata atcaaaatga	3360
aaatttagta cctaacaatc gtgttcaa atacagaagc cctgcacaga atttagttgg	3420

agagaacaat catgatgttg ataacagtga cctcccagta ttgtccactg atcaagatga	3480
aagtttgctg ttatttgaag atgttaatac agagttcgac gatgtgagtc tttcacccctt	3540
gaacagtaaa agcgaatctt tacctgtgtc agacaaaact gctattagtg aaacgcctct	3600
ggctctctcag ttcttaattt ctgatgaact tttgttggac aataattctg aactccaaga	3660
tcaaatcacc cgtgatgcta atagttttaa atctcgtgat cagagaggtg tacaggaaga	3720
aaaagtgaag aatcatgagg atatTTTTga ttgctctagg gatttatttt ctgttacctt	3780
tgatttagga ttctgtagtc cagattctga tgatgaaata ttggaacata catcagatag	3840
caatagacct ctagatgata tatatggaag gtatttggaa attaaggaga taagtgatgc	3900
aaattatgtt tcgaatcaag cactaatacc aagagatcat agtaaaaatt ttactagtgg	3960
aactgttatt atcccatcaa atgaagatat gcagaatcca aattatgtac atttgccact	4020
gagtgacgca aaaaatgaag aattgttatc tcctggttat tctcagtttt ctttaccagt	4080
gcaaaaaaaaa gttatgagta caccactctc taaatcaaac acattgaact cattttctaa	4140
gataagaaag gaaatactta agacaccaga ttctagtaag gaaaaagtaa acctacaaag	4200
attcaaagaa gcattgaatt caacttttga ttattcagaa ttttctctag aaaagtctaa	4260
aagcagtggc ccaatgtatc tgcataaatc ctgtcattct gttgaagatg gacaattatt	4320
aacaagtaac gaaagtgaag atgacgagat tttccgaaga aaagttaaaa gagcaaaagg	4380
aaatgtttta aactctcctg aggatcagaa aaatagtga gttgattctc cacttcatgc	4440
tgtcaaaaag cgcagatttc ctataaacag atcagaatta tcatctagtg atgagagtga	4500
gaattttccc aaaccatgtt cacaattaga agacttcaag gtttgtaacg ggaatgccag	4560
aagaggcatc aaagtcccaa agagacagag tcacttaaag catgtagcta ggaagttttt	4620
agatgatgaa gcagaacttt ctgaagaaga tgcagaatat gtttcatcag atgaaaatga	4680
tgagtcagaa aatgaacaag attcctcatt acttgacttt ttaaagtatg aaactcaact	4740
ttcacaggct ataaatgatt ctgaaatgag agctattttac atgaaatctt tgcgtagtcc	4800
aatgatgaac aataagtaca aaatgattca taagacacat aaaaacataa acattttctc	4860
gcagattcct gaacaagatg aaacctattt agaggatagt ttttgtgttg atgaagagga	4920
gtcttgcaaa ggccaatcaa gtgaagaaga agtttgtgtt gattttaact taataactga	4980
tgattgcttt gcaaatagta aaaagtataa aactcgacgt gcagtaatgc taaaagaaat	5040

gatggaacaa aattgtgcac attcaaaaaa gaaattatcc agaattatTTt taccagatga	5100
ttcaagtgag gaggagaaca atgtaaatga taaaagagaa tctaataattg cggTTaaccC	5160
aagcactgtt aagaagaaca aacaacagga ccattgtTTa aattcagtgc cttctggatC	5220
ttctgCgcag tccaaggTgc gttctactcc aagagttaat ccattagcaa agcagagcaa	5280
acagacatcg ctgaattTaa aggatacaat ttccgaagtc tcagactTca aacctcagaa	5340
tcataatgaa gtccagtcta ccacaccacc cttcactact gttgattcac agaaagactg	5400
tagaaaattt ccagttccac agaaggatgg tagtgctTTg gaggattcta gcacttcagg	5460
ggcatcctgt tccaagtcaa gaccacattt agctgggaca catacttctc ttagacttcc	5520
gcaggaagga aaaggaacct gtattcttgt aggtggTcat gaaatcactt ctggattaga	5580
agtaatttct tccctaagag caattcatgg gttgcaagta gaagtttgtc ctcttaatgg	5640
ctgtgattac atcgtgagta atcgcatggT ggtggaaagg aggtctcaat ctgagatgtt	5700
aaatagtgtc aataagaaca agttcattga gcagatccag cacctgcaga gtatgtttga	5760
aagaatatgt gtgattgtgg aaaaggacag agaaaaaaca ggagacacat caaggatgtt	5820
taggagaaca aagagctatg acagcctgct gactacctta attggcgctg gaatccgaat	5880
tcttttcagt tcctgccaaG aagaaaccgc agatttgcta aaggaactgt ctttagtgga	5940
acaaagaaag aatgtttgga ttcatgttcc aacagtggTg aatagtaata aaagtgaggc	6000
actccagttt tatttaagta ttcccaatat aagttatata actgcattaa atatgtgtca	6060
ccagttttca tctgtgaaaa ggatggctaa cagctcactt caagaaatct ccatgtatgc	6120
acaagtaact catcagaagg ctgaggagat ctatagatat attcactatg tatttgacat	6180
acaaatgtta ccaaTgatc ttaaccaaga tagactgaaa tctgatatat aatcaagctg	6240
ctcaagatgg ggTTTTcaa gacctctcac aatattaaat gcacttcaat aatcattgct	6300
gttttatgtt tatttgtaaa taagagaata ttttattTaa atattttata ttgtatacat	6360
ttttatttat agattataga aattattaaa aaagaaaaat ctgatgttca gtgatcattt	6420
tgactagatt ataaaactaa tttttcttat taaataaaac aaggtttatt aaaagtgtta	6480
ctaaggatag ttttaagaaG taaaagctaa gctagagata tactttggaa tgtttcccaa	6540
aattaaagt gtactgttgt gataaatagt aaagttgaca tgtctatgac tacagccaac	6600
ttgtcgattt tccctatgtg tagatagtat acttttaagt gtactgattc taaatacatg	6660

tacttggttaa ggtgtgggtg atgggtgggt tgtgagataa atgaccagtg aactaggaaa	6720
gtagaaaact taactgaatg tttatctgac caaagggtgtg tcccagttaa gtactgtcaa	6780
atctattaat atgaactctg atatggtttg gctgtgtccc caaccaaagt ctcatcttga	6840
cttgtaatct gaattataat cccaatatat tggggaggga cctcctggaa cgtgattagc	6900
tcatgggggc ggttccccc tgctgttcta gtgatagttc tcagaggatc tgatggtttt	6960
ataagctttt cctctgttca ctctgcagtt ctcttgccca ctgcatgtg gaaaaggaaa	7020
cgtttgcttc ccctccacca tgattgtaag ttcccagggc ctcccagcc atgcaggact	7080
gtgagtcaat taaacatctt ttccttataa a	7111

<210> 44

<211> 3702

<212> DNA

<213> Homo sapiens

<220>

<223> >FBX018|ENSG00000134452|ENST00000379999|3702

<400> 44

ctgctttcaa ggtgtctaga tacagagtga agaatgtcac agtaggaagt gcactggact	60
aggggagtcg ggagaccgg gttcttgtgt tggcaccatc actcatgagc tacgaggtga	120
cttcaggctg ccattggacc tgtcaagtgc ctgagtcatg tgataatggg ctacattgcg	180
cagggcccct gggccatctc cacaggagat gccagaggac gagtgccac ttgctggtct	240
tcacagagca cgctgaaatg agacggttta agcgggaagca tcttactgcc attgactgcc	300
agcatttggc tcggagtcac ttggctgtga cccagccctt cgggtcaaaga tggacaaaca	360
gagatccgaa ccatggcttc tatcctaaac cgagaacaaa aagaggagtg aggggtcagg	420
gaagtcaaag atgcatccct gagttcttcc tagcaggcaa gcagccgtgc accaatgaca	480
tggccaaaag caattctgtt ggccaggaca gctgtcagga ctctgagggt gacatgatct	540
ttcctgcaga gagcagctgt gcactgcctc aggaaggcag tgcagggccg ggctcaccag	600
ggtctgcccc gccctccagg aagcgggtctt ggtcctctga ggaagagagt aaccaggcta	660
ccgggaccag ccgggtgggat ggagtttcta agaaagctcc acggcaccat ttgtctgtgc	720
catgcacaag gcctaggag gccaggcaag aagcagagga cagtacgtct cggctctctg	780
cggagtctgg tgaaaccgac caagatgctg gggacgtggg tcctgatccc attcctgact	840

catactatgg gcttcttggg accttgccct gccaggaagc actgagccac atttgcagcc	900
tgcctagtga ggtcctgagg cacgtgtttg ccttcctccc ggtggaagac ctctattgga	960
acctgagctt ggtgtgccac ttgtggaggg agatcatcag tgacccgctg ttcattcctt	1020
ggaagaagct gtaccatcga tacctgatga atgaagagca agctgtcagc aaagtggacg	1080
gcatcctgtc taactgtggc atagaaaagg agtcagacct gtgtgtgctg aacctcatac	1140
gatacacagc caccactaag tgctctccga gtgtggatcc cgagagggtg ctgtggagtc	1200
tgagggacca cccctcctc cccgaggctg aggcgtgtgt gcggcaacac ctccccgacc	1260
tctacgctgc tgccgggggt gtcaacatct gggccctggg ggcggtgtg gtgctcctct	1320
ccagcagtgt gaatgacatc cagcgactgc tcttctgcct ccggagacc agctccacgg	1380
tgaccatgcc agatgtcacc gagaccctgt actgcatagc cgtgcttctc tacgcatga	1440
gggagaaggg gattaacatc agcaatagga ttcactacaa ctttttctat tgcctatatac	1500
ttcaggagaa ttcctgcact caggccacaa aagttaaaga ggagccatct gtctggccag	1560
gcaagaaaac catccaactt acacatgaac aacagctgat tctgaatcac aagatggaac	1620
ctctccaggt ggtgaaaatt atggcctttg ccggcactgg gaagacctca acgctgggtca	1680
agtatgcaga gaagtggctc cagagcaggt ttctgtatgt gacattcaac aagagcatcg	1740
caaagcaggc cgaacgcgtc ttccccagca acgtcatctg caaaaccttc cactccatgg	1800
cctacgggca catagggcgg aagtaccagt caaagaagaa gttgaatctc ttcaagttaa	1860
caccttcat ggtcaactcc gtccttgctg aagggaaggg tggattcata agagccaagc	1920
ttgtgtgtaa gactctagaa aacttctttg cctcggctga cgaagagctg accattgatc	1980
acgtgcctat ttggtgtaag aacagccaag gacagagagt catggttgag cagagtgaaa	2040
aactgaatgg tgtccttgaa gcgagccgcc tctgggataa catgcggaag ctgggggagt	2100
gcacagaaga ggcgaccag atgactcatg acggctactt gaaactctgg cagctgagca	2160
agccttcgct ggcctctttt gacgccatct ttgtggatga ggcccaggac tgcacaccag	2220
ctatcatgaa catagttctg tctcagccat gtgggaaaat ctttgtaggg gacccgcacc	2280
agcagatcta taccttcggg ggtgcgggtca acgccctgtt cacagtgcc cacaccacg	2340
tcttctatct cacgcagagt tttcggtttg gtgtggaaat agcttatgtg ggagctacta	2400
tcttgatgt ttgcaagaga gtcaggaaaa agactttggg tggaggaaac catcagagtg	2460

gcattagagg tgacgcaaag gggcaagtgg ccttgttgtc ccggaccaac gccaacgtgt	2520
ttgatgaggc cgtacgggtg acggaagggg aattcccttc aaggatacat ttgattgggg	2580
ggattaaatc atttggattg gacagaatca ttgatatttg gatccttctt cagccagagg	2640
aagaacggag gaaacaaaac ctcgtcatta aagacaaatt tadcagaaga tgggtgcaca	2700
aagaaggctt tagtggcttc aagaggtatg tgaccgctgc cgaggacaag gagcttgaag	2760
ccaagatcgc agttgttgaa aagtataaca tcaggattcc agagctgggtg caaaggatag	2820
aaaaatgcc aatagaagat ttggactttg cagagtacat tctgggcact gtgcacaaag	2880
ccaaaggcct ggagtttgac actgtgcatg ttttgatga ttttgtgaaa gtgccttgtg	2940
cccggcataa cctgccccag cttccgcact tcagagttga gtcattttct gaggatgaat	3000
ggaatttact gtatgttgca gtaactcgag ccaagaagcg tctcatcatg accaaatcat	3060
tggaaaacat tttgactttg gctggggagt acttcttgca agcagagctg acaagcaacg	3120
tcttaaaaac aggcgtgggtg cgctgctgcg tgggacagtg caacaatgcc atccctgttg	3180
acaccgtcct taccatgaag aagctgcca tcacctatag caacaggaag gaaaacaagg	3240
ggggctacct ctgccactcc tgtgcggagc agcgcacagg gccctggcg ttcctgacag	3300
cctccccgga gcaggtgcgc gccatggagc gcactgtgga gaacatcgta ctgccccggc	3360
atgaggccct gctcttcctc gtcttctgag gacaaggcgc acgttctccg cagtgcagag	3420
cagcttgccg aggacccgc gtgaagaaag ccagcgaggg gggcttctgc tccctgagac	3480
tctgggttca cccacagcac tttctgagga agaggacacc agcccaagct ggacctgcca	3540
tttctccact ccctacagac agccagtctc cacttgcttc ccctctggat gtatctggtc	3600
agggaaagtgg gggatgttct tttgataaaa aaaaaaaaaa aaatttatgt atttaaactt	3660
ttattacaag atttcaatta aacaggcacc atagcactgg ca	3702

<210> 45

<211> 4977

<212> DNA

<213> Homo sapiens

<220>

<223> >FBXW7|ENSG00000109670|ENST00000281708|4977

<400> 45

cccgcccctc tctctggagt gaggcgagag ccccgcacag agcgagggag acagcgagct	60
gagctccggg cgctgccgt gccgctgccg ccgccgccgc cgctgagact gagagcgaag	120
gagcatccga gagatccagt ccccctgcac tggccgccgc cgagaccttc gctctcacct	180
gggccagcgg gagccgcggc cgcactcctt tccccccctc accttcccgg ccggcagcgg	240
cggctgcaca cgccggagcc ggagccagag ccggagcccg agcctgagcc ggagccggcg	300
gcttgggggg cagggaggcg gctaccacgg gccgggagtg ggtagctgct ccgcggtgag	360
agaacgctga ggaggcgcca gagcttctgc ctctgtccgt ggggcgtggg gcgagacccc	420
caaggtgtag ggaggggggt cccagccgca gcgacacatg cgggagccgg gagcgggggc	480
ggcgccgagc ggagccggcc gggtcctctg ctttgccgcc gactcggcca cccgccggg	540
gccgtagcat cttgccccgg agtgtatgaa ccggggcccc aaccaagctc ggcaaccacc	600
ccccggccgg gggggcgggg accccgatgt gaagcggcgg ctggggcggc ggagagaaca	660
ggaccgacgc cgccgtcctt tcctcacctt cccctcccc tcagccccct ccgggggtct	720
tctcccttgg ccagtcgccg gccccccggc tccttggtg gactccggga ggagttccta	780
gagccccct ccccgcccc agtcccagg gcggcggggc cgggggggac ccggggggcc	840
ggccgcagcc tccaccaga ggaaacttta caaaaacaaa atccggagtc tccaaacct	900
gactgtcccg ggagaagtgg ccctggacgg gcagaagccg cagcctgaaa agaccagga	960
agaggaaaag aggaggtaag cggggccgcc gcctcccctc ccctggcagc gcggaagaga	1020
ccggggttgc cgctggttt agcgacacga gcaccgttc ttcctcagta ccgcgccgga	1080
gccttccgca gctgccgtt cagtccgaag gaggaaggga accaaccac tttctcgcg	1140
ccgcggctct tttctaaaag taatgtgaaa acctttgcat cttctgatag tctagccaag	1200
gtccaagaag tagcaagctg gcttttggaa atgaatcagg aactgctctc tgtgggcagc	1260
aaaagacgac gaactggagg ctctctgaga ggtaaccctt cctcaagcca ggtagatgaa	1320
gaacagatga atcgtgtggt agaggaggaa cagcaacagc aactcagaca acaagaggag	1380
gagcacactg caaggaatgg tgaagttgtt ggagtagaac ctagacctgg aggccaaaat	1440
gattcccagc aaggacagtt ggaagaaaac aataatagat ttatttcggt agatgaggac	1500
tcctcaggaa accaagaaga acaagaggaa gatgaagaac atgctggtga acaagatgag	1560
gaggatgagg aggaggagga gatggaccag gagagtgacg attttgatca gtctgatgat	1620

agtagcagag aagatgaaca tacacatact aacagtgtca cgaactccag tagtattgtg	1680
gacctgccccg ttcaccaact ctctctccca ttctatacaa aaacaacaaa aatgaaaaga	1740
aagttggacc atggttctga ggtccgctct ttttctttgg gaaagaaacc atgcaaagtc	1800
tcagaatata caagtaccac tgggcttgta ccatgttcag caacaccaac aacttttggg	1860
gacctcagag cagccaatgg ccaagggcaa caacgacgcc gaattacatc tgtccagcca	1920
cctacaggcc tccaggaatg gctaaaaatg tttcagagct ggagtggacc agagaaattg	1980
cttgcttttag atgaactcat tgatagttgt gaaccaacac aagtaaaaca tatgatgcaa	2040
gtgatagaac cccagtttca acgagacttc atttcattgc tccctaaaga gttggcactc	2100
tatgtgcttt cattcctgga acccaaagac ctgctacaag cagctcagac atgtcgctac	2160
tggagaattht tggctgaaga caaccttctc tggagagaga aatgcaaaga agaggggatt	2220
gatgaaccat tgcacatcaa gagaagaaaa gtaataaaac caggtttcat acacagtcca	2280
tggaaaagtg catacatcag acagcacaga attgatacta actggaggcg aggagaactc	2340
aaatctccta aggtgctgaa aggacatgat gatcatgtga tcacatgctt acagttttgt	2400
ggtaaccgaa tagttagtgg ttctgatgac aacactttaa aagtttggtc agcagtcaca	2460
ggcaaatgtc tgagaacatt agtgggacat acaggtggag tatgggtcatc acaaatgaga	2520
gacaacatca tcattagtgg atctacagat cggacactca aagtgtggaa tgcagagact	2580
ggagaatgta tacacacctt atatgggcat acttccactg tgcgttgat gcatcttcat	2640
gaaaaaagag ttgttagcgg ttctcgagat gccactctta gggtttggga tattgagaca	2700
ggccagtgtt tacatgtttt gatgggtcat gttgcagcag tccgctgtgt tcaatatgat	2760
ggcaggaggg ttgttagtgg agcatatgat tttatggtaa aggtgtggga tccagagact	2820
gaaacctgtc tacacacgtt gcaggggcat actaatagag tctattcatt acagtttgat	2880
ggtatccatg tggtgagtgg atctcttgat acatcaatcc gtgtttggga tgtggagaca	2940
gggaattgca ttcacacgtt aacagggcac cagtcgttaa caagtggaat ggaactcaaa	3000
gacaatattc ttgtctctgg gaatgcagat tctacagtta aaatctggga tatcaaaaca	3060
ggacagtgtt taaaaacatt gcaaggtccc aacaagcatc agagtgtgtg gacctgttta	3120
cagttcaaca agaactttgt aattaccagc tcagatgatg gaactgtaaa actatgggac	3180
ttgaaaacgg gtgaatttat tcgaaaccta gtcacattgg agagtggggg gagtggggga	3240

gttgtgtggc ggatcagagc ctcaaacaca aagctggtgt gtgcagttgg gagtcggaat	3300
gggactgaag aaaccaagct gctgggtgctg gactttgatg tggacatgaa gtgaagagca	3360
gaaaagatga atttgtccaa ttgtgtagac gatatactcc ctgcccttcc ccctgcaaaa	3420
agaaaaaaag aaaagaaaaa gaaaaaaatc ccttgttctc agtggtgcag gatgttggct	3480
tggggcaaca gattgaaaag acctacagac taagaaggaa aagaagaaga gatgacaaac	3540
cataactgac aagagaggcg tctgctgtct catcacataa aaggcttcac ttttgactga	3600
gggcagcttt gcaaaatgag actttctaaa tcaaaccagg tgcaattatt tctttatttt	3660
cttctccagt ggtcattggg cagtgttaat gctgaaacat cattacagat tctgctagcc	3720
tgttctttta ccactgacag ctagacacct agaaaggaac tgcaataata tcaaaacaag	3780
tactggttga ctttctaatt agagagcatc tgcaacaaaa agtcattttt ctggagtgga	3840
aaagcttaaa aaaattactg tgaattgttt ttgtacagtt atcatgaaaa gctttttttt	3900
tttttttttt gcccaaccatt gccaatgtca atcaatcaca gtattagcct ctgttaatct	3960
atttactgtt gcttccatat acattcttca atgcatatgt tgctcaaagg tggcaagttg	4020
tcctgggttc tgtgagtcct gagatggatt taattcttga tgctggtgct agaagtaggt	4080
cttcaaatat gggattgttg tcccaaccct gtactgtact ccagtggcc aaacttattt	4140
atgctgctaa atgaaagaaa gaaaaaagca aattattttt ttttattttt tttctgctgt	4200
gacgttttag tcccagactg aattccaaat ttgctctagt ttggttatgg aaaaaagact	4260
ttttgccact gaaacttgag ccatctgtgc ctctaagagg ctgagaatgg aagagtttca	4320
gataataaag agtgaagttt gcctgcaagt aaagaattga gagtgtgtgc aaagcttatt	4380
ttcttttatc tgggcaaaaa ttaaaacaca ttccttgga cagagctatt acttgcctgt	4440
tctgtggaga aacttttctt ttgagggct gtggtgaatg gatgaacgta catcgtaaaa	4500
ctgacaaaat attttaaaaa tatataaaac acaaaattaa aataaagttg ctggtcagtc	4560
ttagtgtttt acagtatttg ggaaaacaac tgttacagtt ttattgctct gagtaactga	4620
caaagcagaa actattcagt ttttgtagta aaggcgtcac atgcaaaca acaaaatgaa	4680
tgaaacagtc aaatggtttg cctcattctc caagagccac aactcaagct gaactgtgaa	4740
agtggtttaa cactgtatcc taggcgatct tttttcctcc ttctgtttat tttttgttt	4800
gttttattta tagtctgatt taaaacaatc agattcaagt tggttaattt tagttatgta	4860

acaacctgac atgatggagg aaaacaacct ttaaagggat tgtgtctatg gtttgattca	4920
cttagaaatt ttatitttctt ataacttaag tgcaataaaa tgtgtttttt catgtta	4977

<210> 46
 <211> 2296
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >FEN1|ENSG00000168496|ENST00000305885|2296

<400> 46	
aaaggaagtg cctccggcgc aagtggcatt gagggacttg tagtcctgcg atttcgggtg	60
tagagggagc aggggcctgc ggggacctgg tgtgggtgga gtggggacaa gcggtggaga	120
agggtacgcc agggtcgctg agagactctg ttctccctgg agggactggt tgccatgaga	180
gcagccgtct gaggggacgc agcctgcact acgcgcccc aagggtgtg cgtggcgagc	240
aggtcacgtg acgggagcgc gggctttgga aggcggctga acgtcaggcc acccgccgt	300
aagctgagaa gggagagcga gcttaggacc gcctgcccgg ggcaacccc aaccaagctt	360
tagccgccga ggccgcgtgt ccaaaggcc agtcatccct cctctgtgtt gccatgggaa	420
ttcaaggcct ggccaaacta attgctgatg tggccccag tgccatccgg gagaatgaca	480
tcaagagcta ctttggccgt aaggtggcca ttgatgcctc tatgagcatt tatcagttcc	540
tgattgctgt tcgccagggt ggggatgtgc tgcagaatga ggagggtgag accaccagcc	600
acctgatggg catgttttac cgcaccattc gcatgatgga gaacggcatc aagcccgtgt	660
atgtctttga tggcaagccg ccacagctca agtcaggcga gctggccaaa cgcagtgagc	720
ggcgggctga ggcagagaag cagctgcagc aggctcaggc tgctggggcc gagcaggagg	780
tggaaaaatt cactaagcgg ctggtgaagg tcactaagca gcacaatgat gagtgcaaac	840
atctgctgag cctcatgggc atcccttatc ttgatgcacc cagtgaggca gaggccagct	900
gtgctgccct ggtgaaggct ggcaaagtct atgctgcggc taccgaggac atggactgcc	960
tcaccttcgg cagccctgtg ctaatgcgac acctgactgc cagtgaagcc aaaaagctgc	1020
caatccagga attccacctg agccggattc tgcaggagct gggcctgaac caggaacagt	1080
ttgtggatct gtgcatcctg ctaggcagtg actactgtga gagtatccgg ggtattgggc	1140
ccaagcgggc tgtggacctc atccagaagc acaagagcat cgaggagatc gtgcggcgac	1200

ttgaccccaa caagtaccct gtgccagaaa attggctcca caaggaggct caccagctct	1260
tcttggaacc tgaggtgctg gaccagagt ctgtggagct gaagtggagc gagccaaatg	1320
aagaagagct gatcaagttc atgtgtggtg aaaagcagtt ctctgaggag cgaatccgca	1380
gtggggtcaa gaggctgagt aagagccgcc aaggcagcac ccagggccgc ctggatgatt	1440
tcttcaaggt gaccggctca ctctcttcag ctaagcgcaa ggagccagaa cccaagggat	1500
ccactaagaa gaaggcaaag actggggcag cagggaagtt taaaagggga aaataaatgt	1560
gtttcccat tatacctcct tcacccaga atatttgccg tcttgtacc ttaagagcta	1620
cagctagaga aaccttcacg ggggtggagag aggattctaa ggcttttcta gcgtgaccct	1680
tttcagtagt gctagtccct tttttacttg atcttaatgg caagaaggcc acagaggtac	1740
ttttcctttt ttagctcagg aaaatatgtc aggctcaaac cacttctcag gcagtttaat	1800
ggacactaag tccattgtta catgaaagt atagatagca acaagttttg gagaagagag	1860
aggagataa aagggggaga caaaagatgt acagaaatga tttcctggct ggccaactgg	1920
tggccagtgg gaggtgatgg tggacctaga ctgtgctttt ctgtcttggt cagccttgac	1980
ccaccttgag agagagccac caggaaggcg catcttagca gatgggagga actgctgaga	2040
gaagatgggc agaaagctgg agcccctgga gttggctgtg tctgtgtttg tgactgatta	2100
ctggctgtgt cttgggtggg cagaaactcg aacttgctat gtaatttggt tctagttatt	2160
cagaggagta agatggtgat gttcacctgg caatcagctg agttgagact ttggaataag	2220
acactggttt tcatgctgctg tttttgtttt aaagttatga agaaaaagt caataaaatt	2280
ctaaaagtaa ccaaaa	2296

<210> 47
 <211> 6367
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >GEN1|ENSG00000178295|ENST00000381254|6367

<400> 47	
gaaacgcttc ctggtgctcc tgggcctggc ggttgagccc ggggagctag tctttgcttg	60
ggtaaagaaa gaggactttc cttttttttt ttccttttga gaaaattcag aaatttgag	120

gcacagtagt taggattcag gccaaactgg taaatgacga aggggcttct tccagagcca	180
aggaggaagg gttgtccggc agataatcac cagaatggga gtgaatgact tgtggcaaat	240
tttggagcct gttaagcaac acatccccctt gcgtaatctt ggtgggaaaa ccattgcagt	300
tgatctgagt ctctgggtgt gtgaggcaca gacagtcaaa aaaatgatgg gcagcgtcat	360
gaagccccac ctgaggaact ttttttttcg tatctcatat ttaacacaaa tggatgtaaa	420
actggtattht gttatggaag gggaaccacc aaagctgaaa gctgatgtca taagcaagag	480
gaatcagtct cggtatgggt cttctggaaa atcgtgggtct cagaaaacag ggagatcaca	540
ttttaaatca gtcttaagag agtgcctcca tatgctcgaa tgcttaggaa tcccctgggt	600
tcaggctgct ggggaagctg aagccatgtg tgcttatctc aatgctgggtg gtcatgtcga	660
tggctgcctc accaatgatg gagatacttt cttttatggg gccagactg tttacaggaa	720
tttactatg aatacaaagg acccacatgt tgactgttac acaatgtcat ctatcaagag	780
taaactaggt ttggatagag atgctctgggt tggattagca atacttcttg gctgtgatta	840
tctcccaaag ggagtccctg gagttggaaa agagcaagca ttaaaactta tacagattht	900
gaaagggcaa agtttacttc agaggthta tcggtggaat gaaacatctt gtaactctag	960
tccacaactg ctagtacta aaaaactggc tcattgttcc gtatgttccc atccaggttc	1020
acctaaggat catgaacgta atggatgcag attatgtaaa agtgataaat attgtgagcc	1080
acatgactat gaatactgct gtccttgatg gtggcaccgt acagaacatg ataggcaact	1140
cagtgaagta gagaacaata ttaagaagaa agcttgctgt tgtgagggat tcccattcca	1200
tgaggthatt caagaattcc ttttaaaaca ggataaattg gtgaaggtht tcaggtacca	1260
aagacctgat ttgttattgt ttcagagatt tactcttgaa aaaatggagt ggcccaatca	1320
ctatgcatgt gagaaattgc tgggtacttht gaccattat gacatgatag aaagaaagct	1380
tggtagcaga aactctaadc aactacagcc aattcgaatt gthaagactc gaatcagaaa	1440
tggagthcat tgthttgaaa tagaatggga aaagcctgaa cattatgcta tggaagataa	1500
acaacatgga gaatttgctt tattaacaat tgaggaagaa tcattgtthg aagcagcata	1560
tcctgagatc gttgctgtht accaaaaaca aaagthgaa attaaaggga agaaacaaaa	1620
acgtattaag cctaaagaaa acaatttgcc agaaccagat gaagthaatg gctthtcagtc	1680
acacatgact ttaaaacca catgtgaaat cthtcataag cagaattcca agthaaattc	1740

ggggatttcc cctgataccta cattaccaca ggaatctatt tctgcctcat tgaatagctt	1800
gctttttacct aaaaatactc catgtttgaa tgcacaagaa cagttcatgt cttctctaag	1860
acctttggct atacagcaaa ttaaaagctgt cagtaagtct ctaatttcag aatctagtca	1920
acccaatacc tcatctcata atatatccgt gattgctgat ctacacttga gcactattga	1980
ctgggaaggt acttctttta gtaattctcc agctattcaa aggaataactt tttctcatga	2040
tttaaaatca gaagttgaat cagagctatc agccatccct gatggctttg aaaatatccc	2100
agaacaactg tcctgtgaat cagaaaggta cactgcaaac ataaagaaag tgttggatga	2160
ggattctgat gggattagtc ctgaagagca tctactttct ggcattactg atttatgtct	2220
tcaggatttg cttttaaagg aacgaatatt taaaaatta tcatatcctc aggataatct	2280
acaaccagat gtcaacctga aaactttgtc catacttagt gtaaaagaat cttgtattgc	2340
taacagtggg tctgattgta catcacatct ttcaaaggat cttccaggaa ttcccttgca	2400
aaatgaatcc agagactcta aaattctaaa aggagaccag ctgcttcaag aagactataa	2460
agtcaatact tctgtccctt attctgtcag taacacagtg gtaaagacct gcaatgttag	2520
accaccaaact actgcttttag atcatagtag aaaagttgat atgcaaacca ctcggaaaat	2580
tttaatgaag aagagtgttt gccttgacag acattcctct gatgaacaaa gtgccccagt	2640
gtttgggaaa gctaagtaca caactcaaag aatgaagcac agttctcaaa agcataattc	2700
atcccatttc aaagaaagt gccataacaa gttgagtagc cctaagatac atattaaaga	2760
aactgaacag tgtgtcagat cttatgaaac agctgaaaat gaagaaagct gtttcccaga	2820
ttcaacaaaa agttctctga gttctctaca atgtcataag aaagaaaaca actctgttac	2880
ttgtttggat agccctcttc ctttacgcca gagattaaaa ctaagattcc aaagcacttg	2940
aaatttaaaa cacttaggta taacttaact attttagtac tatcagcaat agcagagaca	3000
gaggggaaggt atctagttca tgtgtggtaa aaattttaat gttctctgtg tcatgaaaca	3060
cttgccattt taatcaaagt tgtaattttt aaaaagtcac ctaaaactct ggtttttaaaa	3120
gacacctgtg attgaaaact tctgataatg tatgtcatta tgtccttact attccttaat	3180
tgtagtttta aaatattggg atagtacttg acagagtaaa tacttcatct gattgttcat	3240
ttttactttt tcttcacaa gcctctaaag tatttatatt ccagcttggt cccaagagga	3300
taattcttta tacttctctt cattctttta aggccttgca aggtcttccg ttataactcg	3360

ctttcctaaa agctattttc tccctcagtg tgaagatacc tttagtgtgc tcttccactt	3420
tggaaggcca tgcaagggtg tccattataa ctctttttcc taagagttat attctccctc	3480
agtttgaaga tacctttagt gtgctcttcc acttttggtg tcttagtctc ttttgagggg	3540
caaaaataga aaaggagaga agatgtcagt atgttttagta aaaatcatgc tcgtaatggc	3600
tgaataaact gagcaaagta actccttatg tatccccaga agttcacagg tatatcgggt	3660
ggaaaaagat ttggaaaatc aaatgtataa ccaaaggat tagaaactag cccaggatca	3720
cgtagctaac taataatcct gtgggaatca gttttcttgc ctgctaattt ttttattttt	3780
ctattttttc cttctatagc acttttcccc cttttgtttt gaatctatgc aatattgact	3840
ttaataccac caaatattaa gtcatgcatt aatttaggtc gcactcaaaa attctagaga	3900
ggcatccaga ttgaaaagga aaatggtgtc tgcagatgac aagattgtat gcatcaaaaa	3960
ttctaagaaa tccactagaa aactattaaa actgataaag cgagttcatc caggtcgtag	4020
aatacaagac caatgtgcaa tgatcattgc attttgtttt tgtttgtttg ttttttttg	4080
agacagagtc tcactcttgt cgagactggg gtgtagtggc gccatcttgg ctactgcaa	4140
cctctgcctc ccaggttcac gcaattcaca tgcctcagcc acctgagtag cagggattac	4200
agggtgtgtc caccatgcct ggctaatttt tagtatTTTT agtagagatg ggattttccc	4260
ttgttggtca ggctagtcac gaactcctgg cctcaagcga tccacccatc ttggcctccc	4320
agagtgtctgg gattacaggt gtgaatcacc atactcagcc tcaattgtat ttctgtacac	4380
ctctaaataa atggaaatac attccatgtt catgaatcag aagacataag atgaccatac	4440
tccccaaact gatctacagg ttaaagttaa ttcccattaa ttccagctgg tttctttgct	4500
gattctaaaa ttcatatgga aatttcagga acacagaata ttaggaagaa tcttgaaaga	4560
gtttgaatga cccacccttc ctgattttta gtatacagaa ctacagtaat gaatgtgttg	4620
tactggtata aaggtagaca tataggtcaa tatagatcaa tggcatagaa ttgagaattc	4680
agaaaaaaag ctttcacgtt tatggtcagt tgatttttaa ccagagtgcc aagacaattc	4740
aatgggaaaa ggaaacagtc tttgaaccaa ctgtcttctg gatctattgg atatctatgt	4800
aaaaaagagt gacatttgat tcctacttca catgatattt aagaagtaac ttaagatgga	4860
tcaaatacct aaatgtaagc taaagctata aaactcttag aagaaaacat agacgtaaat	4920
ctttgtgact ttggattagg caaggatttc ttaaataatac ccaaagcata ggcaacaaaa	4980

gagaatacac aaactggact tcatcaaac taaaaactgt tgtttgaaat gacaccatca	5040
agaaagtga aagacagctt acagaatggg agaaaatatt tgcaaatcat aaatgtgata	5100
agggacttgt atcttggtgg tatataaaga actcgtaact caattataaa aagacaactt	5160
tgaaatgggc aaaatagctc aacaggcttt tctgcaaaga agaaatacaa atgaccaaga	5220
agcacattta aaaattttca gtactattac tcatcaggaa aatgcaaadc aaaaccacaa	5280
gacaccaat gtctacaatc aaaaagataa taactagtat tgatgaggat gtggagaaat	5340
tgaaattctc ataacatgct ggtaggaatg taaaatgggg cagccacttt ggaaaaagtc	5400
tggtagttct tcaaattggtt aaatgtagag ttacgatatg atccagcaat tcctctcca	5460
ggtatatacc caagataaat gaaaacttat atccacataa aaacctgtgc acaaatgtcc	5520
atagcagcgt tattcataat agcctaaaag tggaaacaat cccagttcca gaatgaggaa	5580
ggggagaaac taatgtgtat tagctattgt gtgctaagca ttcaactaga ttatttacia	5640
accttgatc atctcaactc ttttaaggact gtattgcaat gttttgaata ttcagagaga	5700
aaaaagtcgt tgctaaaaca ttttccaagg ttctgcttat tctgatttgt tcagtcgtgg	5760
ctgtgatagt tcaggacat ctagaccagg taaataaaat atctagaggc attcttgaga	5820
ttgtatgaga tgaaaataac aaaattagtt gggagtggcc agtctgagtt cattttgcta	5880
tatagctcag gagtcccaa cccctgggct atggacaggt accagtccat cacctgtcag	5940
gaaccggcct gcacagcagg tgttgagcgg catgcaagcg agaattaccg cctgagctct	6000
gcctcctgtc agatcagcgg tggcattaga gtctcatagg agcatgaacc ctattgtgaa	6060
ctacgcatgc gaggatctag tttgtgcact cctactgaga atctaatacc tgatgatata	6120
aggtggaaca gttcatccc acagccatct cccctcatc atccgtggaa aaactgtttt	6180
ccacgaaagt ggtcctcggg gccagaaagg ttggggactg ctgatatacc taatataggc	6240
ataagtgtaa aattatagtt cctcctctc aaggatttct ctgttttcat tgtccagggt	6300
ataccaattc tttttaagc attccatgtt cttgttcctt gatgtggtca tgttatcact	6360
tgtaaat	6367

<210> 48

<211> 1614

<212> DNA

<213> Homo sapiens

<220>

<223> >H2AFX|ENSG00000188486|ENST00000530167|1614

<400> 48

acagcagtta cactgcggcg ggcgtctgtt ctagtgtttg agccgtcgtg cttcaccggt	60
ctacctcgct agcatgtcgg gccgcggcaa gactggcggc aaggcccgcg ccaaggccaa	120
gtcgcgctcg tcgcgcgccg gcctccagtt cccagtgggc cgtgtacacc ggctgctgcg	180
gaagggccac tacgccgagc gcgttgggcg cggcgcgcca gtgtacctgg cggcagtgct	240
ggagtacctc accgctgaga tcctggagct ggcgggcaat gcggcccgcg acaacaagaa	300
gacgcgaatc atccccgcc acctgcagct ggccatccgc aacgacgagg agctcaacaa	360
gctgctgggc ggcgtgacga tcgcccaggg aggcgtcctg cccaacatcc aggccgtgct	420
gctgccaag aagaccagcg ccaccgtggg gccgaaggcg ccctcgggcg gcaagaaggc	480
caccagggc tcccaggagt actaagaggg cccgcgccgc ggccggccgc caggcctccc	540
catgccacca caaaggccct tttaagggcc accaccgcc tcattggaaag agctgagccg	600
cttcagactg cggggcaagc gggccgcggc tcccttcccc tcccctcccc tcgcccgcct	660
tcgccgcccg gcctcgagtc cccgcccgc cccgctccc tcccgcaccg cctgccgcgt	720
cggcctcggg ccctgccctg tccgccgtcc gccctccgtt agggttcggg ctttcggat	780
gcggccttggg cgctcttcgg ggacctcgt ggcgcggaag acccgagcct gccgggggga	840
ggccggcggc gccgcacctg cccgcctcgg cgttcgtgac tcagccgcc catcccgagt	900
cgctaagggg ctgcggggag gccgcagcac cttctggaag acttggcctt ccgctctgac	960
gcagggccga ggtgggcagt ccaggccgag aggcggcgcg ccctgaaggt gagtgaggcc	1020
ctcggcagct gcagccgggg tgtctggtac cccccggcg tgggtgcttag cccaggactt	1080
tcagacgcgg ccgctggccg ggaggctttg gtgggagaga cgcatcgcc gatttcggtc	1140
tggcgcccct tctgcggccg ggaccaggc ctttcacatc agctctccct ccatcttcat	1200
tcataggtct gcgctggggc cgggacgaag cacttggtaa caggcacatc ttctcccga	1260
gtgactgcct cctaggagga catttagggg agggcagagg cctgcagttt ggcttcacgg	1320
ctggctatgt ggacagcaag agtcgttttc gcggaagccg actggcagcc aggcctgtcg	1380
ggccccccga cgccgcccc tttcccttcc agcaaactca actcggcaat ccaagcacct	1440
agataccagc acaagtcggt taatccctgt ctggactgag cctccgttgg cttctgaact	1500

ggaattctgc agctaaccct tccacgacta gaaccttagg cattggggag ttttagatgg	1560
actaatTTTA ttaaaggatt gttttttttt taaatggagc gtttgttttc attt	1614

<210> 49
 <211> 9874
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >HDAC2|ENSG00000196591|ENST00000519065|9874

<400> 49	
gcaggttgtt caccctaaca gggcctctaa cctcgagccc gaaacggcct ctttctccag	60
ccgtgcccct tcctcccccg tggaggccgc agtggatgcg ctcacctccc tgcggcctcc	120
tgaggTggtt tggTggcccc ctctcgca gttggtgccg ctgccacctc cgattccgag	180
ctttcggcac ctctgccggg tggTaccgag ctttcccggc gccccctcct ctctccac	240
cggcctgccc ttccccgagg gactatcgcc cccacgtttc cctcagccct tttctctccc	300
ggccgagccg cggcggcagc agcagcagca gcagcagcag gaggaggagc ccggtggcgg	360
cggtggccgg ggagcccatg gcgtacagtc aaggaggcgg caaaaaaaaa gtctgctact	420
actacgacgg tgatattgga aattattatt atggacaggg tcatcccatg aagcctcata	480
gaatccgcat gaccataac ttgctgttaa attatggctt atacagaaaa atggaaatat	540
ataggcccca taaagccact gccgaagaaa tgacaaaata tcacagtgat gagtatatca	600
aatttctacg gtcaataaga ccagataaca tgtctgagta tagtaagcag atgcagagat	660
ttaatgttgg agaagattgt ccagtgtttg atggactctt tgagttttgt cagctctcaa	720
ctggcggttc agttgctgga gctgtgaagt taaaccgaca acagactgat atggctgtta	780
attgggctgg aggattacat catgctaaga aatcagaagc atcaggattc tgttacgtta	840
atgatattgt gcttgccatc cttgaattac taaagtatca tcagagagtc ttatatattg	900
atatagatat tcatcatggt gatggtgttg aagaagcttt ttatacaaca gatcgtgtaa	960
tgacggtatc attccataaa tatggggaat actttcctgg cacaggagac ttgagggata	1020
ttggtgctgg aaaaggcaaa tactatgctg tcaattttcc aatgagagat ggtatagatg	1080
atgagtcata tgggcagata tttaagccta ttatctcaaa ggtgatggag atgtatcaac	1140

ctagtgtgt ggtattacag tgtggtgcag actcattatc tggtgataga ctgggttgtt	1200
tcaatctaac agtcaaaggt catgctaaat gtgtagaagt tgtaaaaact ttttaacttac	1260
cattactgat gcttggagga ggtggctaca caatccgtaa tgttgctcga tgttggacat	1320
atgagactgc agttgccctt gattgtgaga ttcccaatga gttgccatat aatgattact	1380
ttgagtattt tggaccagac ttcaaactgc atattagtcc ttcaaacatg acaaaccaga	1440
acactccaga atatatggaa aagataaaac agcgtttgtt tgaaaatttg cgcatgttac	1500
ctcatgcacc tgggtgtccag atgcaagcta ttccagaaga tgctgttcat gaagacagt	1560
gagatgaaga tggagaagat ccagacaaga gaatttctat tcgagcatca gacaagcgga	1620
tagcttgtga tgaagaattc tcagattctg aggatgaagg agaaggaggt cgaagaaatg	1680
tggctgatca taagaaagga gcaaagaaag ctagaattga agaagataag aaagaaacag	1740
aggacaaaaa aacagacgtt aaggaagaag ataaatccaa ggacaacagt ggtgaaaaaa	1800
cagataccaa aggaaccaa tcagaacagc tcagcaaccc ctgaatttga cagtctcacc	1860
aatttcagaa aatcattaaa aagaaaatat tgaaaggaaa atgttttctt tttgaagact	1920
tctggcttca ttttatacta ctttggcatg gactgtattt attttcaa at ggctttttcg	1980
tttttgtttt tcttggcaag ttttattgtg agtttttcta attatgaagc aaaatttctt	2040
ttctccacca tgctttatgt gatagtattt aaaattgatg tgagttatta tgtcaaaaaa	2100
actgatctat taaagaagta attggccttt ctgagctgat ttttccatct tttgtaatta	2160
tctttattaa aaaattgtac ttggattatc ttttgtctgt ttattactac aatatgaagt	2220
cttgtttcag tggctaata ga catcatttct gtagacttac aatacactct aggtgaaaga	2280
taatgattac agcttgaaag ataactattt gctgttttctt tgggaagagt atttatagta	2340
attattactt atctttgcaa tagaaattct accaccttgc cctctatagc ttagccagtt	2400
agtatcagtg aagattaaca tcccattaca atttatgaaa taatacagac tctgcaattg	2460
agatgtagga gttctttgag ttgacccaaa gatttctcaa attgaaatgg aaattctgaa	2520
ttgaaagaag aaactgacca gaacttgat tgaccagact tgcctatagt atattgctgg	2580
tttaaaatgg aacctgcaga caaacctgt ttcttttact gcatttacat ggcattccagg	2640
ttccattatt atttacgtga catccaggtt tctaactcaa ggaaataaac acaactgatt	2700
tatcattcag caactactta ttgtgtgtct gccatttta agactgtagg agtgtaactg	2760

aatataatgg aaaaatgctt tcactcagag cttacactca gagcttacat tctagtagca	2820
ggaagcagac aaatgggggt aactcctgct actaagatgt gcaaagaaga caaaacattt	2880
taggaacttg ccaaaatcag tgaaatctcc ctttttgtca ggcccacatt gattcttttg	2940
agattaaaaa ttacagaatg ccagaagata attcagtcaa aagtatttct cttcagtgca	3000
gtaaaatatt aaaagaaaaa atatttctct acaagcctct taaatgtttc agacattcac	3060
aatagcacct agacttttgt aatgaacagc tgtgcaccca ccccgactta acaaattttt	3120
tacttttgct atatttgctt cgggtgtttt ttcttaaaat acaccgaaat tgaagctctc	3180
tttgtgtact ttgtcgtttg taatcctccc tccctcccag ccttaagaag taatcacct	3240
tgtgattatt ttatccatgt ttttgtgaat attttaccca tgtttttgta cttttgctac	3300
atatttgta tcaatctgta aacctttatg acattaggaa ctaagaaact tagtcccttc	3360
gtaggggga taatgaaatg tatttagtgt ttgtgaaaca tagatggtat gtatttgac	3420
aattctgtaa ctttgctttt tttattttta tttttccata gcttattggg gaacaggtgg	3480
tgtttggtta catgattaag ttcttttagtg gtgatttgtg ggattttggt ggacccatca	3540
cccaagcagt gtacactgca ccctatttgt aatcttttat ccctcgcccc cctcccacca	3600
tgctcccgt ctacatgat gatcctgttt taaataagaa aataccattt cgcaggctcc	3660
agatgttctg gcacccctcc tgtggatttc ccagtcctg cagctcacag gacaacaggg	3720
gctgtggtag agtcacctat gagatcctgg agtagtggat ggaggagatg gaacagtga	3780
gacggaaact gagctcagta tccgggtgcc aggagacaaa ggccctttgc ttttttcat	3840
ttaatatct gatctacccc tgttgacaca tgtaagtat agttcatttt gactgctatg	3900
tattatgttc cattgtgtga acatactgaa attgtacact tcaatactat actggatctc	3960
cttgggtgta ttttaagagg tttgtttttc taagtagttg gttatataca actaaaacct	4020
caagagaact atctaaagca atttcagcaa ggtgatttgg tacagcatta ataaacagaa	4080
atcagtaaca cttagtgacc aagtctgttg gaagaacaaa gaccccat tghtaataaca	4140
aaatttttag aaataatatg taaagaagct atggttcttg tgtctagtaa ggtcaatgta	4200
acatagtaag atgtcagaat accctaatac tttaaaaaat tcatatagga taaaaatgat	4260
atttgaaatt ggcaaggaaa gacattattt tgtaagtgga attgggacaa caactggtaa	4320
ccaaatggaa aaccagttt tctgccctcc actagaagat ttaaatagga aaagataaaa	4380

ctacaaaaa cctaaactct taaagcaa	ggattgaaga aggcctaagt gtgacaccaa	4440
actcaactat aaaagatata tttgataaca	aaaaaaaaat tagttcagtg gaccaacaaa	4500
aacttagaag acaagtcaag aaaaatgaca	aaagacagag tgggaggcag atttgtaact	4560
catccaggtc aaaaggctca tatctaaaga	tagtagagga acaaaatgta taaggatgtg	4620
aactgggaaa caaatacata taaatagttt	gtaaatatga aaagatcttt aacctcagta	4680
aataaaaagc tatagagaga cttatTTTTT	aacttagttt tttaaaccta ttatgtttat	4740
ttatTTTTTt tTTTTtgag acggagtctc	gctgttgccc aggctggtgt gcaatggcgc	4800
aatctcggct cactgcaacc tccgcctccc	aggttgaagc cattctcctg cctcagcctc	4860
ctgagtagct gggattacag ggcctgtca	ccacgccag ctaattgttc gcatttttag	4920
tagagacggg gtttactat gttggccagg	ctggcctcga actcctgacc tcatgatcca	4980
cccaccttgg ctttcaaag tgctgggatt	acaggtgtga gctgccgcac ctggctgttt	5040
atctctTTTT tttagagaaa ggatctcgg	cacccaggat ggagtgtggt agcctgatca	5100
tatctcacta catcttagaa cttctaggct	tagggttattc tcccacctca gcctcccaag	5160
tagctgggac tacaagtgtg caccatcaca	cctagctgat ttttacattt ttatTTTgga	5220
gagatggtgt ctccctgtgt tgcccaggct	ggtcacaaac tcctaggctc aagcgattct	5280
cctgactcag gcatgagcca ccgtaccgg	cctaaaccta tcatgttaca gacttagaaa	5340
gcaactattg tcaagtgttt gaggaaactc	aggtcagggt ttgtaaaacta agatattaac	5400
tcaagtaaag ctctttaatt catttaatga	aggtgccaca ttgtctcagt tctctatggc	5460
atgggtgaat gctgttctaa gtcagcattg	gtactctaag ctagttacat catatctaag	5520
ctttgccctt ctaccagagc tgctagcatt	ctgtcaatgg gcaattattt gaagttctta	5580
cattgaagtt agtcacctac atcttctgtt	ttttatgggt ttgatgtagt aatactgctg	5640
aagTTTTTTt gataactgcg attcataata	tttgttattc atattgtgat ataaatgata	5700
agggtctttg aaaacaagta gtaatcattt	caatagctta ggatctccac cataatctta	5760
ggaaaattac taacctctgt gcctcagttg	cttcatcatt taaaatgagg aaaataatag	5820
tccctactta ataggtttgt tgtgaggatt	gagttaataa catagttaat gctcagtaag	5880
ggttagctgc tatTTTTTTt tctTTTTTTt	ttgaaagagt ctactctgt tgcccagggt	5940
ggagtgcagt ggcatgatct tggctcactg	catcctctgc ttcctagggt cagggtgattc	6000

tcatgcttca gcctcccaag cagctgggag tacaggtgtg caccaccaca cctggctaata	6060
ttttgtatat ttagtagaga cgggttctca ccatattgtc caggctgggc ttgaactcct	6120
tacctcaaata gatccacccg cctgggcctc ccaaagtgt gggattacaa gcatgagcca	6180
ccgcgcctgg cttgctattg ttatgaggta aaggtagata gatgggtgag agtggtgcca	6240
ggggaagtgt taaatttttg agtggtcctt tagatgccag atgggttgta tctgagcctt	6300
ttattgcagt ttgatgccta ctagtgtgaa gactactagg tcatagtgga tagagaagca	6360
atcttttgga gacctgattt tagcaaggat acgaataata ttgacaact ttggggggat	6420
cttgatgcct ctgtaattta ctcaaggata atctcaagaa aaatggcatt aagtagatta	6480
cagaaaaaat agaactatca tattgttatt attggctatt tacatgagca atgcggagaa	6540
atgttttagga ttacagcatt tagaagcttc tcaattgctg ctttctctca ctgtaccaca	6600
agatggcaga tactgcattt aaaatttttt tttctgtgtg ttttctctta tagtcacttg	6660
gtggccatgt aacaagcaga gcaacatgta ttaacagatt ctttttgaat gcaatattgg	6720
attaaaaact ttgaattaaa ctacaattaa gtttactgat ttaggttttc ctataagaat	6780
tataacacta atttctaagt ttacatctg caaaacacag tagagtgtat ctggtagtca	6840
ttgactaatt gtgactttgc tagtcagaca aaatggaact gaatattaga gccagaataa	6900
ttttctgttt tgactgggtg taagaatcag ttagtataca gaatgggaaa gaattgaaat	6960
ttatttctag gaattatagt tatagtgaag ttaatattca ggggcattgg gtgcttgctc	7020
cacgactttt gtttactgaa ctaagtacct aaccctttg gtctttgaac actgcagagt	7080
tatgattagg attattaatg aaaaattctt ttttagtaga acaaacagtt aaggacagaa	7140
gacaccttga ttagattctg ttatttgctc tttttattaa gattgttttt actctaaatg	7200
ttaaagaaat gtataattat ttctaaggta aaatgagtag caggattctg aataggagat	7260
tttttttggc cagaccatt ttgcagccag tggtttgat ttaggagtga gttgaacaaa	7320
gaagactcta atctgttttt gtgttggtga ttttttattc tattgataaa tcagtagctt	7380
ccagcttctc ttgtatcctc tactagtcac ttttttgtt tcttcagaaa taaacatcca	7440
gggtcacttc cctgcccttg ccatatactt actaggatga tgtacaggta aaacttttaa	7500
aggaaagtac ggacaggcaa ttcctaaagt aaaagagagc acattatcaa aactggctgt	7560
acttaacaaa taccaacaaa aagtcattac acttctggca acaaataatg tttgatttaa	7620

aaagacaaga atttcccctt ttgctcttat tttttaaata tgctatgtcc ttaagtgaga	7680
aatgcacttc aaaaattttc ttattctgaa gtctagccgg aaagtttgaa cttgtgttgc	7740
ctttggctag actaggtgaa attaaggggc ttaagagggc actgttgtgt ggccctgtgg	7800
tagataatag ccctccaagc atgaggatgg gagatctacc attcaccagt atttccaaag	7860
ttatatctct aaaaaggcta gatttctaata ttccgtagaa tctaggtatg ttcttttcaa	7920
ctcagcatct gataagaaag tttgggcaag taggtgcgaa aaatttactc tagataccat	7980
tttccaggaa aaaatataat tcaagtccaa cgtattacta gctgtcacct tccccaacaa	8040
tcaaataagc attgttattg tcaaacacat ttacttgtat agttcagtgt tatatgtaat	8100
ggacttaggt ataattcctg tgctattccc ttttgatgcc aaacaccctt ttgtaagctg	8160
ttttgtttga aatccaaatt tatatgtctt aaagtaaata gcagcattgg gctttcttaa	8220
tgtaatgact cttctgcat cctatttcta gagttaggca gaaaattctg taaatctatc	8280
ttatttccag tctgtttgtg actatgggtg gcaataaatc tgatttcca gaaaaacaca	8340
tctggacccc tataggtgaa gtaaacaagc ctcatcttaa attgtttgta cttccgacct	8400
gacctcaaaa caccatcatc atctagagca gcaacaagta aagaacaatg cactggaaat	8460
cacatatatg ggtctagcac cagcctttac attacacaac tttgctttgg gcaagtcatt	8520
taaatttctg tgcttcagtt tactaaatct atgaaactga ctaggtaaca gctgacatgc	8580
ttttccagca ctccgtcttt tacctgtctg atgtcacaaa ttgatcattc ctttctggcc	8640
ttgagatgcg tatcaacca gatagtcatt aaccagata gttggcattc aggataaaat	8700
gtttttgcct ttctgtcct gcaagctctt taaagtccaa aatatgcctt tatatcttta	8760
tgtgtaatag gtatgaaatc tatttcttgg ttgttgaata ggataggctt attagacatt	8820
ctaaagtcaa cttttttctc atatttaaac atttaccgt agaaattcag ttttttaaag	8880
ttttaagtca gttgattgta ccaaacatta ctattaaaaa aaaatcaaag tcagagacca	8940
gagtttccct gcatcagcgt tcagtcagga gacaggaact atactaattt gaacatgtaa	9000
agttaaatat aatccttaat aataacagga ttggaataag ggataaaaga aagagaattc	9060
cttctatata tatgtataca cacacatata tatatataca tgtgtatata tgttgtgtga	9120
tgtatattat ataataca tacacacaca caaagaatgt cagacaagga gcagccacca	9180
ttctgcttac tttcactgga ttaagagttt cataagggca agggttttgt ttgtcttaca	9240

ctccactggg gattaatctc cagtgtctag aacagtgcct aacacataaa taaagcctca	9300
attattattt gtggaatgga atggaatttg agtgaaatgg aatgggaata gaatggactg	9360
gaatggaaat agaagtggga aggggaatgg aaggggatag aagaggattg aagaggacag	9420
aaggagatgg gataagatgg gatgggatag gatcgaatct caggtccagg agtcagaata	9480
actgaatcct attccagttc tgccatcaac aagtcctgtg aaatttgga agtcacttaa	9540
ccctctggcc ttgtagaata tctcagtggg catgtctgat gtttcttcat ggtagatat	9600
aggttgcatg ttttgggcaa gaatatacaca ggaatgatgc tgtgttctca gtgcacccag	9660
tcaggtaatg catgatttca atttctctca ttactgatga tgttaacttt aatcacttag	9720
tgaatgagaa catgagcatc attttgaaag caaactgcaa agttttgatg tcaacatttt	9780
agcaattacg tataatttcc cccaaggata ctggctttga gataatgtaa atatacatgg	9840
gaaagacaaa gctcaaataa aggaaaaaga ccca	9874

<210> 50
 <211> 1151
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >HRAS|ENSG00000174775|ENST00000451590|1151

<400> 50	
tgccctgcgc ccgcaacccg agccgcaccc gccgcggacg gagcccatgc gcggggcgaa	60
ccgcgcgccc ccgccccgc cccgccccgg cctcggcccc ggccctggcc ccgggggcag	120
tcgcgcctgt gaacggtggg gcaggagacc ctgtaggagg accccgggcc gcaggcccct	180
gaggagcgat gacggaatat aagctggtgg tggtagggcg cggcggtgtg ggcaagagt	240
cgctgacat ccagctgatc cagaaccatt ttgtggacga atacgacccc actatagagg	300
attcctaccg gaagcaggtg gtcattgatg gggagacgtg cctgttgac atcctggata	360
ccgccggcca ggaggagtac agcgccatgc gggaccagta catgcgcacc ggggagggct	420
tcctgtgtgt gtttgccatc aacaacacca agtcttttga ggacatccac cagtacaggg	480
agcagatcaa acgggtgaag gactcggatg acgtgcccat ggtgctggtg gggaacaagt	540
gtgacctggc tgcacgcact gtggaatctc ggcaggctca ggacctcgcc cgaagctacg	600
gcatcccccta catcgagacc tcggccaaga cccggcaggg agtggaggat gccttctaca	660

cgttggtgcg tgagatccgg cagcacaagc tgcggaagct gaaccctcct gatgagagtg	720
gccccggctg catgagctgc aagtgtgtgc tctcctgacg caggtgaggg ggactcccag	780
ggcggccgcc acgcccaccg gatgaccccg gctccccgcc cctgccggtc tcctggcctg	840
cggtcagcag cctcccttgt gccccgcca gcacaagctc aggacatgga ggtgccggat	900
gcaggaagga ggtgcagacg gaaggaggag gaaggaagga cggaagcaag gaaggaagga	960
agggctgctg gagcccagtc accccgggac cgtgggccga ggtgactgca gaccctccca	1020
gggaggctgt gcacagactg tcttgaacat cccaaatgcc accggaaccc cagcccttag	1080
ctccccctccc aggccctctgt gggcccttgt cgggcacaga tgggatcaca gtaaattatt	1140
ggatggtctt g	1151

<210> 51
 <211> 1119
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >KRAS|ENSG00000133703|ENST00000256078|1119

<400> 51	
ctgaaggcgg cggcggggcc agaggctcag cggctcccag gtgcgggaga gaggcctgct	60
gaaaatgact gaatataaac ttgtggtagt tggagctggt ggcgtaggca agagtgcctt	120
gacgatacag ctaattcaga atcattttgt ggacgaatat gatccaacaa tagaggattc	180
ctacaggaag caagtagtaa ttgatggaga aacctgtctc ttggatattc tcgacacagc	240
aggtcaagag gagtacagt caatgaggga ccagtacatg aggactgggg agggctttct	300
tttgttattt gccataaata atactaaatc atttgaagat attcaccatt atagagaaca	360
aattaaaaga gttaaggact ctgaagatgt acctatgggt ctagtaggaa ataaatgtga	420
tttgctttct agaacagtag acacaaaaca ggctcaggac ttagcaagaa gttatggaat	480
tccttttatt gaaacatcag caaagacaag acagagagtg gaggatgctt ttatacatt	540
ggtgaggggag atccgacaat acagattgaa aaaaatcagc aaagaagaaa agactcctgg	600
ctgtgtgaaa attaaaaaat gcattataat gtaatctggg tgttgatgat gccttctata	660
cattagttcg agaaattcga aaacataaag aaaagatgag caaagatggt aaaaagaaga	720

aaaagaagtc aaagacaaag tgtgtaatta tgtaaataca atttgtactt ttttcttaag	780
gcatactagt acaagtggta atttttgtac attacactaa attattagca tttgttttag	840
cattacctaa tttttttcct gctccatgca gactgttagc ttttacctta aatgcttatt	900
ttaaaatgac agtgggaagtt tttttttcct ctaagtgcc a gtattcccag agttttggtt	960
tttgaactag caatgcctgt gaaaaagaaa ctgaatacct aagatttctg tcttggggct	1020
tttggtgcat gcagttgatt acttcttatt tttcttacca attgtgaatg ttggtgtgaa	1080
acaaattaat gaagcttttg aatcatccct attctgtgt	1119

<210> 52
 <211> 4103
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >LIG4|ENSG00000174405|ENST00000356922|4103

<400> 52	
ccacagcgct gtagactgcg ccgcattaga agcctggcct cctgatgctg tgctcttcat	60
ctagacccaa gccccaggtc gtgggacgat ttctcccgtt tttgactccc tggaactgta	120
ttgcctgctt tacctgcgta catgttgatt ctttctcatg gcaaccccg c aggaaaccat	180
caagatctca ttttacagct gggattctct ggttcacaga ggtaacggag cttgcccag	240
gccagttaaa cgagaagatt catcaccgct ttgatggctg cctcaciaaac ttcaciaaact	300
gttgcatctc acgttccttt tgcagatttg tgttcaactt tagaacgaat acagaaaagt	360
aaaggacgtg cagaaaaaat cagacacttc agggaatttt tagattcttg gagaaaattt	420
catgatgctc ttcataagaa ccacaaagat gtcacagact ctttttatcc agcaatgaga	480
ctaattcttc ctcagctaga aagagagaga atggcctatg gaattaaaga aactatgctt	540
gctaagcttt atattgagtt gcttaattta cctagagatg gaaaagatgc cttcaaactt	600
ttaaactaca gaacaccac tggaactcat ggagatgctg gagactttgc aatgattgca	660
tattttgtgt tgaagccaag atgtttacag aaaggaagtt taaccataca gcaagtaaac	720
gaccttttag actcaattgc cagcaataat tctgctaaaa gaaaagacct aataaaaaag	780
agccttcttc aacttataac tcagagttca gcacttgagc aaaagtggct tatacggatg	840
atcataaagg atttaaagct tgggtgttagt cagcaaacta tcttttctgt ttttcataat	900

gatgctgctg agttgcataa tgtcactaca gatctggaaa aagtctgtag gcaactgcat	960
gaccccttctg taggactcag tgatatttct atcactttat tttctgcatt taaaccaatg	1020
ctagctgcta ttgcagatat tgagcacatt gagaaggata tgaaacatca gagtttctac	1080
atagaaacca agctagatgg tgaacgtatg caaatgcaca aagatggaga tgtatataaa	1140
tactttcttc gaaatggata taactacact gatcagtttg gtgcttctcc tactgaaggt	1200
tctcttacc ctttcattca taatgcattc aaagcagata tacaaatctg tattcttgat	1260
ggtagatga tggcctataa tcctaataca caaactttca tgcaaaaggg aactaagttt	1320
gatattaaaa gaatggtaga ggattctgat ctgcaaactt gttattgtgt ttttgatgta	1380
ttgatggtta ataataaaaa gctagggcat gagactctga gaaagaggta tgagattctt	1440
agtagtattt ttacaccaat tccaggtaga atagaaatag tgcagaaaac acaagctcat	1500
actaagaatg aagtaattga tgcattgaat gaagcaatag ataaaagaga agagggaatt	1560
atggtaaaac aacctctatc catctacaag ccagacaaaa gaggtgaagg gtggttaaaa	1620
attaaaccag agtatgtcag tggactaatg gatgaattgg acattttaat tgttggagga	1680
tattggggta aaggatcacg ggggtggaatg atgtctcatt ttctgtgtgc agtagcagag	1740
aagccccctc ctgggtagaa gccatctgtg ttccatactc tctctcgtgt tgggtctggc	1800
tgcacatga aagaactgta tgatctgggt ttgaaattgg ccaagtattg gaagcctttt	1860
catagaaaag ctccaccaag cagcatttta tgtggaacag agaagccaga agtatacatt	1920
gaaccttgta attctgtcat tgttcagatt aaagcagcag agatcgtacc cagtgatatg	1980
tataaaactg gctgcacctt gcgttttcca cgaattgaaa agataagaga tgacaaggag	2040
tggcatgagt gcatgaccct ggacgaccta gaacaactta gggggaaggc atctggtaag	2100
ctcgcatcta aacaccttta tataggtggg gatgatgaac cacaagaaaa aaagcggaaa	2160
gctgccccaa agatgaagaa agttattgga attattgagc acttaaaagc acctaacctt	2220
actaacgta acaaaatttc taatatattt gaagatgtag agttttgtgt tatgagtgga	2280
acagatagcc agccaaagcc tgacctggag aacagaattg cagaatttgg tggttatata	2340
gtacaaaatc caggcccaga cacgtactgt gtaattgcag ggtctgagaa catcagagtg	2400
aaaaacataa ttttgtcaaa taaacatgat gttgtcaagc ctgcatggct tttagaatgt	2460
tttaagacca aaagctttgt accatggcag cctcgcttta tgattcatat gtgcccatca	2520

accaaagaac attttgcccg tgaatatgat tgctatggtg atagttatth cattgataca	2580
gacttgaacc aactgaagga agtattctca ggaattaaaa attctaacga gcagactcct	2640
gaagaaatgg cttctctgat tgctgattta gaatatcggg attcctggga ttgctctcct	2700
ctcagtatgt ttcgacgcca caccgtttat ttggactcgt atgctgttat taatgacctg	2760
agtacaaaa atgaggggac aaggtttagct attaaagcct tggagcttcg gtttcatgga	2820
gcaaaagtag tttcttgtht agctgaggga gtgtctcatg taataattgg ggaagatcat	2880
agtcgtgttg cagattthta agctthttaga agaactthta agagaaagtt taaaatccta	2940
aaagaaagtt gggtaactga ttcaatagac aagtgtgaat tacaagaaga aaaccagtat	3000
ttgattthaa gctaggthtc ctagtgagga aagcctctga tctggcagac tcattgcagc	3060
aggtggtaat gataaaatac taaactacat thtaththtg tatctthaaa atctatgcct	3120
aaaaagtatc attacatata ggaaaacaat aaththtaact thtaaggthg aaaagacaat	3180
agcccaaagc caagaaaga aaattatctt gaatgtagta ttcaatgatt thttatgatc	3240
aaggtgaaat aaacagtcta aagaagaggt gthththataa tatccatata gaaatctaga	3300
aththtactt agatactaath aaatacath tagaaactth taaagtcath aaaaagcath	3360
aacctthctaa acagtatath ctaaaaagtc aaacgthta caatagthth tatctataa	3420
aagcactgca agaaaatagg gtagaattgt tacagctgga cttgtaaaaa tatgtcttht	3480
tactcagggt thaaaatgtc ccaththaaat atgaaatgta aacaaathtg thththtaagg	3540
thtaaggcaa atgtaacaath aaaccctgt cgatggttht agctaaatta gaggaagthg	3600
tatgagactt aatgatctaa aaactthaaa ttgaattggt ttgattaaaa ataaagcttg	3660
caaththaaa agtagctcac athtaathtc ttgtgtgaaa tagaacatgc ththaaaggaa	3720
gtathththt gtgaaththg attccagtat aaatagtath cacaaaaag aththcttag	3780
aththtatcta ttgaatagg gtcaatatgg catgcatath gtaactthca ttagaaataa	3840
gttgctthga cththaaaa tgacatagth agattathth aagtcaatgt atatagtata	3900
tattatgtat ggaththtat accaaathth ggaatacagc ctatctcatg accatattga	3960
aatgtacgga atthgatcca tgcgatacta tgtgtgcath atthgaaagt tattggaaat	4020
thtattcaaa ccgtggaaca aatgtatgtg aththgttat actththaat thaaataaaa	4080
taththaatgc actaththaa taa	4103

<210> 53
 <211> 7576
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >MDC1|ENSG00000137337|ENST00000376406|7576

<400> 53
 ttggatttaa gaaaatgacc tcaaaatgtc cgtcagagac gtattcccag gaagaaagat 60
 attacttcta ctacaaacca aatcaaaagg aaatgaaatt ccaatgcaac aggagtgaac 120
 tgccacgcct acgggctgtt ctccaaactg cagcctccag ccacgactgc aacgcgcaac 180
 ccactttcat ttctcatgag tcagcggaca ccatgtctag gaggaccgag gaaaggcgct 240
 ctggccttac cagacacgtc ggacgtctat gacacagccc ctctatccgt tgccggcagc 300
 tggcgccaga ctctctggtc gcggtttgga actgcgcggg aagtgggtgg tgggcgggca 360
 agcggtagtg ggttgtccct tggagctgcc caatcgacgt gcattattct gttggcgcac 420
 ggcggccttc aattaccgtc tcattaactg atctcagcag cctgggagac accacctatt 480
 tgaactctaa gggggcgggg ctttgggtgt gcctccgctc gactggctgc ggttgtgaaa 540
 gacagcggca gaagccaatc agcaaataag ctctttttcg gcacacgcag tcgctccacc 600
 tgggtcgcga ccgttactgg tggcgcgcg cgggacttaa agtagatcat ggaggacacc 660
 caggctattg actgggatgt tgaagaagag gaggagacag agcaatccag tgaatccttg 720
 aggtgtaacg tggagccagt agggcggcta catatcttta gtggtgcca tggaccagaa 780
 aaagatttcc cactacacct cggaagaat gtggtaggcc gaatgcctga ctgctctgtg 840
 gccctgccct ttccatctat ctccaaaca catgcagaga ttgaaatctt agcctgggac 900
 aaggcaccta tcctccgaga ctgtgggagc cttaatggta ctcaaatcct gagacctcct 960
 aaggttttga gccctggggg gagtcaccgt ctgagggacc aggaattgat tctctttgct 1020
 gacttgctct gccagtacca tcgcctggat gtctctctgc ctttgtctc ccggggccct 1080
 ctgacagtag aagagacacc cagagtacag ggagaaactc aacccagag gcttctgttg 1140
 gctgaggact cggaggagga agtagatttt ctttctgaaa ggcgtatggt aaaaaaatca 1200
 aggaccacat cttcctctgt gatagttcca gagagtgatg aagaggggca ttccccggtc 1260

ctgggcggcc ttgggcccgc ttttgccttc aatttgaaca gtgacacaga tgtggaagaa	1320
ggtcagcaac cagccacaga ggaggcctcc tcagctgcca gaagaggtgc cactgtagag	1380
gcaaagcagt ctgaagctga agttgtaact gaaatccagc ttgaaaagga tcagccttta	1440
gtgaaggaga gggacaatga tacaaaagtc aagaggggtg cagggaatgg ggtggttcca	1500
gctgggggtga ttctggagag gagccaacct cctggagagg acagtgacac agatgtggat	1560
gatgacagca ggcctcctgg aaggccagct gaggtccatt tggaaagggc tcagcctttt	1620
ggcttcatcg acagcgacac tgatgcggaa gaagagagga tcccagcaac cccagttgtc	1680
attcctatga agaagaggaa gatcttccat ggagtaggta caaggggtcc tggagcacca	1740
ggcctggccc atctgcagga gagccaggct ggtagtgata cagatgtgga agaaggcaag	1800
gccccacagg ctgtccctct ggagaaaagc caagcttcca tggttatcaa cagcgataca	1860
gatgacgagg aagaagtctc agcagcgctg actttggcac atctgaaaga gagccagcct	1920
gctatatgga acagagatgc agaagaggac atgccccaac gtgtggtcct tctgcagcga	1980
agccaaacca ccactgagag agacagtgc acagacgtgg aggaggaaga gctcccagtg	2040
gaaaatagag aagctgtcct caaggatcac acaaagatta gagcccttgt tagagcacat	2100
tcagaaaagg accaacctcc ttttggggac agtgatgaca gtgtggaagc agataagagc	2160
tcacctggga tccacctgga gagaagccaa gcctccacca cagtggacat caacacacaa	2220
gtggagaagg aagtcccgc agggtcagcc attatacata taaagaagca tcaggtgtct	2280
gtggagggga caaatcaaac agatgtgaaa gcagttgggg gaccagcaaa gctgcttgtg	2340
gtatctctag aggaagcctg gcctctgcat ggggactgtg aaacagatgc agaggagggc	2400
acctccccta cagcctcagt agttgcagat gtaagaaaga gccagcttcc agcagaaggg	2460
gatgctgggg cagagtgggc tgcagctgtt cttaaagcagg agagagctca tgaggtgggg	2520
gcccagggtg ggccacctgt ggcacaagtg gagcaggacc tccctatctc aagagagaac	2580
ctcacagatc tgggtggtga cacagacact ctaggggaat ccaccagcc acagagagag	2640
ggagcccagg tccccacagg aaggagaga gaacaacatg tgggtgggac caaggactct	2700
gaagacaact atggtgattc tgaagatctg gacctacaag ctaccagtg ctttctggag	2760
aatcagggcc tggaagcagt ccagagcatg gaggatgaac ctaccaggc cttcatgttg	2820
actccacccc aagagcttgg cccttcccat tgcagcttcc agacaacagg taccctagat	2880

gaaccatggg aggtcctggc tacacagcca ttctgtctga gagagtctga ggactctgag	2940
acccagcctt ttgacacgca ccttgaggcc tatggacctt gcctgtctcc acctagggca	3000
ataccaggag accaacaatcc agagagccca gttcacacag agccaatggg gattcaaggc	3060
agagggaggc agactgtgga taaagtcatg ggtataccaa aagaaacagc agagagggtg	3120
ggccttgaga gagggccatt ggagagagaa actgagaaac tgctaccaga aagacagaca	3180
gatgtgacag gagaggaaga attaaccaag gggaaacagg acagagaaca aaaacagttg	3240
ttagctagag acaccagag acaagaatct gacaaaaatg gggaaagtgc aagtcctgaa	3300
agagataggg agagtttgaa ggtagaaatt gagacatctg aggaaatata agagaaacaa	3360
gtacagaagc agacccttcc aagcaaagca tttgagagag aagtagagag accagtagca	3420
aacagagagt gcgatccagc cgagttagaa gagaagggtgc ccaaagtgat cctggagaga	3480
gatacacaga gaggggagcc agagggaggg agccaggacc agaaagggca ggcctccagc	3540
ccaacaccag agcctggggg gggggcgggg gaccttccgg gacctacctc agccccgta	3600
ccttctggga gccagtcagg tggaagggga tccccagtga gccccaggag gcatcagaaa	3660
ggcctcctga attgcaagat gccacctgct gagaaggctt ccaggatcag agctgctgag	3720
aaggtttcca ggggcgatca ggaatctcca gatgcttgtc tgccctctac agtacctgaa	3780
gccccagccc cccccaaaa gcccttaac tctcagagcc agaaacatct tgcacctccg	3840
ccccttcttt ctcccctttt accttctatc aagccaaccg ttcgtaagac caggcaagat	3900
gggagtcagg aagctccaga ggctcccttg tcctcagagc tggagccttt ccacccaaag	3960
cctaaaatta gaactcggaa gtcctccaga atgacaccct ttccagctac ctctgctgcc	4020
cctgagcccc acccttccac ctccacagcc cagccagtca ctcccaagcc cacatctcag	4080
gccactagga gcaggacaaa taggtcctct gtcaagacct ctgaaccagt tgtccccaca	4140
gcccctgagc tccagccttc cacctccaca gaccagcctg tcacctctga gccacatct	4200
caggttacta ggggaagaaa aagtagatcc tctgtcaaga cccctgaaac agttgtgccc	4260
acagcccttg agctccagcc ttccacctcc accgaccgac ctgtcacctc tgaaccaccc	4320
tctcaggcta ctaggggaag aaaaaataga tcctctgtca agaccctga accagttgtc	4380
cccacagccc ctgagctcca gccttccacc tccacagacc agcctgtcac ttctgagccc	4440
acatatcagg ctactagggg aagaaaaaat agatcctctg tcaagacccc tgaaccagtt	4500

gtgcccacag cccctgagct ccggccttcc acctccacag accgacctgt ccccccaag	4560
cccacatctc ggaccactag gagcaggaca aatatgtcct ctgtcaagac ccctgaaaca	4620
gttgtcccca cagcccctga gctccagatt tccacctcca cagaccaacc tgtcaccctt	4680
aagcccacat ctcgaccac taggagcagg acaaatatgt cctctgtgaa gaaccctgaa	4740
tcaactgtcc ctatagcccc tgagctccca ctttccacct ccacagagca gcctgtcacc	4800
cctgagccca catctcgggc tactagggga agaaaaata gacacctgtg caagaccctt	4860
gaaacacttg tccccacagc ccctaagctc gagccttcca cttccacaga ccaacctgtc	4920
actcctgagc ccacatctca ggccaccagg ggcaggacaa ataggtcctc tgtgaagacc	4980
cctgaaacag ttgtccccac agcccctgag ctccagcctt ccacctccac agaccagcct	5040
gttacccttg agcctacgtc tcaggctact aggggaagaa cagatagatc ctctgtcaag	5100
actcctgaaa cagttgtccc cacagcccct gagctacagg cttccgcctc cacagaccag	5160
cctgtcacct ctgagccac atctcggacc actaggggaa gaaaaaatcg gtcctctgtc	5220
aagacccttg aaacagttgt gcccgcagcc cctgagctcc agccttccac ctccacagac	5280
caacctgtca cccctgagcc cacatctcgg gccactaggg gcaggacaaa taggtcctct	5340
gtcaagaccc ctgaatcaat tgtccctata gcccctgagc ttcagccttc cacctccaga	5400
aaccagcttg tcaccctga gccacatct cggggcacta ggtgcaggac aaataggtcc	5460
tctgtcaaga cccctgagcc agttgtcccc acagcccctg agcccatcc taccacctcc	5520
acagaccagc ctgtcacccc caagctcaca tctagggcca ctaggagaaa gacaaatagg	5580
tcctctgtca agactcccaa accagttgaa ccagcagcct ctgatcttga gccttttacc	5640
cccacagacc agtccgtcac ccctgaggcc atagctcagg gtggtcagag caaaacactg	5700
aggtcttcca cagtaagagc tatgccggtt cctaccaccc ctgaattcca atctcctgtc	5760
accacagacc agcctatttc ccctgagcct attactcaac ccagttgcat caagaggcag	5820
agagccgctg ggaaccctgg ctccctcgca gctccattg accataagcc ttgctctgca	5880
cccttggaac ctaaattcca ggcctcaagg aaccaaagat ggggagcagt gagagcagct	5940
gaatccctta cagccattcc tgagcctgcc tctccccagc ttcttgagac accaattcat	6000
gcctcccaga tccaaaaggt ggaaccagca ggtagatcta gggtcacccc ggagctccag	6060
cctaaggcct ctcaaagccg caagaggtct ttagctacca tggattcacc accacatcaa	6120

aaacagcccc aaagagggga agtctcccag aagacagtga ttatcaagga agaggaagaa	6180
gatactgcag agaagccagg gaaggaagag gatgtcgtga ctccaaaacc aggcaagaga	6240
aagagagacc aggagagga ggagcccaac agaataccaa gccgcagcct ccgacggacc	6300
aaacttaacc aagaatcaac agccccaaa gtgctcttca caggagtggg ggatgctcgg	6360
ggagagcggg ctgtgctggc actgggggga agtctggctg gttcagcggc agaggcttcc	6420
cacctgggtca ctgatcgcat ccgccggaca gtcaagttcc tgtgtgccct ggggcgggga	6480
atccccattc tgtccctgga ctggctgcat cagtcccgca aggctgggtt cttcttacc	6540
ccggatgaat atgtggtgac cgaccctgag caagagaaga actttggctt tagccttcaa	6600
gacgcactga gcagggtcgc ggagcgaagg ctgctagagg gctatgagat ctatgtgacc	6660
cctggagtcc agccaccacc acctcagatg ggagagatta ttagctgctg tggaggcaca	6720
tacctacca gcatgcctcg gtcctataag cctcagagag ttgtgatcac atgccctcag	6780
gacttccctc attgctccat tccactacgg gttgggctgc cctcctctc gcctgagttc	6840
ctgctgactg gagtgtgaa gcaggaagcc aagccagagg ctttgtcct ctccccttg	6900
gagatgtcat ccacctgaga actccactac cttttccct cccagaccac gaattagaag	6960
atatgtggaa gaaagaactc agggcgtag aaaggattgg ggtatattga tacaacttgt	7020
cctggaacat ggggtgggacc agaaatcttt atgaataaat gaaaagataa gggatttgga	7080
agccacaggt tgttttttgt ttgtttgttt gtttttttaa tggccatttt attttatttg	7140
tatttatagt tttttatttg tatagattta ggggatacaa gatttcttac atgcatgtat	7200
taaatggcca ttttaaaatt agctagtttc atgctcagat gtcataagtg gcagctatct	7260
ttagccagac tgttgcagtt attgctcgat gccactcatg gtgtcctacc tcctatttgg	7320
aaaccatctc tatttttttc ttactgagat tcttactttg gggtcaggaa cttgaaggga	7380
tgcttgagat gagtagattt gaggggtccag ttatggagtg ctactaaaac attttcttct	7440
ctcctggcct ctggaagcat ctttagcttt gactttgggc aagtctctgt acttttctgg	7500
ccagcttttc caggatttat aaaattagag cttcggcttg acctctgtga taaataaata	7560
ttcactctgt gcctta	7576

<210> 54

<211> 2752

<212> DNA

<213> Homo sapiens

<220>

<223> >MLH1|ENSG00000076242|ENST00000231790|2752

<400> 54

gccgcttcag ggagggacga agagaccag caaccacag agttgagaaa tttgactggc	60
attcaagctg tccaatcaat agctgccgct gaagggtggg gctggatggc gtaagctaca	120
gctgaaggaa gaacgtgagc acgaggcact gaggtgattg gctgaaggca cttccgttga	180
gcatctagac gtttccttgg ctcttctggc gccaaaatgt cgttcgtggc aggggttatt	240
cggcggctgg acgagacagt ggtgaaccgc atcgcggcgg gggaagttat ccagcggcca	300
gctaattgcta tcaaagagat gattgagaac tgtttagatg caaaatccac aagtattcaa	360
gtgattgtta aagagggagg cctgaagttg attcagatcc aagacaatgg caccgggatc	420
aggaaagaag atctggatat tgtatgtgaa aggttcacta ctagtaaact gcagtccttt	480
gaggatttag ccagtatttc tacctatggc tttcgaggtg aggctttggc cagcataagc	540
catgtggctc atgttactat tacaacgaaa acagctgatg gaaagtgtgc atacagagca	600
agttactcag atggaaaact gaaagcccct cctaaacat gtgctggcaa tcaagggacc	660
cagatcacgg tggaggacct tttttacaac atagccacga ggagaaaagc tttaaaaaat	720
ccaagtgaag aatatgggaa aatttttgaa gttgttgga ggtattcagt acacaatgca	780
ggcattagtt tctcagttaa aaaacaagga gagacagtag ctgatgttag gacactaccc	840
aatgcctcaa ccgtggacaa tattcgctcc atctttggaa atgctgttag tcgagaactg	900
atagaaattg gatgtgagga taaaacccta gccttcaaaa tgaatggtta catatccaat	960
gcaaactact cagtgaagaa gtgcatcttc ttactcttca tcaaccatcg tctggtagaa	1020
tcaacttcct tgagaaaagc catagaaaca gtgtatgcag cctatttgcc caaaaacaca	1080
caccatttcc tgtacctcag tttagaaatc agtccccaga atgtggatgt taatgtgcac	1140
cccacaaagc atgaagttca cttcctgcac gaggagagca tcctggagcg ggtgcagcag	1200
cacatcgaga gcaagctcct gggctccaat tcctccagga tgtacttcac ccagactttg	1260
ctaccaggac ttgctggccc ctctggggag atgggttaaat ccacaacaag tctgacctcg	1320
tcttctactt ctggaagtag tgataaggtc tatgcccacc agatggttcg tacagattcc	1380
cgggaacaga agcttgatgc atttctgcag cctctgagca aaccctgtc cagtcagccc	1440

caggccattg tcacagagga taagacagat atttctagtgc gcagggctag gcagcaagat	1500
gaggagatgc ttgaactccc agcccctgct gaagtggctg ccaaaaatca gagcttggag	1560
ggggatacaa caaaggggac ttcagaaatg tcagagaaga gaggacctac ttccagcaac	1620
cccagaaaga gacatcggga agattctgat gtggaaatgg tggaagatga ttcccgaag	1680
gaaatgactg cagcttgtac cccccggaga aggatcatta acctcactag tgttttgagt	1740
ctccaggaag aaattaatga gcagggacat gaggttctcc gggagatggt gcataaccac	1800
tccttcgtgg gctgtgtgaa tcctcagtgg gccttggcac agcatcaaac caagttatac	1860
cttctcaaca ccaccaagct tagtgaagaa ctgttctacc agatactcat ttatgatttt	1920
gccaattttg gtgttctcag gttatcggag ccagcaccgc tctttgacct tgccatgctt	1980
gccttagata gtccagagag tggctggaca gaggaagatg gtcccaaaga aggacttgct	2040
gaatacattg ttgagtttct gaagaagaag gctgagatgc ttgcagacta tttctctttg	2100
gaaattgatg aggaagggaa cctgattgga ttacccttc tgattgacaa ctatgtgccc	2160
cctttggagg gactgcctat cttcattctt cgactagcca ctgaggtgaa ttgggacgaa	2220
gaaaaggaat gttttgaaag cctcagtaaa gaatgcgcta tgttctattc catccggaag	2280
cagtacatat ctgaggagtc gaccctctca ggccagcaga gtgaagtgcc tggctccatt	2340
ccaaactcct ggaagtggac tgtggaacac attgtctata aagccttgcg ctcacacatt	2400
ctgcctccta aacatttcac agaagatgga aatatactgc agcttgctaa cctgcctgat	2460
ctatacaaag tctttgagag gtgttaaata tggttattta tgcactgtgg gatgtgttct	2520
tctttctctg tattccgata caaagtgttg tatcaaagtg tgatatacaa agtgtaccaa	2580
cataagtgtt ggtagcactt aagacttata cttgccttct gatagtattc ctttatacac	2640
agtggattga ttataaataa atagatgtgt cttaacataa tttcttattt aattttatta	2700
tgatatatatt gtgtcagttc agatgccaaa aagaggtctt gaacatgtca ca	2752

<210> 55

<211> 7896

<212> DNA

<213> Homo sapiens

<220>

<223> >MLH3|ENSG00000119684|ENST00000355774|7896

<400> 55

aacaactggt gcgcatgcgc actggtgtct cgcggcctgg cgcgccccct ccgaagcgca	60
tgctcgtggg cacgcacgag cctcaagatc caaggtgcgc gcgtcggcgt ccgaggcggt	120
tggtgtcggg gaatttgta agcgggactc caggcaatta tttccagtca gagaaggaaa	180
ccagtgcctg gcattctcac catctttcta cctaccatga tcaagtgctt gtcagttgaa	240
gtacaagcca aattgcgttc tggtttggcc ataagctcct tgggccaatg tgttgaggaa	300
cttgccctca acagtattga tgctgaagca aaatgtgtgg ctgtcagggt gaatatggaa	360
accttccaag ttcaagtgat agacaatgga tttgggatgg ggagtgatga tgtagagaaa	420
gtgggaaatc gttatttcac cagtaaagtc cactcggtag aggacttgga gaatccaagg	480
ttttatgggt tccgaggaga ggccttggca aatattgctg acatggccag tgctgtggaa	540
atttcgtcca agaaaaacag gacaatgaaa acttttgtga aactgtttca gagtggaaaa	600
gccctgaaag cttgtgaagc tgatgtgact agagcaagcg ctgggactac tgtaacagtg	660
tataacctat tttaccagct tcctgtaagg aggaaatgca tggaccctag actggagttt	720
gagaaggtta ggcagagaat agaagctctc tcactcatgc acccttccat ttctttctct	780
ttgagaaatg atgtttctgg ttccatgggt cttcagctcc ctaaaaccaa agacgtatgt	840
tcccgatttt gtcaaattta tggattggga aagtcccaaa agctaagaga aataagtttt	900
aaatataaag agtttgagct tagtggctat atcagctctg aagcacatta caacaagaat	960
atgcagtttt tgtttgtgaa caaaagacta gttttaagga caaagctaca taaactcatt	1020
gactttttat taaggaaaga aagtattata tgcaagccaa agaatgggtcc caccagtagg	1080
caaatgaatt caagtcttcg gcaccggtct accccagaac tctatggcat atatgttaatt	1140
aatgtgcagt gccaatcttg tgagtatgat gtgtgcatgg agccagccaa aactctgatt	1200
gaatttcaga actgggacac tctcttgttt tgcatcagg aaggagtga aatgttttta	1260
aagcaagaaa aattatttgt ggaattatca ggtgaggata ttaaggaatt tagtgaagat	1320
aatggtttta gtttatttga tgctactctt cagaagcgtg tgacttccga tgagaggagc	1380
aatttccagg aagcatgtaa taatatttta gattcctatg agatgtttta tttgcagtca	1440
aaagctgtga aaagaaaaac tactgcagaa aacgtaaaca cacagagttc tagggattca	1500
gaagctacca gaaaaaatac aaatgatgca tttttgtaca tttatgaatc aggtggtcca	1560

ggccatagca	aaatgacaga	gccatcttta	caaaacaaag	acagctcttg	ctcagaatca	1620
aagatgttag	aacaagagac	aattgtagca	tcagaagctg	gagaaaatga	gaaacataaa	1680
aaatctttcc	tggaacatag	ctcttttagaa	aatccgtgtg	gaaccagttt	agaaatgttt	1740
ttaagccctt	ttcagacacc	atgtcacttt	gaggagagtg	ggcaggatct	agaaatatgg	1800
aaagaaagta	ctactgttaa	tggcatggct	gccaacatct	tgaaaaataa	tagaattcag	1860
aatcaaccaa	agagatttaa	agatgctact	gaagtgggat	gccagcctct	gccttttgca	1920
acaacattat	ggggagtaca	tagtgctcag	acagagaaag	agaaaaaaaa	agaatctagc	1980
aattgtggaa	gaagaaatgt	ttttagttat	gggcgagtta	aattatgttc	cactggcttt	2040
ataactcatg	tagtacaaaa	tgaaaaaact	aatcaactg	aaacagaaca	ttcattttaa	2100
aattatgtta	gacctggtcc	cacacgtgcc	caagaaacat	ttggaaatag	aacacgtcat	2160
tcagttgaaa	ctccagacat	caaagattta	gccagcactt	taagtaaaga	atctgggtcaa	2220
ttgccaaca	aaaaaaattg	cagaacgaat	ataagttatg	ggctagagaa	tgaacctaca	2280
gcaacttata	caatgttttc	tgcttttcag	gaaggtagca	aaaaatcaca	aacagattgc	2340
atattatctg	atacatcccc	ctctttcccc	tggtatagac	acgtttccaa	tgatagtagg	2400
aaaacagata	aattaattgg	tttctccaaa	ccaatcgtcc	gtaagaagct	aagcttgagt	2460
tcacagctag	gatctttaga	gaagtttaag	aggcaatatg	ggaaggttga	aaatcctctg	2520
gatacagaag	tagaggaaag	taatggagtc	actaccaatc	tcagtcttca	agttgaacct	2580
gacattctgc	tgaaggacaa	gaaccgctta	gagaactctg	atgtttgtaa	aatcactact	2640
atggagcata	gtgattcaga	tagtagttgt	caaccagcaa	gccacatcct	taactcagag	2700
aagtttccat	tctccaagga	tgaagattgt	ttagaacaac	agatgcctag	tttgagagaa	2760
agtcctatga	ccctgaagga	gttatctctc	tttaatagaa	aacctttgga	ccttgagaag	2820
tcatctgaat	cactagcctc	taaattatcc	agactgaagg	gttccgaaag	agaaactcaa	2880
acaatgggga	tgatgagtcg	ttttaatgaa	cttccaaatt	cagattccag	taggaaagac	2940
agcaagttgt	gcagtgtgtt	aacacaagat	ttttgtatgt	tatttaacaa	caagcatgaa	3000
aaaacagaga	atgggtgtcat	ccaacatca	gattctgcca	cacaggataa	ttcctttaat	3060
aaaaatagta	aaacacattc	taacagcaat	acaacagaga	actgtgtgat	atcagaaact	3120
cctttggtat	tgccctataa	taattctaaa	gttaccggta	aagattcaga	tgttcttatc	3180

agagcctcag aacaacagat aggaagtctt gactctccca gtggaatggt aatgaatccg	3240
gtagaagatg ccacaggtga ccaaaatgga atttgttttc agagtgagga atctaaagca	3300
agagcttggt ctgaaactga agagtcaaac acgtgttggt cagattggca gcggcatttc	3360
gatgtagccc tgggaagaat ggtttatgtc aacaaaatga ctggactcag cacattcatt	3420
gccccaaactg aggacattca ggctgcttgt actaaagacc tgacaactgt ggctgtggat	3480
gttgtacttg agaatgggtc tcagtacagg tgtcaacctt ttagaagcga cttgtttctt	3540
cctttccttc cgagagctcg agcagagagg actgtgatga gacaggataa cagagatact	3600
gtggatgata ctgttagtag cgaatcgctt cagtctttgt tctcagaatg ggacaatcca	3660
gtatttgccc gttatccaga ggttgctgtt gatgtaagca gtggccaggc tgagagctta	3720
gcagttaaaa ttcacaacat cttgtatccc tatcgtttca ccaaaggaat gattcattca	3780
atgcagggtc tccagcaagt agataacaag tttattgcct gtttgatgag cactaagact	3840
gaagagaatg gcgaggcagg tgggaacctg ctcgtgctgg tggatcagca cgctgcccac	3900
gagcgtatac gtctggagca gcttatcatt gattcctacg agaagcaaca ggcacaaggc	3960
tctggtcgga aaaaattact gtcttctact ctaattcctc cgctagagat aacagtgaca	4020
gaggaacaaa ggagactctt atggtgttac cacaaaaatc tggaagatct gggccttgaa	4080
tttgtatttc cagacactag tgattctctg gtccttgtgg gaaaagtacc actatgtttt	4140
gtggaaagag aagccaatga acttcggaga ggaagatcta ctgtgaccaa gagtattgtg	4200
gaggaattta tccgagaaca actggagcta ctccagacca ccggaggcat ccaagggaca	4260
ttgccactga ctgtccagaa ggtgttgga tcccaagcct gccatggggc cattaagttt	4320
aatgatggcc tgagcttaca ggaaagttgc cgccttattg aagctctgtc ctcatgccag	4380
ctgccattcc agtgtgctca cgggagacct tctatgctgc cgttagctga catagaccac	4440
ttggaacagg aaaaacagat taaaccaaac ctactaaac ttcgcaaat ggcccaggcc	4500
tggcgtctct ttggaaaagc agagtgtgat acaaggcaga gcctgcagca atccatgcct	4560
ccctgtgagc caccatgaga acagaatcac tgggtctaaaa ggaacaaagg gatgttcact	4620
gtatgcctct gagcagagag cagcagcagc aggtaccagc acggccctga ctgaatcagc	4680
ccagtgtccc tgagcagctt agacagcagg gctctctgta tcagtctttc ttgagcagat	4740
gattccccta gttgagtagc cagatgaaat tcaagcctaa agacaattca ttcatttgca	4800

tccatgggca cagaaggttg ctatatagta tctacctttt gctacttatt taatgataaa	4860
atttaatgac agtttgattg gttgcttggg ttgttatttg aagggtgtga tttttgtttt	4920
tgtacagttt tttttcaagc ttcacatttg cgtgtatcta attcagctga tgctcaagtc	4980
caaggggtag tctgccttcc caggctgccc ccagggtttc tgcactggtc ccctcttttc	5040
ccttcagtct tcttcacttc cctatgctgc tgcttcatgt gctacatctc agacttaaag	5100
agttttctcta ctacagtga aacattctct aggggtctttc atcaggcctt tagttatttt	5160
agggataaaa actattgata aaaaggacaa ggatagaaca gagaaaattt aaagtcctgt	5220
tccgggtttt ttgttatgtt ttcttttaaaa actcagagac tgatgttcaa tatcccaaac	5280
cagtaaaatg gtgaaaatac tatgagcttg ttttttaaaa tatgattttt tttggtactt	5340
tataaagtat ctctttatgt gaaagcaatt gtcatatcaa aacacagcat acatacgttc	5400
aacctaacca aatatcttta cactttttct ttcaggagac aagggttctt tgggtccctt	5460
tcaaacggta tcttgggtgtt attacattat gcctatctat tgcccttata atatcacttg	5520
ggaccaggac tgatcgttct gcaaagctt gttatgccat tctcaatcta tttttccgc	5580
accttttcac atgatttggt gttaatagga ctcaacagac taaaattgca tagtagaaaa	5640
aaaatgcaaa aagccagctg gtaatgttta ttgcaactgg ggtgctatac aattagtaag	5700
atgatgcaat gagaatttct acttttgtat ttcctgacca gcctgctcaa agtggctttt	5760
atatcaattg aatgattttc ctcatTTTTT aatacaggaa accaattcgt gctcatggaa	5820
gaaaagtcc tttgccagca gccttgaagt gaatcttaca ggagcaatga aagtattgca	5880
ttcattagcg tctgccccag agaaggttca gagaaaacct tcacttgttt tcaaggggat	5940
ccttgtagat ttacgtaatt ggaatcctga agaacaggcc ctactgtcta aaaaatggct	6000
tttattcttc taaatacata taaacggatg ttttatagat gggaagacat gaccttagaa	6060
aggagagagt tttcagagga tttgccaggc tgtcaggggc tctgcctcca ggcccagtgt	6120
ggcagtgtgg cctcagggcc tccgcctccc tgcttgaggg ctgcatggag gccaaactgtc	6180
ctggggagttg taaaaatctt ttaaggccag accaatttga gggattttta aaagtgtctc	6240
agtgcctctt atgatttcag aaggttttgc tatatgtaat cccaactact gttttcttga	6300
gagtagcaga ggattagaaa aagtcctcca taaattatgt aaccggcctt cctgactagc	6360
ctgactcaag caatgtaaga gataattatt ctgttttcat aatttataag tgtgggggca	6420

tgccctcagca taaaaacaac ctattaggga aaaatatcta atagattacc tttatcgctt	6480
gtaggggttt tatgttggtt ttaactcaga tgccataaga acaaagatac atgtaattta	6540
taatagtaat cattaatacc tatattgtgc ttttaagggtt aaaaaataat ttttctcata	6600
ctttatctta gtttagtttc ttgacagtcc atgaggtaag gtggtagctt tatcaccatt	6660
ttacaaagtg ggaaacgaag gttcctctta ggaacctagt tgtcaccttt gtataataaa	6720
acttcgaagc tcggagctgt taactgggtt gctgaaggct tagctgtaag agccagaatt	6780
cagaccagg tctgagtgac ttcaaactgc acagtccttc ccattattac ccatatgcta	6840
tcccttatat ttttaattta ttaggaattc attcatttat aaacttggtg attcaccttt	6900
attagattct ggtcgctgaa ggcttttagta acttcagagt aaaacttgag agatgagatg	6960
taaaatgcag ccattcttga gagttccttt ttctgtaaca ttcatacaaca cttcattgag	7020
aagtgaaggt tcctatggct gtctctacct tcaagaggct tagctttagt cactgagaaa	7080
gacaaggaaa ctaatgatag aatatagtag cttcttctgg cgttaggtat cacagagtca	7140
cagctagtta cagctagccc tttattattg aaagaagagg agctagcagt cccactatca	7200
gaattaagac tagagatggg aataggagct agtatcagaa aagcttaagg caaagcataa	7260
agtgtaggct agaatgaagc tggagaatgg ggagggggct tgggtaacat ccagaacctg	7320
gctggggacc tggaactaca tgagatgtaa gaatggagag gttctagcag tcagaggtca	7380
ggtacaaatg aacagctggg atctgcgcac ggcagacagt gaaaaaaccc aggcaagcaa	7440
aatggctcaga gcagaaaggg gcccaaggcc acgttcttga gatgtggagg gggctgagga	7500
agccacgcca agtaaggaca gatgcagctc agcagttcct agcgagccct gacaagccag	7560
ctcagctgaa gcttcgggtg ggagccagtc atggcacagt ggagtgaagg aagagcagtt	7620
tcaggcacc aaaaactgac cccacgacc tgttttccac ctgaagagcc acccattcca	7680
tccaaaccct tggcaaaagt ctgctaacag agagaaccgg ccagtatgct ggccagtcgc	7740
gatcatgcct gtctttaccc tctaagctga agctgctcat caacggtgag atggcaaaaa	7800
ggtgggtcca gaagagggga aaagaaggga gtctgtgaaa aaaaaatgct gaagaatctg	7860
catcaaataa acccttcctt ccttcctttt tccttc	7896

<210> 56

<211> 2746

<212> DNA

<213> Homo sapiens

<220>

<223> >MRE11A|ENSG0000020922|ENST00000323929|2746

<400> 56

agccaatgag agccgaactg gacttgaagc atctacgtta tccatgaagt gtcgcgagag	60
aaacggacgc cgttctctcc cgcggaattc aggtttacgg ccctgcgggt tctcagagaa	120
tttctagaat ttggaatcga gtgcattttc tgacatttga gtacagtacc caggggttct	180
tggagaagaa cctggtccca gaggagcttg actgaccata aaaatgagta ctgcagatgc	240
acttgatgat gaaaacacat ttaaaatatt agttgcaaca gatattcatc ttggatttat	300
ggagaaagat gcagtcagag gaaatgatac gtttgtaaca ctcatgaaa ttttaagact	360
tgcccaggaa aatgaagtgg attttatatt gttaggtggg gatctttttc atgaaaataa	420
gccctcaagg aaaacattac atacctgcct cgagttatta agaaaatatt gtatgggtga	480
tcggcctgtc cagtttgaaa ttctcagtga tcagtcagtc aactttggtt ttagtaagtt	540
tccatgggtg aactatcaag atggcaacct caacatttca attccagtgt ttagtattca	600
tggcaatcat gacgatccca caggggcaga tgcactttgt gccttggaaca ttttaagttg	660
tgctggattt gtaaatcact ttggacgttc aatgtctgtg gagaagatag acattagtcc	720
ggttttgctt caaaaaggaa gcacaaagat tgcgctatat ggtttaggat ccattccaga	780
tgaaaggctc tatcgaatgt ttgtcaataa aaaagtaaca atgttgagac caaaggaaga	840
tgagaactct tggtttaact tatttgtgat tcatcagaac aggagtaaac atggaagtac	900
taacttcatt ccagaacaat ttttgatga cttcattgat cttgttatct ggggccatga	960
acatgagtgt aaaatagctc caacaaaaa tgaacaacag ctgttttata tctcacaacc	1020
tggaagctca gtggttactt ctctttcccc aggagaagct gtaaagaaac atgttggttt	1080
gctgcgtatt aaaggaggga agatgaatat gcataaaatt cctcttcaca cagtgcggca	1140
gtttttcatg gaggatattg ttctagctaa tcatccagac atttttaacc cagataatcc	1200
taaagtaacc caagccatac aaagcttctg tttggagaag attgaagaaa tgcttgaaaa	1260
tgctgaacgg gaacgtctgg gtaattctca ccagccagag aagcctcttg tacgactgcg	1320
agtggactat agtggagggt ttgaacctt cagtgttctt cgcttttagcc agaaatttgt	1380
ggatcgggta gctaattcaa aagacattat ccattttttc aggcatagag aacaaaagga	1440

aaaaacagga gaagagatca actttgggaa acttatcaca aagccttcag aaggaacaac	1500
tttaagggtta gaagatcttg taaaacagta ctttcaaacc gcagagaaga atgtgcagct	1560
ctcactgcta acagaaagag ggatgggtga agcagtacaa gaatttgttg acaaggagga	1620
gaaagatgcc attgaggaat tagtgaaata ccagttggaa aaaacacagc gatttcttaa	1680
agaacgtcat attgatgccc tcgaagacaa aatcgatgag gaggtacgtc gtttcagaga	1740
aaccagacaa aaaaatacta atgaagaaga tgatgaagtc cgtgaggcta tgaccagggc	1800
cagagcactc agatctcagt cagaggagtc tgcttctgcc tttagtgtg atgaccttat	1860
gagtatagat ttagcagaac agatggctaa tgactctgat gatagcatct cagcagcaac	1920
caacaaagga agaggccgag gaagaggtcg aagaggtgga agagggcaga attcagcatc	1980
gagaggaggg tctcaaagag gaagagcaga cactgggtctg gagacttcta cccgtagcag	2040
gaactcaaag actgctgtgt cagcatctag aaatatgtct attatagatg cctttaaatc	2100
tacaagacag cagccttccc gaaatgtcac tactaagaat tattcagagg tgattgaggt	2160
agatgaatca gatgtggaag aagacatctt tcctaccact tcaaagacag atcaaaggtg	2220
gtccagcaca tcatccagca aaatcatgtc ccagagtcaa gtatcgaaag gggttgattt	2280
tgaatcaagt gaggatgatg atgatgatcc ttttatgaac actagttctt taagaagaaa	2340
tagaagataa tatatttaat ggcactgaga aacatgcaag atacaggaaa aatgaaaatg	2400
ttacaagcta agagtttaca gtttaagatt ttaagtattg tttcctgagc ataactccat	2460
aagtaagaaa tttctagtct acagacatac aatagcattg attcaccttg tttttttaac	2520
ctggttgttg tagtaagagc tttgtttcaa tatcactctt gagtaaagat taaaataaag	2580
ctaccatttt acatttctat ttcataatga aaaactatgt cagtatttta atatggttac	2640
atttagccaa agttgaggga aagagcttat aaaatttaac ttcttcataa ttttagtaat	2700
ttcctagagg ttctgggttt tctgaaagta aaacaattta tgcgaa	2746

<210> 57

<211> 3307

<212> DNA

<213> Homo sapiens

<220>

<223> >MSH2|ENSG00000095002|ENST00000233146|3307

<400> 57

gcagtagcta aagtcaccag cgtgcgcggg aagctgggcc gcgtctgctt atgattggtt	60
gccgcggcag actcccaccc accgaaacgc agccctggaa gctgattggg tgtggtcgcc	120
gtggccggac gccgctcggg ggacgtggga ggggaggcgg gaaacagctt agtgggtgtg	180
gggtcgcgca ttttcttcaa ccaggaggtg aggaggtttc gacatggcgg tgcagccgaa	240
ggagacgctg cagttggaga gcgcggccga ggtcggcttc gtgcgcttct ttcagggcat	300
gccggagaag ccgaccacca cagtgcgcct tttcgaccgg ggcgacttct atacggcgca	360
cggcgaggac gcgctgctgg ccgcccggga ggtgttcaag acccaggggg tgatcaagta	420
catggggccg gcaggagcaa agaacttgca gagtgttgtg cttagtaaaa tgaattttga	480
atcttttgta aaagatcttc ttctggttcg tcagtataga gttgaagttt ataagaatag	540
agctggaaat aaggcatcca aggagaaatga ttggtatttg gcatataagg cttctcctgg	600
caatctctct cagtttgaag acattctctt tggttaacaat gatatgtcag cttccattgg	660
tgttgtgggt gttaaaatgt ccgcagttga tggccagaga caggttggag ttgggtatgt	720
ggattccata cagaggaaac taggactgtg tgaattccct gataatgatc agttctccaa	780
tcttgaggct ctctcatcc agattggacc aaaggaatgt gttttaccg gaggagagac	840
tgctggagac atggggaaac tgagacagat aattcaaaga ggaggaattc tgatcacaga	900
aagaaaaaaa gctgactttt ccacaaaaga catttatcag gacctcaacc ggttgttgaa	960
aggcaaaaag ggagagcaga tgaatagtgc tgtattgcc aaatggaga atcaggttgc	1020
agtttcatca ctgtctgcgg taatcaagtt tttagaactc ttatcagatg attccaactt	1080
tggacagttt gaactgacta cttttgactt cagccagtat atgaaattgg atattgcagc	1140
agtcagagcc cttaaccttt ttcagggttc tgttgaagat accactggct ctcagtctct	1200
ggctgccttg ctgaataagt gtaaaacccc tcaaggacaa agacttgta accagtggat	1260
taagcagcct ctcatggata agaacagaat agaggagaga ttgaatttag tggaagcttt	1320
tgtagaagat gcagaattga ggcagacttt acaagaagat ttacttcgtc gattcccaga	1380
tcttaaccga cttgccaaga agtttcaaag acaagcagca aacttacaag attgttaccg	1440
actctatcag ggtataaatc aactaccta tgttatacag gctctggaaa aacatgaagg	1500
aaaacaccag aaattattgt tggcagtttt tgtgactcct cttactgatc ttcgttctga	1560

cttctccaag tttcaggaaa tgatagaaac aacttttagat atggatcagg tggaaaacca	1620
tgaattcctt gtaaaacctt catttgatcc taatctcagt gaattaagag aaataatgaa	1680
tgacttggaa aagaagatgc agtcaacatt aataagtga gccagagatc ttggcttggaa	1740
ccctggcaaa cagattaaac tggattccag tgcacagttt ggatattact ttcgtgtaac	1800
ctgtaaggaa gaaaaagtcc ttcgtaacaa taaaaacttt agtactgtag atatccagaa	1860
gaatggtgtt aaatttacca acagcaaatt gacttcttta aatgaagagt ataccaaaaa	1920
taaaacagaa tatgaagaag cccaggatgc cattgttaaa gaaattgtca atatttcttc	1980
aggctatgta gaaccaatgc agacactcaa tgatgtgtta gctcagctag atgctgttgt	2040
cagctttgct cacgtgtcaa atggagcacc tgttccatat gtacgaccag ccattttgga	2100
gaaaggacaa ggaagaatta tattaaaagc atccaggcat gcttgtgttg aagttcaaga	2160
tgaaattgca tttattccta atgacgtata ctttgaaaaa gataaacaga tgttccacat	2220
cattactggc cccaatatgg gaggtaaatc aacatatatt cgacaaactg gggatgatagt	2280
actcatggcc caaattgggt gttttgtgcc atgtgagtca gcagaagtgt ccattgtgga	2340
ctgcatctta gcccagtagt gggctggtga cagtcaattg aaaggagtct ccacgttcat	2400
ggctgaaatg ttggaaactg cttctatcct caggtctgca accaaagatt cattaataat	2460
catagatgaa ttgggaagag gaacttctac ctacgatgga tttgggttag catgggctat	2520
atcagaatac attgcaacaa agattggtgc tttttgcatg tttgcaacc attttcatga	2580
acttactgcc ttggccaatc agataccaac tgtaataat ctacatgtca cagcactcac	2640
cactgaagag accttaacta tgctttatca ggtgaagaaa ggtgtctgtg atcaaagttt	2700
tgggattcat gttgcagagc ttgctaattt ccctaagcat gtaatagagt gtgctaaaca	2760
gaaagccctg gaacttgagg agtttcagta tattggagaa tcgcaaggat atgatatcat	2820
ggaaccagca gcaaagaagt gctatctgga aagagagcaa ggtgaaaaaa ttattcagga	2880
gttcctgtcc aaggtgaaac aaatgccctt tactgaaatg tcagaagaaa acatcacaat	2940
aaagttaaaa cagctaaaag ctgaagtaat agcaaagaat aatagctttg taaatgaaat	3000
catttcacga ataaaagtta ctacgtgaaa aatcccagta atggaatgaa ggtaatatgt	3060
ataagctatt gtctgtaata gttttatatt gttttatatt aacccttttt ccatagtgtt	3120
aactgtcagt gcccatgggc tatcaactta ataagatatt tagtaatat ttactttgag	3180

gacattttca aagattttta ttttgaaaaa tgagagctgt aactgaggac tgtttgcaat	3240
tgacataggc aataataagt gatgtgctga attttataaa taaaatcatg tagtttgtgg	3300
aatttga	3307

<210> 58
 <211> 4092
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >MSH3|ENSG00000113318|ENST00000265081|4092

<400> 58	
ccgcagacgc ctgggaactg cggccgcggg ctcgcgctcc tcgccaggcc ctgccgccgg	60
gctgccatcc ttgccctgcc atgtctcgcc ggaagcctgc gtcgggcggc ctcgctgcct	120
ccagctcagc ccctgcgagg caagcggttt tgagccgatt cttccagtct acgggaagcc	180
tgaaatccac ctctctctcc acaggtgcag ccgaccaggt ggaccctggc gctgcagcgg	240
ctgcagcggc cgcagcggcc gcagcgcccc cagcgcccc agctcccgcc ttcccgcgcc	300
agctgccgcc gcacatagct acagaaattg acagaagaaa gaagagacca ttggaaaatg	360
atgggcctgt taaaaagaaa gtaaagaaag tccaacaaaa ggaaggagga agtgatctgg	420
gaatgtctgg caactctgag ccaaagaaat gtctgaggac caggaatgtt tcaaagtctc	480
tggaaaaatt gaaagaattc tgctgcgatt ctgcccttcc tcaaagtaga gtccagacag	540
aatctctgca ggagagattt gcagttctgc caaatgtac tgattttgat gatatcagtc	600
ttctacacgc aaagaatgca gtttcttctg aagattcgaa acgtcaaatt aatcaaaagg	660
acacaacact tttgatctc agtcagtttg gatcatcaaa tacaagtcac gaaaatttac	720
agaaaactgc ttccaaatca gctaacaaac ggtccaaaag catctatacg ccgctagaat	780
tacaatacat agaaatgaag cagcagcaca aagatgcagt tttgtgtgtg gaatgtggat	840
ataagtatag attctttggg gaagatgcag agattgcagc ccgagagctc aatattttatt	900
gccattttaga tcacaacttt atgacagcaa gtatacctac tcacagactg tttgttcatg	960
tacgccgcct ggtggcaaaa ggatataagg tgggagttgt gaagcaaact gaaactgcag	1020
cattaaaggc cattggagac aacagaagtt cactcttttc ccggaaattg actgcccttt	1080
atacaaaatc tacacttatt ggagaagatg tgaatcccct aatcaagctg gatgatgctg	1140

taaatgttga tgagataatg actgatactt ctaccagcta tcttctgtgc atctctgaaa	1200
ataaggaaaa tgttagggac aaaaaaaagg gcaacatttt tattggcatt gtgggagtg	1260
agcctgccac aggcgagggt gtgtttgata gtttccagga ctctgcttct cgttcagagc	1320
tagaaacccg gatgtcaagc ctgcagccag tagagctgct gcttccttcg gccttgtccg	1380
agcaaacaga ggcgctcatc cacagagcca catctgttag tgtgcaggat gacagaattc	1440
gagtcgaaag gatggataac atttattttg aatacagcca tgctttccag gcagttacag	1500
agttttatgc aaaagataca gttgacatca aaggtttctca aattatttct ggcattgtta	1560
acttagagaa gcctgtgatt tgctctttgg ctgccatcat aaaatacctc aaagaattca	1620
acttggaaaa gatgctctcc aaacctgaga attttaaaca gctatcaagt aaaatggaat	1680
ttatgacaat taatggaaca acattaagga atctggaaat cctacagaat cagactgata	1740
tgaaaaccaa aggaagtttg ctgtgggttt tagaccacac taaaacttca tttgggagac	1800
ggaagttaaa gaagtgggtg acccagccac tccttaaatt aagggaata aatgcccggc	1860
ttgatgctgt atcggaagtt ctccattcag aatctagtgt gtttggtcag atagaaaatc	1920
atctacgtaa attgcccagc atagagaggg gactctgtag catttatcac aaaaaatgtt	1980
ctaccaaga gttcttcttg attgtcaaaa ctttatatca cctaaagtca gaatttcaag	2040
caataatacc tgctgttaat tcccacattc agtcagactt gctccggacc gttattttag	2100
aaattcctga actcctcagt ccagtggagc attacttaaa gatactcaat gaacaagctg	2160
ccaaagttag ggataaaact gaattattta aagacctttc tgacttccct ttaataaaaa	2220
agaggaagga tgaaattcaa ggtgttattg acgagatccg aatgcatttg caagaaatac	2280
gaaaaatact aaaaaatcct tctgcacaat atgtgacagt atcaggacag gagtttatga	2340
tagaaataaa gaactctgct gtatcttgta taccaactga ttgggtaaag gttggaagca	2400
caaaagctgt gagccgcttt cactctcctt ttattgtaga aaattacaga catctgaatc	2460
agctccggga gcagctagtc cttgactgca gtgctgaatg gcttgatttt ctagagaaat	2520
tcagtgaaca ttatcactcc ttgtgtaaag cagtgcacat cctagcaact gttgactgca	2580
ttttctccct ggccaaggct gctaagcaag gagattactg cagaccaact gtacaagaag	2640
aaagaaaaat tgtaataaaa aatggaaggc accctgtgat tgatgtgttg ctgggagaac	2700
aggatcaata tgtcccaaat aatacagatt tatcagagga ctacagagaga gtaatgataa	2760

ttaccggacc aaacatgggt ggaaagagct cctacataaa acaagttgca ttgattacca	2820
tcatggctca gattggctcc tatgttcctg cagaagaagc gacaattggg attgtggatg	2880
gcattttcac aaggatgggt gctgcagaca atatatataa aggacagagt acatttatgg	2940
aagaactgac tgacacagca gaaataatca gaaaagcaac atcacagtcc ttggttatct	3000
tggatgaact aggaagaggg acgagcactc atgatggaat tgccattgcc tatgctacac	3060
ttgagtatth catcagagat gtgaaatcct taaccctgtt tgtcacccat tatccgccag	3120
tttgtgaact agaaaaaaat tactcacacc aggtggggaa ttaccacatg ggattcttgg	3180
tcagtgagga tgaaagcaaa ctggatccag gcgcagcaga acaagtcctt gatthttgtca	3240
ccttccttta ccaaataact agaggaattg cagcaaggag ttatggatta aatgtggcta	3300
aactagcaga tgttcctgga gaaatthtga agaaagcagc tcacaagtca aaagagctgg	3360
aaggattaat aaatacgaaa agaaagagac tcaagtatth tgcaaagtta tggacgatgc	3420
ataatgcaca agacctgcag aagtggacag aggagthcaa catggaagaa acacagactt	3480
ctcttcttca ttaaaatgaa gactacatth gtgaacaaaa aatggagaat taaaaatacc	3540
aactgtacaa aataactctc cagtaacagc ctatctthgt gtgacatgtg agcataaaat	3600
tatgaccatg gtatatthct atttgaaaca gagaggttht tctgaagaca gtctthtttca	3660
agthttctgtc thcctaactt thctacgtat aaacactctt gaatagactt ccactthtga	3720
attagaaaat thtatggaca gtaagtccag taaagcctta agtggcagaa tataattccc	3780
aagctthttg aggggtgatat aaaaatthac thgatattht tathttgttc agttcagata	3840
attggcaact ggggtgaatct ggcaggaatc tatccattga actaaaataa thttattatg	3900
caaccagtht atccaccaag aacataagaa thththtataa gtagaaagaa ttggccaggc	3960
atgggtggctc atgcctgtaa tcccagcact ttggggaggcc aaggtaggca gatcacctga	4020
ggtcaggagt tcaagaccag cctggccaac atggcaaaac cccatcttht ctaaaaatat	4080
aaagtacatc tc	4092

<210> 59

<211> 7476

<212> DNA

<213> Homo sapiens

<220>

<223> >MSH6|ENSG00000116062|ENST00000234420|7476

<400> 59

ggcgaggcgc ctgttgattg gccactgggg cccgggttcc tccggcggag cgcgctccc	60
cccagatttc ccgccagcag gagccgcgcg gtagatgcgg tgcttttagg agctccgtcc	120
gacagaacgg ttgggccttg ccggctgtcg gtatgtcgcg acagagcacc ctgtacagct	180
tcttcccaa gtctccggcg ctgagtgatg ccaacaaggc ctcggccagg gcctcacgcg	240
aaggcggccg tgccgccgct gccccgggg cctctccttc cccaggcggg gatgcggcct	300
ggagcgaggc tgggcctggg cccaggccct tggcgcgctc cgcgtcaccg cccaaggcga	360
agaacctcaa cggagggctg cggagatcgg tagcgcctgc tgccccacc agttgtgact	420
tctcaccagg agatttggtt tgggccaaga tggagggtta cccctggtgg cttgtctgg	480
tttacaacca ccccttgat ggaacattca tccgcgagaa agggaaatca gtccgtgttc	540
atgtacagtt ttttgatgac agcccaacaa ggggctgggt tagcaaaagg cttttaaagc	600
catatacagg ttcaaaatca aaggaagccc agaaggagg tcatttttac agtgcaaagc	660
ctgaaatact gagagcaatg caacgtgcag atgaagcctt aaataaagac aagattaaga	720
ggcttgaatt ggcagtttgt gatgagccct cagagccaga agaggaagaa gagatggagg	780
taggcacaac ttacgtaaca gataagagtg aagaagataa tgaaattgag agtgaagagg	840
aagtacagcc taagacacaa ggatctaggc gaagtagccg ccaaataaaa aaacgaaggg	900
tcatatcaga ttctgagagt gacattggtg gctctgatgt ggaatttaag ccagacacta	960
aggaggaagg aagcagtgat gaaataagca gtggagtggg ggatagtgag agtgaaggcc	1020
tgaacagccc tgtcaaagtt gctcgaaagc ggaagagaat ggtgactgga aatggctctc	1080
ttaaaaggaa aagctctagg aaggaaacgc cctcagccac caaacaagca actagcattt	1140
catcagaaac caagaatact ttgagagctt tctctgcccc tcaaaattct gaatcccaag	1200
cccacgttag tggaggtggg gatgacagta gtcgccctac tgtttggtat catgaaactt	1260
tagaatggct taaggaggaa aagagaagag atgagcacag gaggaggcct gatcaccccg	1320
attttgatgc atctacactc tatgtgcctg aggatttcct caattcttgt actcctggga	1380
tgaggaagtg gtggcagatt aagtctcaga actttgatct tgtcatctgt tacaaggtgg	1440
ggaaatttta tgagctgtac cacatggatg ctcttattgg agtcagtga ctggggctgg	1500

tattcatgaa aggcaactgg gccattctg gctttcctga aattgcattt ggccgttatt	1560
cagattccct ggtgcagaag ggctataaag tagcacgagt ggaacagact gagactccag	1620
aaatgatgga ggcacgatgt agaaagatgg cacatatatc caagtatgat agagtgggtga	1680
ggagggagat ctgtaggatac attaccaagg gtacacagac ttacagtgtg ctggaagggtg	1740
atccctctga gaactacagt aagtatcttc ttagcctcaa agaaaaagag gaagattctt	1800
ctggccatac tcgtgcatat ggtgtgtgct ttgttgatac ttcactggga aagtttttca	1860
taggtcagtt ttcagatgat cgccattgtt cgagatttag gactctagtg gcacactatc	1920
ccccagtaca agttttattt gaaaaaggaa atctctcaaa ggaaactaaa acaattctaa	1980
agagttcatt gtcctgttct cttcaggaag gtctgatacc cggctcccag ttttgggatg	2040
catccaaaac tttgagaact ctccttgagg aagaatattt tagggaaaag ctaagtgatg	2100
gcattgggggt gatgttacc caggtgctta aaggtatgac ttcagagtct gattccattg	2160
ggttgacacc aggagagaaa agtgaattgg ccctctctgc tctaggtgggt tgtgtcttct	2220
acctcaaaaa atgccttatt gatcaggagc ttttatcaat ggctaatttt gaagaatata	2280
ttcccttgga ttctgacaca gtcagcacta caagatctgg tgctatcttc accaaagcct	2340
atcaacgaat ggtgctagat gcagtacat taaacaactt ggagattttt ctgaatggaa	2400
caaatgggtt tactgaagga accctactag agagggttga tacttgccat actccttttg	2460
gtaagcggct cctaaagcaa tggctttgtg cccactctg taaccattat gctattaatg	2520
atcgtctaga tgccatagaa gacctatgg ttgtgcctga caaatctcc gaagttgtag	2580
agcttctaaa gaagcttcca gatcttgaga ggctactcag taaaattcat aatgttggtt	2640
ctcccctgaa gagtccagac caccagaca gcagggtat aatgtatgaa gaaactacat	2700
acagcaagaa gaagattatt gattttcttt ctgctctgga aggattcaaa gtaatgtgta	2760
aaattatagg gatcatggaa gaagttgctg atggttttta gtctaaaatc cttagcagg	2820
tcattctctt gcagacaaaa aatcctgaag gtcgttttcc tgatttgact gtagaattga	2880
accgatggga tacagccttt gaccatgaaa aggctcgaaa gactggactt attactccca	2940
aagcaggctt tgactctgat tatgaccaag ctcttgctga cataagagaa aatgaacaga	3000
gcctcctgga atacctagag aaacagcgca acagaattgg ctgtaggacc atagtctatt	3060
gggggattgg taggaaccgt taccagctgg aaattcctga gaatttcacc actcgcaatt	3120

tgccagaaga atacgagttg aaatctacca agaagggtg taaacgatac tggaccaaaa	3180
ctattgaaaa gaagttggct aatctcataa atgctgaaga acggagggat gtatcattga	3240
aggactgcat gcggcgactg ttctataact ttgataaaaa ttacaaggac tggcagtctg	3300
ctgtagagtg tatcgcagtg ttggatgttt tactgtgcct ggctaactat agtcgagggg	3360
gtgatggtcc tatgtgtcgc ccagtaattc tgttgccgga agataccccc cccttcttag	3420
agcttaaagg atcacgcat ccttgcatta cgaagacttt ttttggagat gattttattc	3480
ctaatacat tctaataaggc tgtgaggaag aggagcagga aaatggcaaa gcctattgtg	3540
tgcttggtac tggaccaa atggggggca agtctacgt tatgagacag gctggcttat	3600
tagctgtaat ggcccagatg ggttggttacg tccctgctga agtgtgcagg ctcacaccaa	3660
ttgatagagt gtttactaga cttgggtgcct cagacagaat aatgtcaggt gaaagtacat	3720
tttttggtga attaagtga actgccagca tactcatgca tgcaacagca catttcttg	3780
tgcttggtga tgaattagga agaggtactg caacatttga tgggacggca atagcaa atg	3840
cagttgttaa agaacttgct gagactataa aatgtcgtac attattttca actcactacc	3900
attcattagt agaagattat tctcaaaatg ttgctgtgcg cctaggacat atggcatgca	3960
tggtagaaaa tgaatgtga gaccccagcc aggagactat tacgttcctc tataaattca	4020
ttaaggagc ttgtcctaaa agctatggct ttaatgcagc aaggcttgct aatctcccag	4080
aggaagtat tcaaaagga catagaaaag caagagaatt tgagaagatg aatcagtcac	4140
tacgattatt tcgggaagtt tgcctggcta gtgaaaggtc aactgtagat gctgaagctg	4200
tccataaatt gctgactttg attaaggaat tatagactga ctacattgga agctttgagt	4260
tgacttctga caaaggtggt aaattcagac aacattatga tctaataaac tttatttttt	4320
aaaaatgacc atttttccat tttctttcta ggaaattaaa ccctttta at tcttatctac	4380
cttctacata atggttattg aatactccac aatatattaa gtctagatgt tatggtacat	4440
gcatacactt tcaggctggt ttataccac tgtcaccaat acacataaat gggggaggaa	4500
aagctatgaa actgtatagg gctgtatata tacttgtctc agcttaatgc aggaaattgg	4560
tttaatttcc agcagttttg tctaaactgt tcaaaaaaaaa actatgaaca gagttcaaat	4620
acaggactgt ttgttttgaa gagactttct aaagtgtact taaaacatag tagtttttta	4680
cccttcacaa aactgagtta caagaatact tttgttttac agtgcattccc ttcctaggaa	4740

gtctcattaa aacactcact ttttctaggg gtgattttga atgctgcaca gggaagggaa	4800
ggaaataata gtcttaactt ttcttaaagg ataccagaaa cattgctgga tataatttaa	4860
gattagtgtt ttctctttca tagaaagaac gtacatactg ggacatgagt acagttacag	4920
caagtctagg tgtgctaaca aaacagggca cattcaagta cagtaagatt ttgcttgaaa	4980
ttaaaaacaa actacatgag attaaagcat taaaatcata tttctcaatc tgaatacatg	5040
ttaaaaaaaaa aaaatcaaaa ggaacgcaga agtgctagct cacattttta ccatattaca	5100
aaagcaattg gtacccatgt ccataaaggc agcaacaaag ctgcttgtct attgaagatt	5160
actactgcaa attggactgc attcaatgct agttgtaaaa acaccagctt ttcagaagtt	5220
ggtatctgta caaaattgca gcttattttc ttcacttctg tcccttcaag tctttacaca	5280
gtaatgctaa aacaccagc tttgagatcc tgagtcaata tattgccact ttctttttgg	5340
tagcttgagc ttcatagtgt caactgacct tgtgtatcca tttttaatac agtctcttcc	5400
tgtagcatgg gcaaatatth taaatcttct tccaaaaaag tgttttaagt tatgatgtta	5460
caatggcagg actttttctt tagggaagga attcagttgt gctgcaatgt attagattct	5520
ataggtggag cagagtcata tagtgtatct gtatcatgtg taggctcacc agctaagtta	5580
caaggattag acagtgttcc agcaccacag tcacagaaaa acctaaagca aaatgaaacc	5640
caaatattag aaaagtgagg gggaaagtaa ttgggtaata tatcaagcaa gtgtgctaca	5700
tacctatcat gtctaataaa ctctacatca tgtccctgat ggcaattctt aatgcagttc	5760
acacatatgg catttcgatc tgtgggtgta caagtatgac atctaataag caaaagctta	5820
aattactttt ctcaaactg tcattaatgc aaaacattcc attctgttta tatattacta	5880
tgacctttgg ctttaagagg accaaaacaa aattctttgt ggctccagcc cagattaatt	5940
ctgaaaagga actttaatgg agtaagtgat tttcctgtca tctgtgtctt cggaggggaag	6000
agaaatgatt tgtaaattgt ataaaggcag ttctttccac tttaaaagcc tctcaaatgt	6060
ttctgggctg aaaacaatth ttggaggcgt gaagagtcaa aactgtcaca gtgactggga	6120
tatatcaaac acttaacccc gacatcttta cttgaaatt tctaggaaaa cattacacaa	6180
catgagttac atgaatgaca tcagttactg tagcattagg tttttccata gttatggtct	6240
ttgttttgtt ttgtagagac agggctctccc tatgttgccc aggctggttt agaactcctg	6300
ggctccagtg atcctcccac ttcagcctcc caaagtgcta ggattacagg cataagccac	6360

cacgcctacc cacagttaca gtcttaaaca cgatcttcaa gtagattgat gataaaat	6420
ttcagttagtt atagtctcaa caccggcaaa tagccaaaaa tgctaggcat tgctaattta	6480
aaaaggaaat cagtcttcct cttttcagga ctcaaataa tttctaagtt acctgtagaa	6540
atcatgcatg ggatagctgg tataacttga ttttttatat aaacattggc ctctactaac	6600
agccttttct atggcatctt gattgttcat tttttgtta tctgtaataa aagaaagaat	6660
aagtaaaaat tcagaggaat gttaatatat taaaaacaa agattatagg attattctaa	6720
cagaagagcc actattttta agagctttaa atgaagctaa ccaatgaagt aattgtaaga	6780
aatcagctaa gaatagaatt ttccttgat aagatactcc aaccatttag aaccaaagct	6840
ctgtttcttt caaaatctat cttaaactgt tgctaacttg gagagtgaca taaggaatca	6900
agttataaaa cggcttctga ttatctttca tggcatattg catatattta taggtatagc	6960
agactccaac atacctttca ttgtcacatt aacaccagat gctaaaaata agcctccaaa	7020
ccggttgta aaaatctgat tgccttctag tgttgcagtt gcgtgatttg taatttcaat	7080
acctgaagta aaatttaca acaagtagat acatcacttt atactgcttc ttaaaaacct	7140
gaaattagca agcaaagtga aactgcttct tttatagaag tacattaacc ctcttaatgt	7200
ctactgaata aaatgtagat acctatttca accaccaaca gtaacattca cttatcaatg	7260
actatgggtca aaactgcaat taactttcgc accaacctaa ctgtcttaaa gtttaaatac	7320
atgatacttg gatttcattt gcatccattt taacatctct ttttctgttg cagatttaaa	7380
ctggtaaaat catctgagga attgaatcta tctgtattcc tagtggtaat acaagcctgc	7440
atttattcta tccaataaa tgtttcataa tcacga	7476

<210> 60
 <211> 2345
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >MYC|ENSG00000136997|ENST00000377970|2345

<400> 60	
ctgctcgcgg ccgccaccgc cgggccccgg ccgtccctgg ctcccctcct gcctcgagaa	60
gggcagggct tctcagaggc ttggcgggaa aaagaacgga gggagggatc gcgctgagta	120
taaaagccgg ttttcggggc tttatctaac tcgctgtagt aattccagcg agaggcagag	180

ggagcgagcg ggcggccggc tagggtggaa gagccgggcg agcagagctg cgctgcgggc	240
gtcctgggaa gggagatccg gagcgaatag ggggcttcgc ctctggccca gccctcccgc	300
tgatcccca gccagcggtc cgcaaccctt gccgcatcca cgaaactttg cccatagcag	360
cgggcgggca ctttgactg gaacttacia caccgagca aggacgcgac tctcccgacg	420
cggggaggct attctgcca tttggggaca cttccccgcc gctgccagga cccgcttctc	480
tgaaaggctc tccttgacg tgcttagacg ctggatTTTT ttcgggtagt ggaaaaccag	540
cagcctcccg cgacgatgcc cctcaacgtt agcttcacca acaggaacta tgacctcgac	600
tacgactcgg tgcagccgta tttctactgc gacgaggagg agaacttcta ccagcagcag	660
cagcagagcg agctgcagcc cccggcgccc agcgaggata tctggaagaa attcgagctg	720
ctgcccacc cgcccctgtc ccctagccgc cgctccgggc tctgctcgcc ctctacgtt	780
gcggtcacac ctttctccct tcggggagac aacgacggcg gtggcgggag cttctccacg	840
gccgaccagc tggagatggt gaccgagctg ctgggaggag acatggtgaa ccagagtttc	900
atctgcgacc cggacgacga gaccttcac aaaaacatca tcatccagga ctgtatgtgg	960
agcggcttct cggccgccgc caagctcgtc tcagagaagc tggcctccta ccaggctgcg	1020
cgaaaagaca gcggcagccc gaaccccgcc cgcggccaca gcgtctgctc cacctccagc	1080
ttgtacctgc aggatctgag cgccgccgcc tcagagtga tcgaccctc ggtggtcttc	1140
ccctaccctc tcaacgacag cagctcgccc aagtcctgcg cctcgcaaga ctccagcgcc	1200
ttctctccgt cctcgattc tctgctctcc tcgacggagt cttccccgca gggcagcccc	1260
gagcccctgg tgctccatga ggagacaccg cccaccacca gcagcgactc tgaggaggaa	1320
caagaagatg aggaagaaat cgatgttggt tctgtggaaa agaggcaggc tcctggcaaa	1380
aggtcagagt ctggatcacc ttctgctgga ggccacagca aacctcctca cagcccactg	1440
gtcctcaaga ggtgccacgt ctccacacat cagcacaact acgcagcgcc tccctccact	1500
cggaaggact atcctgctgc caagagggtc aagttggaca gtgtcagagt cctgagacag	1560
atcagcaaca accgaaaatg caccagcccc aggtcctcgg acaccgagga gaatgtcaag	1620
aggcgaacac acaacgtctt ggagcgccag aggaggaacg agctaaaacg gagctttttt	1680
gccctgcgtg accagatccc ggagttggaa aacaatgaaa aggccccaa ggtagttatc	1740
cttaaaaaag ccacagcata catcctgtcc gtccaagcag aggagcaaaa gctcatttct	1800

gaagaggact tgttgcggaa acgacgagaa cagttgaaac acaaacttga acagctacgg	1860
aactcttgtg cgtaaggaaa agtaaggaaa acgattcctt ctaacagaaa tgtcctgagc	1920
aatcacctat gaacttgttt caaatgcatg atcaaatgca acctcacaac cttggctgag	1980
tcttgagact gaaagattta gccataatgt aaactgcctc aaattggact ttgggcataa	2040
aagaactttt ttatgcttac catctttttt ttttctttaa cagatttgta tttagaatt	2100
gtttttaaaa aattttaaga ttacacaaat gtttctctgt aaatattgcc attaaatgta	2160
aataacttta ataaaacgtt tatagcagtt acacagaatt tcaatcctag tatatagtac	2220
ctagtattat aggtactata aaccctaatt ttttttattt aagtacattt tgctttttaa	2280
agttgatttt tttctattgt ttttagaaaa aataaaataa ctggcaaata tatcattgag	2340
ccaaa	2345

<210> 61
 <211> 4666
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >NBN|ENSG00000104320|ENST00000265433|4666

<400> 61	
aagtcgcact cccgcctcat ccaaggcagc ctgcgtggct cccgggagcg cgcacgtccc	60
ggagcccatg ccgaccgcag gcgccgtatc cgcgctcgtc tagcagcccc ggttacgcgg	120
ttgcacgtcg gccccagccc tgaggagccg gaccgatgtg gaaactgctg cccgccgcgg	180
gccccggcagg aggagaacca tacagacttt tgactggcgt tgagtacgtt gttggaagga	240
aaaactgtgc cattctgatt gaaaatgata agtcgatcag ccgaaatcat gctgtgttaa	300
ctgctaactt ttctgtaacc aacctgagtc aaacagatga aatccctgta ttgacattaa	360
aagataattc taagtatggg acctttgtta atgaggaaaa aatgcagaat ggcttttccc	420
gaactttgaa gtcgggggat ggtattactt ttggagtgtt tggaagtaaa ttcagaatag	480
agtatgagcc tttggttgca tgctcttctt gtttagatgt ctctgggaaa actgctttaa	540
atcaagctat attgcaactt ggaggattta ctgtaaacia ttggacagaa gaatgcactc	600
accttgtcat ggtatcagtg aaagttacca ttaaaacaat atgtgcactc atttgtggac	660

gtccaattgt aaagccagaa tattttactg aattcctgaa agcagttgag tccaagaagc	720
agcctccaca aattgaaagt ttttaccac ctcttgatga accatctatt ggaagtaaaa	780
atgttgatct gtcaggacgg caggaaagaa aacaaatctt caaagggaaa acatttatat	840
ttttgaatgc caaacagcat aagaaattga gttccgcagt tgtctttgga ggtggggaag	900
ctaggttgat aacagaagag aatgaagaag aacataatth ctttttggct ccgggaacgt	960
gtgttggtga tacaggaata acaaactcac agaccttaat tcctgactgt cagaagaaat	1020
ggattcagtc aataatggat atgctccaaa ggcaaggtct tagacctatt cctgaagcag	1080
aaattggatt ggcggtgatt ttcattgacta caaagaatta ctgtgacatc cagggccatc	1140
ccagtacagg attaaagaca acaactccag gaccaagcct ttcacaaggc gtgtcagttg	1200
atgaaaaact aatgccaagc gccccagtga acactacaac atacgtagct gacacagaat	1260
cagagcaagc agatacatgg gatttgagtg aaaggccaaa agaaatcaaa gtctccaaaa	1320
tggaacaaaa attcagaatg ctttcacaag atgcaccac tgtaaaggag tcctgcaaaa	1380
caagctctaa taataatagt atggtatcaa atactttggc taagatgaga atcccaaact	1440
atcagctttc accaactaaa ttgccaagta taaataaaaag taaagatagg gcttctcagc	1500
agcagcagac caactccatc agaaactact ttcagccgtc taccaaaaaa agggaaaggg	1560
atgaagaaaa tcaagaaatg tcttcatgca aatcagcaag aatagaaacg tcttgttctc	1620
ttttagaaca aacacaacct gctacaccct cattgtggaa aaataaggag cagcatctat	1680
ctgagaatga gcctgtggac acaaactcag acaataactt atttacagat acagatttaa	1740
aatctattgt gaaaaattct gccagtaaat ctcatgctgc agaaaagcta agatcaaata	1800
aaaaaaggga aatggatgat gtggccatag aagatgaagt attggaacag ttattcaagg	1860
acacaaaacc agagttagaa attgatgtga aagttcaaaa acaggaggaa gatgtcaatg	1920
ttagaaaaag gccaaaggatg gatatagaaa caaatgacac tttcagtgat gaagcagtac	1980
cagaaagtag caaaatatct caagaaaatg aaattgggaa gaaacgtgaa ctcaaggaag	2040
actcactatg gtcagctaaa gaaatatcta acaatgacaa acttcaggat gatagtgaga	2100
tgcttccaaa aaagctgtta ttgactgaat ttagatcact ggtgattaaa aactctactt	2160
ccagaaatcc atctggcata aatgatgatt atggtcaact aaaaaatttc aagaaattca	2220
aaaaggtcac atatcctgga gcaggaaaac ttccacacat cattggagga tcagatctaa	2280

tagctcatca tgctcgaaag aatacagaac tagaagagtg gctaaggcag gaaatggagg	2340
tacaaaatca acatgcaaaa gaagagtctc ttgctgatga tcttttttaga tacaatcctt	2400
atttaaaaag gagaagataa ctgaggattt taaaaagaag ccatggaaaa acttcctagt	2460
aagcatctac ttcaggccaa caagggtata tgaatatata gtgtatagaa gcgatttaag	2520
ttacaatgtt ttatggccta aatttattaa ataaaatgca caaaactttg attcctttgt	2580
atgtaacaat tgtttgttct gttttcaggc tttgtcattg catctttttt tcatttttaa	2640
atgtgttttg tttattaaat agttaatata gtcacagttc aaaattctaa atgtacgtaa	2700
ggtaaagact aaagtcaccc ttccaccatt gtcctagcta cttggttccc ctcagaaaaa	2760
aattcatgat actcatttct tatgaatctt tccagggatt tttgagtcct attcaaattc	2820
ctatttttaa ataatttcct acacaaatga tagcataaca tatgcagtgt tctacacctt	2880
gcttttttac ttagtagatt aaaaattata ggaatatcaa tataatgttt ttaatatttt	2940
ttcttttcca ttatgctgta gtcttaccta aactctggtg atccaaacaa aatggcttca	3000
gtggtgcaga tgtcacctac atgttattct agtactagaa actgaagacc atgtggagac	3060
ttcatcaaac atgggttttag ttttcaccag aatggaaaga cctgtacccc tttttggtgg	3120
tcttactgag ctgggtgggt gtctgttttg agcttattta gagtcctagt tttcctactt	3180
ataaagtaga aatggtgaga ttgttttctt tttctacctt aaaggagat ggtaagaaac	3240
aatgaatgtc ttttttcaa ctttattgac aagtgatttt caagtctgtg ttcaaaaata	3300
tattcatgta cctgtgatcc agcaagaagg gagttccagt caagagtcac tacaactgat	3360
tagttgttta gagaatgaga aatggaacag tgaggaatgg aggccatatt tccatgactt	3420
cccttgtaaa cagaagcaac agaagggaca agaggctggc ctctacatca ctctcacctt	3480
ccaaatcttg tggaagtgca tctacttgcc agaaccaaat taacttactt ccaagttctg	3540
gctgcttgca ggtggaactc cagctgcaag ggagttaggg aaatgaaggt ctttttttaa	3600
aagcttctca gccttcctag ggaacagaaa ttgggtgagc caatctgcaa tttctactac	3660
aggcattgag accagttaga ttattgaaat attatagaga gttatgaaca cttaaattat	3720
gatagtggta tgacattgga tagaacatgg gatactttag aagtagaatt gacagggcat	3780
attagttgat gaaatggagt catttgagtc tcttaatagc catgtatcat aattaccaag	3840
tgaagctggg ggaacatatg gtctccattt tacagttaag gaatataatg gacagattaa	3900

tattgttctc tgtcatgccc acaatccctt tctaaggaag actgccctac tatagcagtt	3960
tttatatttg tcaatttatg aatataatga atgagagttc tggtagctcc tgtctttaca	4020
aatattgggtg ttgtcagtat ttttcctttt taaccattcc aatcggtgtg tagtgatggt	4080
tcattttggt ttttaatttg ataccctga tagctataat tgggtcatag aaattcttta	4140
tacattctag atgcaagtct cttgtcggat atatgtattg agatattaca cctagtctgt	4200
ggcttgactg ttttctttat gtcttttgat gaatagaagt tttaaatttt gacaaggcca	4260
aatttatattt tttcttttggt ttgatatttt ttctctccaa ttttaaccca agatttcaga	4320
tattctgctc tattatataa actttatatt tttatatttg tgatctacct tgaattgata	4380
tgtatgttgt gaattatgga tcagggttct ttttttcccc catacaagta tccagtcatt	4440
gtaacactgt ttattgaaag aattatcctt tcctcattaa attaccttgc caattagtaa	4500
aaaatcaatt aaccataatg gtggatctgt ttctggactt tctgtttggt tacactgaaa	4560
tgtttgcca tccttgact cactcatacc atactgcctt gaattactgt agctgcatag	4620
atgctccta agttgggatt acattgtaat aaacgcaatg taagtt	4666

<210> 62

<211> 4449

<212> DNA

<213> Homo sapiens

<220>

<223> >NRAS|ENSG00000213281|ENST00000369535|4449

<400> 62

gaaacgtccc gtgtgggagg ggcgggtctg ggtgcggcct gccgcatgac tcgtggttcg	60
gaggccacg tggccggggc ggggactcag gcgcctgggg cgccgactga ttacgtagcg	120
ggcggggccg gaagtgccgc tccttggtgg gggctgttca tggcggttcc ggggtctcca	180
acatttttcc cggctgtggt cctaaatctg tccaaagcag aggcagtgga gcttgagggt	240
cttgctggtg tgaaatgact gagtacaaac tgggtggtggt tggagcaggt ggtgttgga	300
aaagcgact gacaatccag ctaatccaga accactttgt agatgaatat gatccacca	360
tagaggattc ttacagaaaa caagtgggta tagatggtga aacctgtttg ttggacatac	420
tggatacagc tggacaagaa gagtacagt ccatgagaga ccaatacatg aggacaggcg	480
aaggcttcct ctgtgtattt gccatcaata atagcaagtc atttgcggat attaacctct	540

acagggagca gattaagcga gtaaaagact cggatgatgt acctatggtg ctagtgggaa	600
acaagtgtga tttgccaaca aggacagttg atacaaaaca agcccacgaa ctggccaaga	660
gttacgggat tccattcatt gaaacctcag ccaagaccag acagggtgtt gaagatgctt	720
tttacacact ggtaagagaa atacgccagt accgaatgaa aaaactcaac agcagtgatg	780
atgggactca gggttgtatg ggattgccat gtgtggtgat gtaacaagat acttttaaag	840
ttttgtcaga aaagagccac tttcaagctg cactgacacc ctggtcctga cttccctgga	900
ggagaagtat tcctgttgct gtcttcagtc tcacagagaa gtcctgcta cttccccagc	960
tctcagtagt ttagtacaat aatctctatt tgagaagttc tcagaataac tacctcctca	1020
cttggctgtc tgaccagaga atgcacctct tgttactccc tgttatTTTT ctgccctggg	1080
ttcttcaca gcacaaacac acctctgcca cccaggttt ttcattctgaa aagcagttca	1140
tgtctgaaac agagaaccaa accgcaaacg tgaaattcta ttgaaaacag tgtcttgagc	1200
tctaaagtag caactgctgg tgattTTTT tttctTTTTa ctgttgaact tagaactatg	1260
ctaattTTTT gagaaatgtc ataaattact gttttgcaa gaatatagtt attattgctg	1320
tttggtttgt ttataatgtt atcggctcta ttctctaaac tggcatctgc tctagattca	1380
taaatacaaa aatgaatact gaattttgag tctatcctag tcttcacaac tttgacgtaa	1440
ttaaatccaa ctttcacagt gaagtgcctt tttcctagaa gtggtttgta gacttccttt	1500
ataatatttc agtggaatag atgtctcaaa aatccttatg catgaaatga atgtctgaga	1560
tacgtctgtg acttatctac cattgaagga aagctatatc tatttgagag cagatgccat	1620
tttgtacatg tatgaaattg gttttccaga ggcctgtttt ggggctttcc caggagaaag	1680
atgaaactga aagcacatga ataatttcac ttaataattt ttacctaatc tccacttttt	1740
tcataggtta ctacctatac aatgtatgta atttgtttcc cctagcttac tgataaacct	1800
aatattcaat gaacttccat ttgtattcaa atttgtgtca taccagaaag ctctacattt	1860
gcagatgttc aaatattgta aaactttggt gcattgttat ttaatagctg tgatcagtga	1920
ttttcaaacc tcaaatatag tatattaaca aattacattt tcaactgtata tcatggatatc	1980
ttaatgatgt atataattgc cttcaatccc cttctcacc caccctctac agcttcccc	2040
acagcaatag gggcttgatt atttcagttg agtaaagcat ggtgctaattg gaccagggtc	2100
acagtttcaa aacttgaaca atccagttag catcacagag aaagaaattc ttctgcattt	2160

gctcattgca ccagtaactc cagctagtaa ttttgctagg tagctgcagt tagccctgca	2220
aggaaagaag aggtcagtta gcacaaaccc tttaccatga ctggaaaact cagtatcacg	2280
tatttaaaca ttttttttct ttttagccat gtagaaactc taaattaagc caatattctc	2340
atttgagaat gaggatgtct cagctgagaa acgttttaaa ttctctttat tcataatgtt	2400
ctttgaaggg tttaaaacaa gatgttgata aatctaagct gatgagtttg ctcaaaacag	2460
gaagttgaaa ttgttgagac aggaatggaa aatataatta attgatacct atgaggattt	2520
ggaggcttgg cattttaatt tgcagataat accctggtaa ttctcatgaa aaatagactt	2580
ggataacttt tgataaaaaga ctaattccaa aatggccact ttgttcctgt ctttaatatc	2640
taaatactta ctgaggctct ccattcttcta tattatgaat tttcatttat taagcaaag	2700
tcataattacc ttgaaattca gaagagaaga aacatatact gtgtccagag tataatgaac	2760
ctgcagagtt gtgcttctta ctgctaattc tgggagcttt cacagtactg tcatcatttg	2820
taaatggaaa ttctgctttt ctgtttctgc tccttctgga gcagtgtac tctgtaattt	2880
tcctgaggct tatcacctca gtcatttctt ttttaaagt ctgtgactgg cagtgattct	2940
ttttcttaaa aatctattaa atttgatgtc aaattagga gaaagatagt tactcatctt	3000
gggctcttgt gccaatagcc cttgtatgta tgtacttaga gttttccaag tatgttctaa	3060
gcacagaagt ttctaaatgg ggccaaaatt cagacttgag tatgttcttt gaatacctta	3120
agaagttaca attagccggg catggtggcc cgtgcctgta gtcccagcta cttgagaggc	3180
tgaggcagga gaatcacttc aaccaggag gtggaggtta cagtgagcag agatcgtgcc	3240
actgcactcc agcctgggtg acaagagaga cttgtctcca aaaaaaagt tacacctagg	3300
tgtgaatttt ggcacaaagg agtgacaaac ttatagttaa aagctgaata acttcagtgt	3360
ggtataaaac gtggttttta ggctatgttt gtgattgctg aaaagaattc tagtttacct	3420
caaaatcctt ctctttcccc aaattaagt cctggccagc tgtcataaat tacatattcc	3480
ttttggtttt tttaaagggt acatgttcaa gagtgaaaat aagatgttct gtctgaaggc	3540
taccatgccg gatctgtaaa tgaacctgtt aaatgctgta tttgctcaa cggcttacta	3600
tagaatgtta cttaatataa tatcatactt attacaattt ttactatagg agtgtaatag	3660
gtaaaattaa tctctatttt agtgggcca tgtttagtct ttcaccatcc tttaaactgc	3720
tgtgaatttt tttgtcatga cttgaaagca aggatagaga aacactttag agatatgtgg	3780

ggtttttttta ccattccaga gcttgtgagc ataatcatat ttgctttata tttatagtca	3840
tgaactccta agttggcagc tacaaccaag aaccaaaaaa tgggtgcgttc tgcttcttgt	3900
aattcatctc tgctaataaa ttataagaag caaggaaaat tagggaaaat attttatttg	3960
gatggtttct ataaacaagg gactataatt cttgtacatt atttttcatc tttgctgttt	4020
ctttgagcag tctaattgtc cacacaatta tctaaggtat ttgttttcta taagaattgt	4080
tttaaaagta ttcttgttac cagagtagtt gtattatatt tcaaaacgta agatgatttt	4140
taaaagcctg agtactgacc taagatggaa ttgtatgaac tctgctctgg agggagggga	4200
ggatgtccgt ggaagttgta agacttttat ttttttgtgc catcaaata aggtaaaaat	4260
aattgtgcaa ttctgctgtt taaacaggaa ctattggcct ccttggccct aaatggaagg	4320
gccgatattt taagttgatt attttattgt aaattaatcc aacctagttc tttttaattt	4380
ggttgaatgt ttttcttgt taaatgatgt ttaaaaaata aaaactggaa gttcttggct	4440
tagtcataa	4449

<210> 63
 <211> 4003
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PALB2|ENSG00000083093|ENST00000261584|4003

<400> 63	
gccccgctgg gtcagctgat cgcgcactga ggggtgcgatc ccgggctccc cattccttcc	60
tggggcgcct ccccgccca gggccaactg ggtcccgggtg tcggcaggcc tggggtcggc	120
gacggctgct cttttcgttc tgtcgcctgc ccgatggacg agcctcccgg gaagcccctc	180
agctgtgagg agaaggaaaa gttaaaggag aaattagcat tcttgaaaag ggaatacagc	240
aagacactag cccgccttca gcgtgccccaa agagctgaaa agattaagca ttctattaag	300
aaaacagtag aagaacaaga ttgtttgtct cagcaggatc tctcaccgca gctaaaacac	360
tcagaacctt aaaataaaat atgtgtttat gacaagttac acatcaaaac ccatcttgat	420
gaagaaactg gagaaaagac atctatcaca cttgatgttg ggcctgagtc ctttaaccct	480
ggagatggcc caggaggatt acctatacaa agaacagatg acaccaaga acattttccc	540

cacaggggtca gtagaccctag tggtagagcaa aagcagaagc tgccaagcag aagaaagaag	600
cagcagaaga ggacatttat ttcacaggag agagactgtg tctttggcac tgattcactc	660
agattgtctg ggaaaagact aaaggaacag gaagaaatca gtagcaaaaa tcctgctaga	720
tcaccagtaa ctgaaataag aactcacctt ttaagtctta aatctgaact tccagattct	780
ccagaaccag ttacagaaat taatgaagac agtgtattaa ttccaccaac tgcccaacca	840
gaaaaaggtg ttgatacatt cctaagaaga cctaatttca ccagggcgac tacagttcct	900
ttacagactc tatcagatag cggtagtagt cagcaccttg aacacattcc tcctaaaggt	960
agcagtgaac ttactactca cgacctaaaa aacattagat ttacttcacc tgtaagtttg	1020
gaggcacaag gcaaaaaaat gactgtctct acagataacc tccttgtaaa taaagctata	1080
agtaaaagtg gccaaactgcc cacaagttct aatttagagg caaatatttc atgttctcta	1140
aatgaactca cctacaataa cttaccagca aatgaaaacc aaaacttaaa agaacaaaat	1200
caaacagaga aatctttaaa atctcccagt gacactcttg atggcaggaa tgaaaatctt	1260
caggaaagtg agattctaag tcaacctaag agtcttagcc tggaagcaac ctctcctctt	1320
tctgcagaaa aacattcttg cacagtgcct gaaggccttc tgtttcctgc agaataattat	1380
gttagaacia cacgaagcat gtccaattgc cagaggaaag tagccgtgga ggctgtcatt	1440
cagagtcatt tggatgtcaa gaaaaaaggg tttaaaaata aaaataagga tgcaagtaaa	1500
aatttaaacc tttccaatga ggaaactgac caaagtgaag ttaggatgtc tggcacatgc	1560
acaggacaac caagttcaag aacctctcag aaacttctct cattaactaa agtcagctct	1620
cccgctgggc ccactgaaga taatgacttg tctaggaagg cagttgcca agcacctggt	1680
agaagataca caggaaaaag aaaatcagcc tgcaccccag catcagatca ttgtgaacca	1740
cttttgccaa cttctagcct gtcgattgtt aacagggtcca aggaagaagt cacctcacac	1800
aaatatcagc acgaaaaatt atttattcaa gtgaaaggga agaaaagtcg tcatcaaaaa	1860
gaggattccc tttcttgag taatagtgtt tatttatcct tggatgatga tgctttcacg	1920
gctccatttc atagggatgg aatgctgagt ttaaagcaac tactgtcttt tctcagtatc	1980
acagactttc agttacctga tgaagacttt ggacctctta agcttgaaaa agtgaagtcc	2040
tgctcagaaa aaccagtgga gccctttgag tcaaaaatgt ttggagagag acatcttaaa	2100
gagggaagct gtatttttcc agaggaactg agtcctaaac gcatggatac agaaatggag	2160

gacttagaag aggaccttat tgttctacca ggaaaatcac atcccaaaag gccaaactcg	2220
caaagccagc atacaaagac gggcctttct tcatccatat tactttatac tcctttaaat	2280
acggttgcg ctagatgataa tgacaggcct accacagaca tgtgttcacc tgctttcccc	2340
atcttaggta ctactccagc ctttggccct caaggctcct atgaaaaagc atctacagaa	2400
gttgctggac gaacttgctg cacaccccaa cttgctcatt tgaaagactc agtctgtctt	2460
gccagtata ctaaacaatt cgacagttca ggcagcccag caaaaccaca taccaccctg	2520
caagtgtcag gcaggcaagg acaacctacc tgtgactgtg actctgtccc gccaggaaca	2580
cctccacca ttgagtcatt cactttttaa gaaaatcagc tctgtagaaa cacatgccag	2640
gagctgcata aacattccgt cgaacagact gaaacagcag agcttcctgc ttctgatagc	2700
ataaaccag gcaacctaca attggtttca gagttaaaga atccttcagg ttcctgttcc	2760
gtagatgtga gtgccatgtt ttgggaaaga gccggttgta aagagccatg tatcataact	2820
gcttgcaag atgtagtttc tctttgaaa gctctggatg cttggcagtg ggaaaaactt	2880
tatactggc acttcgcaga ggttcagta ttacagatag ttccagtgcc tgatgtgtat	2940
aatctcgtgt gtgtagcttt gggaaatttg gaaatcagag agatcagggc attgttttgt	3000
tcctctgatg atgaaagtga aaagcaagta ctactgaagt ctggaaatat aaaagctgtg	3060
cttggcctga caaagaggag gctagttagt agcagtggga ccctttctga tcaacaagta	3120
gaagtcatga cgtttgcaga agatggagga ggcaaagaaa accaattttt gatgccccct	3180
gaggagacta tactaacttt tgctgaggtc caagggatgc aagaagctct gcttggtact	3240
actattatga acaacattgt tatttggaat ttaaaaactg gtcaactcct gaaaaagatg	3300
cacattgatg attcttacca agcttcagtc tgtcacaaag cctattctga aatggggctt	3360
ctctttattg tcctgagtca tccctgtgcc aaagagagtg agtcgttgcg aagccctgtg	3420
tttcagctca ttgtgattaa ccctaagacg actctcagcg tgggtgtgat gctgtactgt	3480
cttcctccag ggcaggctgg caggttcctg gaaggtgacg tgaaagatca ctgtgcagca	3540
gcaatcttga cttctggaac aattgccatt tgggacttac ttctcggta gtgtactgcc	3600
ctcctccac ctgtctctga ccaacattgg tcttttgtga aatggtcggg tacagactct	3660
catttgctgg ctggacaaaa agatggaaat atatttgtat accactattc ataagttagg	3720
gtaaagtga aacacaattt tctggatata ttgggcctct tagtattttt tggagtttta	3780

aatataaagg agaatatctg aatgacactt aaaatgattg cttgtttatg tccagacaga	3840
cttattttttt attctaataga tggtagcacc actgatcttg gatgtacatt tatgtataact	3900
ttgagaaaaa gggtttttagg ttgatttttg taatttccca catttgtaca tgtgctttta	3960
aagggtgtaca taaagcttca aatggcaata aatattttatt ttt	4003

<210> 64
 <211> 3958
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PARP1|ENSG00000143799|ENST00000366794|3958

<400> 64	
cggtaggcccgg tgcggcggtg tcggtaggccc ctctggccgc tcaggcgccct gcggctgggt	60
gagcgcacgc gaggcggcga ggcggcagcg tgtttctagg tcgtggcgctc gggcttccgg	120
agctttggcg gcagctaggg gaggatggcg gagtcttcgg ataagctcta tcgagtcgag	180
tacgccaaaga gcgggcgcgc ctcttgcaag aaatgcagcg agagcatccc caaggactcg	240
ctccggatgg ccatcatggt gcagtcgccc atgtttgatg gaaaagtccc aactggtac	300
cacttctcct gcttctggaa ggtgggccac tccatccggc accctgacgt tgaggtggat	360
gggttctctg agcttcggtg ggatgaccag cagaaagtca agaagacagc ggaagctgga	420
ggagtgcagc gcaaaggcca ggatggaatt ggtagcaagg cagagaagac tctgggtgac	480
tttgacgagc agtatgccaa gtccaacaga agtacgtgca aggggtgtat ggagaagata	540
gaaaagggcc aggtgcccct gtccaagaag atggtggacc cggagaagcc acagctaggc	600
atgattgacc gctggtacca tccaggctgc tttgtcaaga acagggagga gctgggtttc	660
cggcccagat acagtgcgag tcagctcaag ggcttcagcc tccttgctac agaggataaa	720
gaagccctga agaagcagct cccaggagtc aagagtgaag gaaagagaaa aggcgatgag	780
gtggatggag tggatgaagt ggcgaagaag aaatctaaaa aagaaaaaga caaggatagt	840
aagcttgaag aagccctaaa ggctcagaac gacctgatct ggaacatcaa ggacgagcta	900
aagaaagtgt gttcaactaa tgacctgaag gagctactca tcttcaacaa gcagcaagtg	960
ccttctgggg agtcggcgat cttggaccga gtagctgatg gcatggtgtt cggtgccctc	1020
cttcctgctg aggaatgctc gggtcagctg gtcttcaaga gcgatgccta ttactgcact	1080

ggggacgtca ctgcctggac caagtgtatg gtcaagacac agacacccaa ccggaaggag	1140
tgggtaaccc caaaggaatt ccgagaaatc tcttacctca agaaattgaa ggtaaataaaa	1200
caggaccgta tattcccccc agaaaccagc gcctccgtgg cggccacgcc tccgcctcc	1260
acagcctcgg ctctgtctgc tgtgaactcc tctgcttcag cagataagcc attatccaac	1320
atgaagatcc tgactctcgg gaagctgtcc cggaacaagg atgaagtga ggccatgatt	1380
gagaaactcg gggggaagtt gacggggacg gccacaagg cttccctgtg catcagcacc	1440
aaaaaggagg tggaaaagat gaataagaag atggaggaag taaaggaagc caacatccga	1500
gttgtgtctg aggacttcct ccaggacgtc tccgcctcca ccaagagcct tcaggagtgtg	1560
ttcttagcgc acatcttgct cccttggggg gcagaggtga aggcagagcc tgttgaagtt	1620
gtggcccaa gagggaagtc aggggctgctg ctctccaaa aaagcaaggg ccaggtcaag	1680
gaggaaggta tcaacaaatc tgaaaagaga atgaaattaa ctcttaaagg aggagcagct	1740
gtggatcctg attctggact ggaacactct gcgcatgtcc tggagaaagg tgggaaggtc	1800
ttcagtcca cccttggcct ggtggacatc gttaaaggaa ccaactccta ctacaagctg	1860
cagcttctgg aggacgaca ggaaaacagg tattggatat tcaggctctg gggccgtgtg	1920
ggtacgggta tcggtagcaa caaactggaa cagatgccgt ccaaggagga tgccattgag	1980
cacttcatga aattatatga agaaaaaac gggaacgctt ggcactcaa aaatttcacg	2040
aagtatccca aaaagttcta cccctggag attgactatg gccaggatga agaggcagtg	2100
aagaagctga cagtaaatcc tggcaccaag tccaagctcc ccaagccagt tcaggacctc	2160
atcaagatga tctttgatgt ggaaagtatg aagaaagcca tgggtggagta tgagatcgac	2220
cttcagaaga tgcccttggg gaagctgagc aaaaggcaga tccaggccgc atactccatc	2280
ctcagtgagg tccagcaggc ggtgtctcag ggcagcagcg actctcagat cctggatctc	2340
tcaaactgct ttacaccct gatccccac gactttggga tgaagaagcc tccgctcctg	2400
aacaatgcag acagtgtgca ggccaagggtg gaaatgcttg acaacctgct ggacatcgag	2460
gtggcctaca gtctgctcag gggaggggtct gatgatagca gcaaggatcc catcgatgtc	2520
aactatgaga agctcaaac tgacattaag gtggttgaca gagattctga agaagccgag	2580
atcatcagga agtatgttaa gaacactcat gcaaccacac acaatgcgta tgacttgga	2640
gtcatcgata tctttaagat agagcgtgaa ggcgaatgcc agcgttaca gccctttaag	2700

cagcttcata accgaagatt gctgtggcac ggggccagga ccaccaactt tgctgggatc	2760
ctgtcccagg gtcttcggat agccccgcct gaagcgcccg tgacaggcta catgttttgt	2820
aaagggatct atttcgctga catgggtctcc aagagtgcc aactactgcca tacgtctcag	2880
ggagacccaa taggcttaat cctgttggga gaagttgccc ttggaaacat gtatgaactg	2940
aagcacgctt cacatatcag caagttaccc aagggaagc acagtgtcaa aggtttgggc	3000
aaaactaccc ctgatccttc agctaacatt agtctggatg gtgtagacgt tcctcttggg	3060
accgggattt catctgggtg gaatgacacc tctctactat ataacgagta cattgtctat	3120
gatattgctc aggtaaatct gaagtatctg ctgaaactga aattcaattt taagacctcc	3180
ctgtggtaat tgggagagg agccgagtca caccgggtgg ctctggtagt aattcacccg	3240
aagcgcttct gcaccaactc acctggccgc taagttgctg atgggtagta cctgtactaa	3300
accacctcag aaaggatttt acagaaacgt gttaaagggt ttctctaact tctcaagtcc	3360
cttgttttgt gttgtgtctg tggggagggg ttgttttggg gttgtttttg ttttttcttg	3420
ccaggtagat aaaactgaca tagagaaaag gctggagaga gattctgttg catagactag	3480
tcctatggaa aaaaccaagc ttcgttagaa tgtctgcctt actggtttcc ccagggaagg	3540
aaaaatacac ttccaccctt ttttctaagt gttcgtcttt agttttgatt ttggaaagat	3600
gttaagcatt tatttttagt taaaaataaa aactaatttc atactattta gattttcttt	3660
tttatcttgc acttattgtc ccctttttag ttttttttgt ttgcctcttg tggtagggg	3720
tgtgggaaga ccaaaggaag gaacgctaac aatttctcat acttagaaac aaaaagagct	3780
ttccttctcc aggaatactg aacatgggag ctcttgaaat atgtagtatt aaaagttgca	3840
tttgaaattc ttgactttct tatgggcact tttgtcttcc aaattaaaac tctaccacaa	3900
atatacttac ccaagggcta atagtaatac tcgattaaaa atgcagatgc cttctcta	3958

<210> 65

<211> 1886

<212> DNA

<213> Homo sapiens

<220>

<223> >PARP2|ENSG00000129484|ENST00000250416|1886

<400> 65

ggttgatgac gtcagcgttc gaattccatg gcggcgcggc ggcgacggag caccggcggc	60
ggcagggcgga gagcattaaa tgaaagcaaa agagttaata atggcaacac ggctccagaa	120
gactcttccc ctgccaagaa aactcgtaga tgccagagac aggagtcgaa aaagatgcct	180
gtggctggag gaaaagctaa taaggacagg acagaagaca agcaagatgg tatgccagga	240
aggtcatggg ccagcaaaag ggtctctgaa tctgtgaagg ccttgctgtt aaagggcaaa	300
gctcctgtgg acccagagtg tacagccaag gtggggaagg ctcatgtgta ttgtgaagga	360
aatgatgtct atgatgtcat gctaaatcag accaatctcc agttcaacaa caacaagtac	420
tatctgattc agctattaga agatgatgcc cagaggaact tcagtgtttg gatgagatgg	480
ggccgagttg ggaaaatggg acagcacagc ctggtggctt gttcaggcaa tctcaacaag	540
gccaaggaaa tctttcagaa gaaattcctt gacaaaacga aaaacaattg ggaagatcga	600
gaaaagtttg agaagggtgcc tggaaaatat gatatgctac agatggacta tgccaccaat	660
actcaggatg aagaggaaac aaagaaagag gaatctctta aatctccctt gaagccagag	720
tcacagctag atcttcgggt acaggagtta ataaagttga tctgtaatgt tcaggccatg	780
gaagaaatga tgatggaaat gaagtataat accaagaaag cccacttgg gaagctgaca	840
gtggcacaaa tcaaggcagg ttaccagtct cttaagaaga ttgaggattg tattcgggct	900
ggccagcatg gacgagctct catggaagca tgcaatgaat tctacaccag gattccgcat	960
gactttggac tccgtactcc tccactaatc cggacacaga aggaactgtc agaaaaaata	1020
caattactag aggctttggg agacattgaa attgctatta agctggtgaa aacagagcta	1080
caaagcccag aacacccatt ggaccaacac tatagaaacc tacattgtgc cttgcgcccc	1140
cttgaccatg aaagttatga gttcaaagtg atttcccagt acctacaatc taccatgct	1200
cccacacaca gcgactatac catgaccttg ctggatttgt ttgaagtgga gaaggatggt	1260
gagaaagaag cttcagaga ggaccttcat aacaggatgc ttctatggca tggttccagg	1320
atgagtaact ggggtgggaat cttgagccat gggcttcgaa ttgccccacc tgaagctccc	1380
atcacagggtt acatgtttgg gaaaggaatc tactttgctg acatgtcttc caagagtgcc	1440
aattactgct ttgcctctcg cctaaagaat acaggactgc tgctcttatc agaggtagct	1500
ctaggtcagt gtaatgaact actagaggcc aatcctaagg ccgaaggatt gcttcaaggt	1560
aaacatagca ccaaggggct gggcaagatg gctcccagtt ctgcccactt cgtcacctg	1620

aatgggagta cagtgccatt aggaccagca agtgacacag gaattctgaa tccagatggt	1680
tataccctca actacaatga atatattgta tataacccca accaggtccg tatgcggtac	1740
cttttaaagg ttcagtttaa tttccttcag ctgtgggtgaa tgttgatatt aaataaacca	1800
gagatctgat cttcaagcaa gaaaataagc agtggtgtac ttgtgaattt tgtgatattt	1860
tatgtaataa aaactgtaca ggtcta	1886

<210> 66
 <211> 2336
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PARP3|ENSG0000041880|ENST00000398755|2336

<400> 66	
ggtcaccgcg cgaccggcag atgcgtgctg caggccccgg ccacatgagc agcgctacgg	60
acgcgactgc cccggccttg gatatgccag atcgagtgtc caccgtccg tgggactggt	120
cgcctgactc ggcctgcccc agcctctgct tcaccccact ggtggccaaa tagccgatgt	180
ctaatacccc acacaagctc atccccggcc tctggcgatt gttgggaatt ctctccctaa	240
ttcacgcctg aggctcatgg agagttgcta gacctgggac tgcctggga ggcgcacaca	300
accaggccgg gtggcagcca ggacctctcc catgtccctg cttttcttgg ccatggctcc	360
aaagccgaag ccctgggtac agactgaggg ccctgagaag aagaagggcc ggcaggcagg	420
aaggaggag gaccccttc gctccaccgc tgaggccctc aaggccatac ccgcagagaa	480
gcgcataatc cgcgtggatc caacatgtcc actcagcagc aaccccgga cccaggtgta	540
tgaggactac aactgcaccc tgaaccagac caacatcgag aacaacaaca acaagttcta	600
catcatccag ctgctccaag acagcaaccg cttcttcacc tgctggaacc actggggccg	660
tgtgggagag gtcggccagt caaagatcaa ccacttcaca aggctagaag atgcaaagaa	720
ggactttgag aagaaatttc gggaaaagac caagaacaac tgggcagagc gggaccactt	780
tgtgtctcac ccgggcaagt acacacttat cgaagtacag gcagaggatg aggcccagga	840
agctgtggtg aaggtggaca gaggcccagt gaggactgtg actaagcggg tgcagccctg	900
ctccctggac ccagccacgc agaagctcat cactaacatc ttcagcaagg agatgttcaa	960
gaacaccatg gccctcatgg acctggatgt gaagaagatg cccctgggaa agctgagcaa	1020

gcaacagatt gcacggggtt tcgaggcctt ggaggcgctg gaggaggccc tgaaaggccc	1080
cacggatggt ggccaaagcc tggaggagct gtcctcacac ttttacaccg tcatcccga	1140
caacttcggc cacagccagc ccccgcccat caattcccct gagcttctgc aggccaagaa	1200
ggacatgctg ctggtgctgg cggacatcga gctggcccag gccctgcagg cagtctctga	1260
gcaggagaag acggtggagg aggtgccaca cccctggac cgagactacc agcttctcaa	1320
gtgccagctg cagctgctag actctggagc acctgagtac aaggtgatac agacctactt	1380
agaacagact ggcagcaacc acaggtgccc tacacttcaa cacatctgga aagtaaacca	1440
agaaggggag gaagacagat tccaggccca ctccaaactg ggtaatcgga agctgctgtg	1500
gcatggcacc aacatggccg tgggtggccgc catcctcact agtgggctcc gcatcatgcc	1560
acattctggt gggcgtgttg gcaagggcat ctactttgcc tcagagaaca gcaagtcagc	1620
tggatatgtt attggcatga agtgtggggc ccaccatgtc ggctacatgt tcctgggtga	1680
ggtggccctg ggcagagagc accatatcaa cacggacaac cccagcttga agagcccacc	1740
tcctggcttc gacagtgtca ttgcccagg ccacaccgag cctgatccga cccaggacac	1800
tgagttggag ctggatggcc agcaagtggg ggtgccccag ggccagcctg tgccctgccc	1860
agagttcagc agctccacat tctcccagag cgagtacctc atctaccagg agagccagtg	1920
tcgcctgcgc tacctgctgg aggtccacct ctgagtgccc gccctgtccc ccggggtcct	1980
gcaaggctgg actgtgatct tcaatcatcc tgcccattct tggtagccct atatcactcc	2040
tttttttcaa gaatacaata cgttggtgtt aactatagtc accatgctgt acaagatccc	2100
tgaacttatg cctcctaact gaaattttgt attctttgac acatctgccc agtccctctc	2160
ctcccagccc atggtaacca gcatttgact ctttacttgt ataaggcag cttttatagg	2220
ttccacatgt aagtgagatc atgcagtgtt tgtctttctg tgcctggctt atttcactca	2280
gcataatgtg caccgggttc acccatgttt tcataaatga caagatttcc tccttt	2336

<210> 67

<211> 5474

<212> DNA

<213> Homo sapiens

<220>

<223> >PARP4|ENSG00000102699|ENST00000381989|5474

<400> 67

cgcccgccca gccccggggg caggagagc ctagattacg gaagtaccgc gagcaaggag	60
cgcggaatcg gggagcgtcc ggagctagct ggatcctcta ggcaggatgg tgatgggaat	120
ctttgcaaat tgtatcttct gtttgaaagt gaagtactta cctcagcagc agaagaaaaa	180
gctacaaact gacattaagg aaaatggcgg aaagttttcc ttttcgttaa atcctcagtg	240
cacacatata atcttagata atgctgatgt tctgagtcag taccaactga attctatcca	300
aaagaaccac gttcatattg caaacccaga ttttatatgg aaatctatca gggaaaagag	360
actcttgat gtaaagaatt atgatcctta taagcccctg gacatcacac cacctcctga	420
tcagaaggcg agcagttctg aagtgaaaac agaaggtcta tgcccggaca gtgccacaga	480
ggaggaagac actgtggaac tctactgagtt tggatgcag aatgttgaaa ttcctcatct	540
tcctcaagat tttgaagttg caaaatataa caccttgag aaagtgggaa tggagggagg	600
ccaggaagct gtggtggtgg agcttcagtg ttcgcgggac tccagggact gtcctttcct	660
gatatcctca cacttcctcc tggatgatgg catggagact agaagacagt ttgctataaa	720
gaaaacctct gaagatgcaa gtgaatactt tgaaaattac attgaagaac tgaagaaaca	780
aggatttcta ctaagagaac atttcacacc tgaagcaacc caattagcat ctgaacaatt	840
gcaagcattg cttttggagg aagtcatgaa ttcaagcact ctgagccaag aggtgagcga	900
tttagtagag atgatttggg cagaggccct gggccacctg gaacacatgc ttctcaagcc	960
agtgaacagg attagcctca acgatgtgag caaggcagag gggatttctc ttctagtaaa	1020
ggcagcactg aaaaatggag aaacagcaga gcaattgcaa aagatgatga cagagtttta	1080
cagactgata cctcaciaag gcacaatgcc caaagaagtg aacctgggac tattggctaa	1140
gaaagcagac ctctgccagc taataagaga catggttaat gtctgtgaaa ctaatttgtc	1200
caaaccaac ccaccatccc tggccaaata ccgagctttg aggtgcaaaa ttgagcatgt	1260
tgaacagaat actgaagaat ttctcagggt tagaaaagag gttttgcaga atcatcacag	1320
taagagccca gtggatgtct tgcagatatt tagagttggc agagtgaatg aaaccacaga	1380
gtttttgagc aaacttggtg atgtgaggcc cttgttgcat gggttctcctg taaaaacat	1440
cgtgggaatc ttgtgtcgag ggttgctttt acccaaagta gtggaagatc gtggtgtgca	1500
aagaacagac gtcggaaacc ttggaagtgg gatttatttc agtgattcgc tcagtacaag	1560

tatcaagtac tcacacccgg gagagacaga tggcaccaga ctcttgctca tttgtgacgt	1620
agccctcggg aagtgtatgg acttacatga gaaggacttt tccttaactg aagcaccacc	1680
aggctacgac agtgtgcatg gagtttcgca aacagcctct gtcaccacag actttgagga	1740
tgatgaattt gttgtctata aaaccaatca ggttaaaatg aaatatatta ttaaattttc	1800
catgcctgga gatcagataa aggactttca tcctagtgat catactgaat tagaggaata	1860
cagacctgag ttttcaaatt tttcaaaggt tgaagattac cagttaccag atgccaaaac	1920
ttccagcagc accaaggccg gcctccagga tgcctctggg aacttggttc ctctggagga	1980
tgtccacatc aaaggagaa tcatagacac tgtagcccag gtcattgttt ttcagacata	2040
cacaaataaa agtcacgtgc ccattgaggc aaaatatatc tttcctttgg atgacaaggc	2100
cgctgtgtgt ggcttcgaag cttcatcaa tgggaagcac atagttggag agattaaaga	2160
gaaggaagaa gccagcaag agtacctaga agccgtgacc cagggccatg gcgcttacct	2220
gatgagtcag gatgctccgg acgtttttac tgtaagtgtt ggaaacttac cccctaaggc	2280
taaggttctt ataaaaatta cctacatcac agaactcagc atcctgggca ctgttggtgt	2340
ctttttcatg cccgccaccg tagcaccctg gcaacaggac aaggctttga atgaaaacct	2400
tcaggataca gtagagaaga tttgtataaa agaaatagga acaaagcaaa gcttctcttt	2460
gactatgtct attgagatgc cgtatgtgat tgaattcatt ttcagtgata cacatgaact	2520
gaaacaaaag cgcacagact gcaaagctgt cattagcacc atggaaggca gctccttaga	2580
cagcagtgga ttttctctcc acatcggttt gtctgctgcc tatctcccaa gaatgtgggt	2640
tgaaaaacat ccagaaaaag aaagcgaggc ttgcatgctt gtctttcaac ccgatctcga	2700
tgtcgacctc cctgacctag ccagtgagag cgaagtgatt atttgtcttg actgctccag	2760
ttccatggag ggtgtgacat tcttgcaagc caagcaaatc gccttgcatg cgctgtcctt	2820
ggtgggtgag aagcagaaag taaatattat ccagttcggc acaggttaca aggagctatt	2880
ttcgtatcct aagcatatca caagcaatac catggcagca gagttcatca tgtctgccac	2940
acctaccatg gggaacacag acttctggaa aacactccga tatcttagct tattgtaccc	3000
tgctcgaggg tcacggaaca tcctcctggg gtctgatggg cacctccagg atgagagcct	3060
gacattacag ctctgaaga ggagccgcc gcacaccagg ttattcgcct gcggtatcgg	3120
ttctacagca aatcgtcacg tcttaaggat tttgtcccag tgtggtgccg gagtatttga	3180

atattttaat gcaaaatcca agcatagttg gagaaaacag atagaagacc aaatgaccag	3240
gctatgttct ccgagttgcc actctgtctc cgtcaaattgg cagcaactca atccagatgt	3300
gcccagggcc ctgcaggccc cagcccaggt gccgtccttg tttctcaatg atcgactcct	3360
tgtctatgga ttcattcctc actgcacaca ggcaactctg tgtgcactaa ttcaagagaa	3420
agaatttcgt acaatggtgt cgactactga gcttcagaag acaactggaa ctatgatcca	3480
caagctggca gcccagctc taatcagaga ttatgaagat ggcatcttc acgaaaatga	3540
aaccagtcac gagatgaaaa aacaaacctt gaaatctctg attattaaac tcagtaaaga	3600
aaactctctc ataacacaat ttacaagctt tgtggcagtt gagaaaagg atgagaatga	3660
gtcgctttt cctgatattc caaaagtttc tgaacttatt gccaaagaag atgtagactt	3720
cctgccctac atgagctggc agggggagcc ccaagaagcc gtcaggaacc agtctctttt	3780
agcatcctct gagtggccag aattacgttt atccaaacga aaacatagga aaattccatt	3840
ttccaaaaga aaaatggaat tatctcagcc agaagtttct gaagattttg aagaggatgg	3900
cttaggtgta ctaccagctt tcacatcaaa tttggaacgt ggaggtgtgg aaaagctatt	3960
ggatttaagt tggacagagt catgtaaacc aacagcaact gaaccactat ttaagaaagt	4020
cagtccatgg gaaacatcta cttctagctt ttttcctatt ttggctccgg ccgttggttc	4080
ctatcttccc ccgactgccc gcgctcacag tcctgcttcc ttgtcttttg cctcatatcg	4140
tcaggtagct agtttcggtt cagctgctcc tcccagacag tttgatgcat ctcaattcag	4200
ccaaggccct gtgcctggca cttgtgctga ctggatccca cagtcggcgt cttgtcccac	4260
aggacctccc cagaaccac cttcttcacc ctattgtggc attgtttttt caggagctc	4320
attaagctct gcacagtctg ctccactgca acatcctgga ggctttacta ccaggccttc	4380
tgctggcacc ttccctgagc tggattctcc ccagcttcac ttctctcttc ctacagaccc	4440
tgatcccatc agaggttttg ggtcttatca tccctctgct tcctctcctt ttcattttca	4500
accttccgca gcctctttga ctgccaacct taggctgcca atggcctctg ctttacctga	4560
ggctctttgc agtcagtccc ggactacccc agtagatctc tgtcttctag aagaatcagt	4620
aggcagtctc gaaggaagtc gatgtcctgt ctttgctttt caaagttctg acacagaaag	4680
tgatgagcta tcagaagtac ttcaagacag ctgcttttta caaataaaat gtgatacaaa	4740
agatgacagt atcctgtgct ttctggaagt aaaagaagag gatgaaatag tgtgcataca	4800

acactggcag gatgctgtgc cttggacaga actcctcagt ctacagacag aggatggctt	4860
ctggaaactt acaccagaac tgggacttat attaaatctt aatacaaag gtttgcacag	4920
ctttcttaaa caaaaaggca ttcaatctct aggtgtaaaa ggaagagaat gtctcctgga	4980
cctaattgcc acaatgctgg tactacagtt tattcgcacc aggttggaag aagagggaat	5040
agtgttcaaa tcaatgatga aaatggatga cgcttctatt tccaggaata ttccctgggc	5100
ttttgaggca ataaagcaag caagtgaatg ggtaagaaga actgaaggac agtaccatc	5160
tatctgcca cggcttgaac tggggaacga ctgggactct gccaccaagc agttgctggg	5220
actccagccc ataagcactg tgtcccctct tcatagagtc ctccattaca gtcaaggcta	5280
agtcaaagta aactgaattt taaacttttt gcatgcttct atgtagaaaa taatcaaag	5340
ataatagata cttataatga aacttcatta aggtttcatt cagtgtagca attactgtct	5400
ttaaaaatta agtggaagaa gaattacttt aatcaactaa caagcaataa taaaatgaaa	5460
cttaaaatat ttca	5474

<210> 68
 <211> 1359
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PCNA|ENSG00000132646|ENST00000379160|1359

<400> 68	
ccgcggatgg ccggagctgg cgccctgggt ctggaggtaa ccggttactg agggcgagaa	60
gcgccacccg gaggtcttag cctgacaaat gcttgctgac ctgggccaga gctcttcct	120
tacgcaagtc tcagccggtc gtcgcgacgt tcgcccgtc gctctgaggc tcctgaagcc	180
gaaaccagct agactttcct cttcccgcc tgcctgtagc ggcgttggtg ccactccgcc	240
accatgttcg aggcgcgcct ggtccagggc tccatcctca agaaggtgtt ggaggcactc	300
aaggacctca tcaacgaggc ctgctgggat attagctcca gcggtgtaaa cctgcagagc	360
atggactcgt cccacgtctc tttggtgcag ctcaccctgc ggtctgaggg cttcgacacc	420
taccgctgcg accgcaacct ggccatgggc gtgaacctca ccagtatgtc caaaatacta	480
aaatgcgccg gcaatgaaga tatcattaca ctaagggccg aagataacgc ggataccttg	540
gcgctagtat ttgaagcacc aaaccaggag aaagtttcag actatgaaat gaagttgatg	600

gatttagatg ttgaacaact tggaattcca gaacaggagt acagctgtgt agtaaagatg	660
ccttctggtg aatttgcacg tatatgccga gatctcagcc atattggaga tgctgttgta	720
atttcctgtg caaaagacgg agtgaaattt tctgcaagtg gagaacttgg aaatggaaac	780
attaaattgt cacagacaag taatgtcgat aaagaggagg aagctgttac catagagatg	840
aatgaaccag ttcaactaac ttttgcactg aggtacctga acttctttac aaaagccact	900
ccactctctt caacggtgac actcagtatg tctgcagatg tacccttgt tgtagagtat	960
aaaattgcgg atatgggaca cttaaaatac tacttggctc ccaagatcga ggatgaagaa	1020
ggatccttagg cattcttaaa attcaagaaa ataaaactaa gctctttgag aactgcttct	1080
aagatgccag catatactga agtcttttct gtcaccaaatt ttgtacctct aagtacatat	1140
gtagatattg ttttctgtaa ataacctatt ttttctcta ttctctgcaa tttgtttaaa	1200
gaataaagtc caaagtcaga tctggtctag ttaacctaga agtatttttg tctcttagaa	1260
atacttgtga tttttataat acaaaagggt cttgactcta aatgcagttt taagaattgt	1320
ttttgaattt aaataaagtt acttgaattt caaacatca	1359

<210> 69

<211> 9093

<212> DNA

<213> Homo sapiens

<220>

<223> >PIK3CA|ENSG00000121879|ENST00000263967|9093

<400> 69

tctccctcgg cgccgccgcc gccgcccgcg gggctgggac ccgatgcggt tagagccgcg	60
gagcctggaa gagccccgag cgtttctgct ttgggacaac catacatcta attccttaaa	120
gtagttttat atgtaaaact tgcaaagaat cagaacaatg cctccacgac catcatcagg	180
tgaactgtgg ggcattccact tgatgcccc aagaatccta gtagaatgtt tactaccaaa	240
tggaatgata gtgacttttag aatgcctccg tgaggctaca ttaataacca taaagcatga	300
actattttaa gaagcaagaa aataccccct ccatcaactt cttcaagatg aatcttctta	360
cattttcgta agtgttactc aagaagcaga aagggaagaa ttttttgatg aaacaagacg	420
actttgtgac cttcggcttt ttcaaccctt tttaaaagta attgaaccag taggcaaccg	480

tgaagaaaag atcctcaatc gagaaattgg ttttgctatc ggcatgccag tgtgtgaatt	540
tgatatgggt aaagatccag aagtacagga cttccgaaga aatattctga acgtttgtaa	600
agaagctgtg gatcttaggg acctcaattc acctcatagt agagcaatgt atgtctatcc	660
tccaaatgta gaatcttcac cagaattgcc aaagcacata tataataaat tagataaagg	720
gcaaataata gtggatgatct gggtaatagt ttctccaaat aatgacaagc agaagtatac	780
tctgaaaatc aacctgact gtgtaccaga acaagtaatt gctgaagcaa tcaggaaaaa	840
aactcgaagt atgttgctat cctctgaaca actaaaactc tgtgttttag aatatcaggg	900
caagtatatt ttaaaagtgt gtggatgtga tgaatacttc ctagaaaaat atcctctgag	960
tcagtataag tatataagaa gctgtataat gcttgggagg atgcccaatt tgatgttgat	1020
ggctaaagaa agcctttatt ctcaactgcc aatggactgt ttacaatgc catcttattc	1080
cagacgcatt tccacagcta caccatatat gaatggagaa acatctacaa aatcccttg	1140
ggttataaat agtgcactca gaataaaaaat tctttgtgca acctacgtga atgtaaatat	1200
tcgagacatt gataagatct atgttcgaac aggtatctac catggaggag aacccttatg	1260
tgacaatgtg aacactcaaa gagtaccttg ttccaatccc aggtggaatg aatggctgaa	1320
ttatgatata tacattcctg atcttcctcg tgctgctcga ctttgccttt ccatttgctc	1380
tgttaaaggc cgaaagggtg cttaaagagga acactgtcca ttggcatggg gaaatataaa	1440
cttgtttgat tacacagaca ctctagtatc tggaaaaatg gctttgaatc tttggccagt	1500
acctcatgga ttagaagatt tgctgaaccc tattgggtgtt actggatcaa atccaaataa	1560
agaaactcca tgcttagagt tggagtttga ctggttcagc agtgttggtaa agttcccaga	1620
tatgtcagtg attgaagagc atgccaattg gtctgtatcc cgagaagcag gatttagcta	1680
ttcccacgca ggactgagta acagactagc tagagacaat gaattaaggg aaaatgacaa	1740
agaacagctc aaagcaattt ctacacgaga tcctctctct gaaatcactg agcaggagaa	1800
agattttcta tggagtcaca gacactattg tgtaactatc cccgaaattc tacccaaatt	1860
gcttctgtct gttaaattgga attctagaga tgaagtagcc cagatgtatt gcttggtaaa	1920
agattggcct ccaatcaaac ctgaacaggc tatggaactt ctggactgta attaccaga	1980
tcctatgggt cgaggttttg ctgttcggtg cttggaaaaa tatttaacag atgacaaact	2040
ttctcagtat ttaattcagc tagtacaggt ctaaaaatat gaacaatatt tggataactt	2100

gcttgtgaga tttttactga agaaagcatt gactaatcaa aggattgggc actttttctt	2160
ttggcattta aaatctgaga tgcacaataa aacagttagc cagaggtttg gcctgctttt	2220
ggagtcctat tgtcgtgcat gtgggatgta tttgaagcac ctgaataggc aagtcgaggc	2280
aatggaaaag ctcatctaact taactgacat tctcaaacag gagaagaagg atgaaacaca	2340
aaaggtacag atgaagtttt tagttgagca aatgaggcga ccagatttca tggatgctct	2400
acagggcctt ctgtctcctc taaaccctgc tcatcaacta ggaaacctca ggcttgaaga	2460
gtgtcgaatt atgtcctctg caaaaaggcc actgtggttg aattgggaga acccagacat	2520
catgtcagag ttactgtttc agaacaatga gatcatcttt aaaaatgggg atgatttacg	2580
gcaagatatg ctaacacttc aaattattcg tattatggaa aatatctggc aaaatcaagg	2640
tcttgatctt cgaatgttac cttatggttg tctgtcaatc ggtgactgtg tgggacttat	2700
tgaggtgggtg cgaaattctc acactattat gcaaattcag tgcaaaggcg gcttgaaagg	2760
tgactgcag ttcaacagcc acacactaca tcagtggctc aaagacaaga acaaaggaga	2820
aatatatgat gcagccattg acctgtttac acgttcatgt gctggatact gtgtagctac	2880
cttcattttg ggaattggag atcgtcacao tagtaacatc atggtgaaag acgatggaca	2940
actgtttcat atagattttg gacacttttt ggatcacaag aagaaaaaat ttggttataa	3000
acgagaacgt gtgccatttg ttttgacaca ggatttctta atagtgatta gtaaaggagc	3060
ccaagaatgc acaaagacaa gagaatttga gaggtttcag gagatgtgtt acaaggctta	3120
tctagctatt cgacagcatg ccaatctctt cataaatctt ttctcaatga tgcttggctc	3180
tggaatgcca gaactacaat cttttgatga cattgcatac attcgaaaga ccctagcctt	3240
agataaaact gagcaagagg ctttggagta tttcatgaaa caaatgaatg atgcacatca	3300
tggtggctgg acaacaaaaa tggattggat cttccacaca attaaacagc atgcattgaa	3360
ctgaaaagat aactgagaaa atgaaagctc actctggatt ccacactgca ctgttaataa	3420
ctctcagcag gcaaagaccg attgcatagg aattgcacao tccatgaaca gcattagaat	3480
ttacagcaag aacagaaata aaatactata taattttaat aatgtaaacg caaacagggt	3540
ttgatagcac ttaaactagt tcatttcaaa attaagcttt agaataatgc gcaatttcat	3600
gttatgcctt aagtcacaaa aggtaaactt tgaagattgt ttgtatcttt ttttaaaaaa	3660
caaaacaaaa caaaaatccc caaaatatat agaaatgatg gagaaggaaa aagtgatggt	3720

ttttttgtc ttgcaaatgt tctatgtttt gaaatgtgga cacaacaaag gctgttattg	3780
cattaggtgt aagtaaactg gagtttatgt taaattacat tgattggaaa agaataaaaa	3840
tttcttattt ttccattgct gttcaattta tagtttgaag tgggtttttg actgcttggt	3900
taatgaagaa aaatgcttgg ggtggaaggg actcttgaga tttcaccaga gactttttct	3960
ttttaataaa tcaaaccttt tgatgatttg aggttttatc tgcagttttg gaagcagtca	4020
caaatgagac ctgttataag gtggtatttt tttttttctt ctggacagta tttaaaggat	4080
cttattctta tttcccaggg aaattctggg ctcccacaaa gtaaaaaaaaa aaaaaaatca	4140
tagaaaaaga atgagcagga atagttctta ttccagaatt gtacagtatt caccttaagt	4200
tgattttttt tctccttctg caattgaact gaatacattt ttcattgcatg ttttccagaa	4260
aatagaagta ttaatgttat taaaaagatt atttttttta ttaaaggcta tttatattat	4320
agaaactatc attaatatat attctttatt tacatgatct gtcccatagt catgcattgt	4380
tttgcacccc aaatttttta ttgttcatag cagcatggtc agctttcttc ttgatctata	4440
gatgaggctc aggcaactat ccatttatac caataaccag tgtataacta ctttaaggaaa	4500
acataaaaaac ttcatcttct ttccttttat ttcttatgtg aatctcccgt cttccattct	4560
cttttataat tgagaatgtc tcaatcatat gaaattagtt accagaatta acacaattta	4620
gactatcttc ctgattcctt aaaccccttt actgaagtat actcatgaat aatactttta	4680
aatatggggg aatagaaacc atgaactttt taccttttta aactatttat ccatatctcc	4740
aaagtagaac attaaacat ttttaagatat gtctcattcc caagtagtca gagctcactc	4800
tccaacttta ttaaatacta tttgagcaca ggacacattc ttaaacattt tgaaaaacat	4860
taaccaaga tgtagaggct actgctagtc gtcattctag aatctgatat tttactctgt	4920
atttgaaatg aatgattaat gtcctaggaa attagcttta gcagatgtcc aggtgccaca	4980
tcaaaaaagt gcaataatta ttgacagttt tttagattag gcatattatt ggaaaacaac	5040
tttataaaga gtgaacattg tatactctag taaaacagca tcactttaaa aatattcatt	5100
tatgaaatct gttacctata gttgaagtct tgagtagtga acaagggact ctaataccaa	5160
tactcttaat atctggctat tttagatccc ttaaagggca taattattgg aaatttaggt	5220
atttcactaa agcatgtata taatattgcc aacaagaaaa gtaaatttga agattaaggg	5280
aacttacttc tgcaaactgt cttgcgatag ttaagcagaa tttaaactct gttttaagca	5340

ggaaaccaga aagattat	ttt tgcagttgta gaagatttca taacttatta aaacttatta	5400
acattttgtg ttgttttagat ataggcagtt gatacatact aacatcccag ctttttcaat		5460
atcaggggta aattatagga aaactcagta aaatgggtaca aatctgaaag tttgatggta		5520
gaaactgaag atttaacaga gaactgtgtt ttacccgagt gccaaaaatg ctgtgagcct		5580
ccttgacaaa aatttatacc acttttgcac ttttatctat cagtccagat agttgtctcc		5640
cctccttctc ccaggacctc tccaccatta aaatgcacaa accacatggc cgatttcacc		5700
atttacattt attttcaaaa gttactacaa ccaaattaat tctattagaa gaaatgtaga		5760
caaattctat aaagactata gattgtgacc taagaaagaa atgaggcaaa gaaccaaaaca		5820
ttgaattaaa tgctacatgg gtgactaaga tctgtttcaa gtcagtgata atatagccac		5880
ttctgggtac ttcagtatca gagatcagtt ctcgtggttt agacagttcc tatctatagc		5940
tgactatcct tgtccttgaa tatgggtgtaa ctgactattg gctctacagt tttattgggc		6000
cacttaagaa atatttcctt gaataattat tttgagaaaa agtctaaaag taataaaaat		6060
aattttaaac aactgtagt aagaaatgac tgttggaaaa ttatgctttc actttctacc		6120
atattctcag ctatacaaaa ccatttattt tgaagatttt tagactactg ttaatttgaa		6180
atctgttact cttattgtgg aatttgtttt tttaaaaaag atgtttctaa ttggattttt		6240
aaaagaagaa tggaatttgg ttgctatttt acaatagaac ctaagctttt tgtggttctt		6300
agtgtcctat gtaaaactta gtgtcaaagt aatcaacttt gagattttcc cttctattct		6360
gctttatatt aaaagcccat tagaaaatgg gaacctgggtg aatatataat gaattgtaaa		6420
atattttaat gtgtaacttt ttcaactgtg aaactgactt gattttttga tgaaaacagc		6480
tgctgataaa gtattttgtg taaagtgtag ttcttattaa tcaggaaaat gatgacttga		6540
ttagactgta tatgccctct tggattttat tttaaatgga ttgggtgactt tcacataggt		6600
aaaacacagt ccatctgtat tcttttttcc atcaaaaatc gagtgatttg gaattataaa		6660
aaaattgtga gcagcctatt tgaaaggcat catggaaatt tcacagcaca ataacacgga		6720
tttgtttttt cttaatgatg taaatccgtt taattcatac tttgatcaat agcccatgct		6780
tgccaactct gaagaaattt aatttccagc agtatttttaa agctagcctg ttaacttttt		6840
ctgaatattt aaagttcctc ttttttctat gtctgcacaa actgcagacc tgggctggac		6900
ccacatactc aagagtccac cttagaaat tattttgatg tccaagacat cactaaaata		6960

tttaagttta aagataatat gtggtgttaa tagattgtgg tgcttttact atttaaagac	7020
aactttcata cttcagatgt ttttgagaag aggggaatgt gaggggaggg ggcagaacag	7080
ggaggagtgtg tttgaatgaa ttacattctt tatatccatc ctgctcattt ggggcatgtc	7140
tttaagagaa ggctgaaagt tgtgagagta tattgtatac cgtaagagaa tcaactcttc	7200
atcatggatg ggattgtgaa ggctgaacta taaaattcag cattgacagc atcctcaatt	7260
aataattctt ggtgacagaa taatacagct gggctgtttt ttaaaatata aacaatacca	7320
ttttaatta ttacattaaa aattgtaaat atatctatgt gccatggcct gggaagcctg	7380
ctttcttttt tcataaaaaat tattttttact gtatgaaaag atcatgggggt ttagctcaaa	7440
atatctgtgg tcctgataaa attggattgg taactctacc tcagaaggaa aatgggaaaa	7500
aaaaatagat gagtcacaat tcaatacttc aagctcagaa actgtgcaga tcaactgaatt	7560
ttagatttat aaagtcagag ttggcatgcc ttgtttttta tgatatggaa gaccttaaga	7620
aaaaaacttg gctgaagttt aatcgttgggt ccagccattt gaaaaaggca atagtttgag	7680
gaggttcccg aattcggcat ttgaaattca ttttgttctc tcttcttcat tattagtga	7740
tttgggtgtgt gtatacttgc acacaattct gtttgtgtac acactgcttg cttagcccta	7800
gtcaagaggc atcttttata aaagggtgtaa agaaatatca aggttctaaa attcggaaga	7860
gtttagaatt tattaggagt ttcccaagtt gggatgttag tctttaaata aacttcatgc	7920
acctattcca ctttaaggttt tgcacctcct ttttattagt gcagtgccat ttcttctgct	7980
tgattttagg tatgttaata ttccagcctt gctagttagc ataaagtgac aggtgtgagc	8040
catgaggaaa ttttctgact taatttgtac acaactacat ataagagttt tagtggagga	8100
aaaaaattag tcccttgtgc gtatacagta gttaggtaaa tgatttttct accaacagta	8160
tactccattc ctcatgtagg taagtacaga aaaggttttt aaatgtattt ttttagccag	8220
ttaaagtcta tgaatctatc tgcaacctta tttaatctgt cactataata attttgtggt	8280
tatgctaaga accatgtata cttttaggta ttcttatttt tgtcaatttt tctaggttgg	8340
caaggaggca gaaaaccttc attgtttcat attaaaatat aattagacta aacttaattc	8400
tagtatgaat ttccaaaatc attatctatt tatttcattt ttatttaatt ttgtttttat	8460
ttcattttta aaagtcctt gttcaattta acttatgttc ctaagagagg ttggagaact	8520
tggccttcat ctgatttcaa aaatgttttg agtttcaaat gaagttaatg gtttcagtgt	8580

gattcagtcc tcagacctaa ttgggttgaa taaaatctaa aagaatatac ctttttgag	8640
cataacattt taataccttg gggaatgtgg cactaccaa agaagactac taacacgtca	8700
gatgttcacc tggaagcttt atcaagaaat tcgaaccacc cttttggccc cattaattgt	8760
agcaagtta tttctctata tttgtcatt cagtgaattg aagtcctgtg gtatactgca	8820
ttcattagaa gaaaaacgtt tttaatgtcc ttttaatgat ggcccagaaa gcatttgaca	8880
cagcaagatg catgtgttac tatattgaga atatagaata ataacagtat cactaaattt	8940
aagacctctt cccagtcttg ctgttcctag caagaagttt ggctgtgac tgcacttact	9000
gtttatgctc atcagaaact gtcaatgtct gcttttcttt aactctgcag tctgtaacat	9060
cacgctgttt attaaaaaaaa aaaagaaaaa tta	9093

<210> 70

<211> 2855

<212> DNA

<213> Homo sapiens

<220>

<223> >PMS2|ENSG00000122512|ENST00000265849|2855

<400> 70

ggagcacaac gtcgaaagca gccaatggga gttcaggagg cggagcgcct gtgggagccc	60
tggaggggaac tttcccagtc cccgaggcgg atcgggtggt gcatccatgg agcgagctga	120
gagctcgagt acagaacctg ctaaggccat caaacctatt gatcggaagt cagtccatca	180
gatttgctct gggcaggtgg tactgagtct aagcactgcg gtaaaggagt tagtagaaaa	240
cagtctggat gctggtgcca ctaatattga tctaaagctt aaggactatg gagtggatct	300
tattgaagtt tcagacaatg gatgtggggg agaagaagaa aacttcgaag gcttaactct	360
gaaacatcac acatctaaga ttcaagagtt tgccgaccta actcaggttg aaacttttgg	420
ctttcggggg gaagctctga gctcactttg tgcactgagc gatgtcacca tttctacctg	480
ccacgcatcg gcgaaggttg gaactcgact gatgtttgat cacaatggga aaattatcca	540
gaaaaccccc taccctcgcc ccagagggac cacagtcagc gtgcagcagt tattttccac	600
actacctgtg cgccataagg aatttcaaag gaatattaag aaggagtatg ccaaatggt	660
ccaggtctta catgcatact gtatcatttc agcaggcatc cgtgtaagtt gcaccaatca	720
gcttggacaa ggaaaacgac agcctgtggt atgcacaggt ggaagcccca gcataaagga	780

aaatatcggc tctgtgtttg ggcagaagca gttgcaaagc ctcattcctt ttgttcagct	840
gccccctagt gactccgtgt gtgaagagta cggtttgagc tgttccgatg ctctgcataa	900
tcttttttac atctcagggt tcatttcaca atgcacgcat ggagttggaa ggagttcaac	960
agacagacag tttttcttta tcaaccggcg gccttgtgac ccagcaaagg tctgcagact	1020
cgtgaatgag gtctaccaca tgtataatcg acaccagtat ccatttgttg ttcttaacat	1080
ttctgttgat tcagaatgcg ttgatatcaa tgttactcca gataaaaggc aaattttgct	1140
acaagaggaa aagcttttgt tggcagtttt aaagacctct ttgataggaa tgtttgatag	1200
tgatgtcaac aagctaaatg tcagtcagca gccactgctg gatgttgaag gtaacttaat	1260
aaaaatgcat gcagcggatt tggaaaagcc catggtagaa aagcaggatc aatccccttc	1320
attaaggact ggagaagaaa aaaaagacgt gtccatttcc agactgcgag aggccctttc	1380
tcttcgtcac acaacagaga acaagcctca cagcccaaag actccagaac caagaaggag	1440
ccctctagga cagaaaaggg gtatgctgtc ttctagcact tcaggtgcca tctctgacaa	1500
aggcgtcctg agacctcaga aagaggcagt gagttccagt cacggacca gtgaccctac	1560
ggacagagcg gaggtggaga aggactcggg gcacggcagc acttccgtgg attctgaggg	1620
gttcagcatc ccagacacgg gcagtcactg cagcagcgag tatgcggcca gctcccagg	1680
ggacaggggc tcgcaggaac atgtggactc tcaggagaaa ggcctaaaa ctgacgactc	1740
tttttcagat gtggactgcc attcaaacca ggaagatacc ggatgtaaat ttcgagtttt	1800
gcctcagcca actaatctcg caaccccaa cacaagcgt tttaaaaaag aagaaattct	1860
ttccagttct gacatttgtc aaaagttagt aaatactcag gacatgtcag cctctcaggt	1920
tgatgtagct gtgaaaatta ataagaaagt tgtgcccctg gacttttcta tgagttcttt	1980
agctaaacga ataaagcagt tacatcatga agcacagcaa agtgaagggg aacagaatta	2040
caggaaagttt agggcaaaga tttgtcctgg agaaaatcaa gcagccgaag atgaactaag	2100
aaaagagata agtaaaacga tgtttgcaga aatggaaatc attggtcagt ttaacctggg	2160
atttataata accaaactga atgaggatat cttcatagtg gaccagcatg ccacggacga	2220
gaagtataac ttcgagatgc tgcagcagca caccgtgctc caggggcaga ggctcatagc	2280
acctcagact ctcaacttaa ctgctgttaa tgaagctgtt ctgatagaaa atctggaaat	2340
atttagaaag aatggctttg attttgttat cgatgaaaat gctccagtca ctgaaagggc	2400

taaactgatt tccttgccaa ctagtaaaaa ctggaccttc ggaccccagg acgtcgatga	2460
actgatcttc atgctgagcg acagccctgg ggtcatgtgc cggccttccc gagtcaagca	2520
gatgtttgcc tccagagcct gccggaagtc ggtgatgatt gggactgctc ttaacacaag	2580
cgagatgaag aaactgatca cccacatggg ggagatggac caccctgga actgtcccca	2640
tggaaggcca accatgagac acatcgccaa cctgggtgtc atttctcaga actgaccgta	2700
gtcactgtat ggaataattg gttttatcgc agatTTTTat gttttgaaag acagagtctt	2760
cactaacctt ttttgtttta aaatgaacct gctacttaaa aaaaatacac atcacacca	2820
tttaaaagtg atcttgagaa ctttttcaaa ccaga	2855

<210> 71

<211> 5440

<212> DNA

<213> Homo sapiens

<220>

<223> >POLA1|ENSG00000101868|ENST00000379059|5440

<400> 71

gggagattcg ggaccatggc acctgtgcac ggcgacgact ctctgtcaga ttcagggagt	60
tttgtatctt ctcgagcccg gcgagaaaaa aaatcaaaga aggggcgcca agaagcccta	120
gaaagactga aaaaggctaa agctggtgag aagtataaat atgaagtcga ggacttcaca	180
ggtgtttatg aagaagttga tgaagaacag tattcgaagc tggttcaggc acgccaggat	240
gatgactgga ttgtggatga tgatggtatt ggctatgtgg aagatggccg agagatTTTT	300
gatgatgacc ttgaagatga tgcccttgat gctgatgaga aaggaaaaga tggtaaagca	360
cgcaataaag acaagaggaa tgtaaagaag ctcgcagtga caaaaccgaa caacattaag	420
tcaatgttca ttgcttgtgc tggaaagaaa actgcagata aagctgtaga cttgtccaag	480
gatggtctgc taggtgacat tctacaggat cttaacactg agacacctca aataactcca	540
ccacctgtaa tgatactgaa gaagaaaaga tccattggag cttcaccgaa tcctttctct	600
gtgcacaccg ccacggcagt tccttcagga aaaattgctt cccctgtctc cagaaaggag	660
cctccattaa ctctgtttcc tcttaaactg gctgaatttg ctggcgatga tgtacaggtc	720
gagagtacag aagaagagca ggagtcaggg gcaatggagt ttgaagatgg tgactttgat	780

gagcccatgg aagttgaaga ggtggacctg gagcctatgg ctgccaaggc ttgggacaaa	840
gagagtgagc cagcagagga agtgaaacaa gaggcggatt ctgggaaagg gaccgtgtcc	900
tacttaggaa gttttctccc ggatgtctct tgttgggaca ttgatcaaga aggtgatagc	960
agtttctcag tgcaagaagt tcaagtggat tccagtcacc tcccattggt aaaaggggca	1020
gatgaggaac aagtattcca cttttattgg ttggatgctt atgaggatca gtacaaccaa	1080
ccaggtgtgg tattttctgtt tgggaaagt ttgattgaat cagccgagac ccatgtgagc	1140
tgttgtgtca tggtgaaaaa tatcgagcga acgctttact tccttccccg tgaaatgaaa	1200
attgatctaa atacggggaa agaaacagga actccaattt caatgaagga tgtttatgag	1260
gaatttgatg agaaaatagc aacaaaatat aaaattatga agttcaagtc taagccagtg	1320
gaaaagaact atgcttttga gatactgat gttccagaaa aatctgagta cttggaagtt	1380
aaatactcgg ctgaaatgcc acagcttcct caagatttga aaggagaaac tttttctcat	1440
gtatttggga ccaacacatc tagcctggaa ctgttcttga tgaacagaaa gatcaaagga	1500
ccttgttggc ttgaagtaaa aagtccacag ctcttgaatc agccagtcag ttggtgtaaa	1560
gttgaggcaa tggctttgaa accagacctg gtgaatgtaa ttaaggatgt cagtccacca	1620
ccgcttgtcg tgatggcttt cagcatgaag acaatgcaga atgcaaagaa ccatcaaaat	1680
gagattattg ctatggcagc tttggtccat cacagttttg cattggataa agcagcccca	1740
aagcctccct ttcagtcaca cttctgtgtt gtgtctaaac caaaggactg tattttttcca	1800
tatgctttca aagaagtcac tgagaaaaag aatgtgaagg ttgaggttgc tgcaacagaa	1860
agaacactgc taggtttttt ccttgcaaaa gttcacaaaa ttgatcctga tatcattgtg	1920
ggtcataata tttatgggtt tgaactggaa gtactactgc agagaattaa tgtgtgcaaa	1980
gctcctcact ggtccaagat aggtcgactg aagcgatcca acatgcaaaa gcttgggggc	2040
cggagtggat ttggtgaaag aaatgctacc tgtggtcgaa tgatctgtga tgtggaaatt	2100
tcagcaaagg aattgattcg ttgtaaaagc taccatctgt ctgaacttgt tcagcagatt	2160
ctaaaaactg aaagggttgt aatcccaatg gaaaatatac aaaatatgta cagtgaatct	2220
tctcaactgt tatacctgtt ggaacacacc tggaaagatg ccaagttcat tttgcagatc	2280
atgtgtgagc taaatgttct tccattagca ttgcagatca ctaacatcgc tgggaacatt	2340
atgtccagga cgctgatggg tggacgatcc gagcgtaacg agttcttggt gcttcatgca	2400

ttttacgaaa acaactatat tgtgcctgac aagcagattt tcagaaagcc tcagcaaaaa	2460
ctgggagatg aagatgaaga aattgatgga gataccaata aatacaagaa aggacgtaag	2520
aaagcagctt atgctggagg cttggttttg gaccccaaag ttggttttta tgataagttc	2580
attttgcttc tggacttcaa cagtctatat ctttccatca ttcaggaatt taacatttgt	2640
tttacaacag taaaaagagt tgcttcagag gcacagaaag ttacagagga tggagaacaa	2700
gaacagatcc ctgagttgcc agatccaagc ttagaaatgg gcattttgcc cagagagatc	2760
cggaaactgg tagaacggag aaaacaagtc aaacagctaa tgaaacagca agacttaaat	2820
ccagacctta ttcttcagta tgacattcga cagaaggctt tgaagctcac agcgaacagt	2880
atgtatggtt gcctgggatt ttcctatagc agatttttacg ccaaaccact ggctgccttg	2940
gtgacataca aaggaaggga gattttgatg catacgaaag agatggtaca aaagatgaat	3000
cttgaagtta tttatggaga tacagattca attatgataa acaccaatag caccaatctg	3060
gaagaagtat ttaagttggg aaacaaggta aaaagtgaag tgaataagtt gtacaaactg	3120
cttgaaatag acattgatgg ggttttcaag tctctgctac tgctgaaaaa aaagaagtac	3180
gctgctctgg ttgttgagcc aacgtcggat gggaattatg tcaccaaaca ggagctcaaa	3240
ggattagata tagttagaag agattggtgt gatcttgcta aagacactgg aaactttgtg	3300
attggccaga ttctttctga tcaaagccgg gacactatag tggaaaacat tcagaagagg	3360
ctgatagaaa ttggagaaaa tgtgctaaat ggcagtgctc cagtgagcca gtttgaaatt	3420
aacaaggcat tgacaaagga tccccaggat taccctgata aaaaaagcct acctcatgta	3480
catgttgccc tctggataaa ttctcaagga ggcagaaagg tgaaagctgg agatactgtg	3540
tcatatgtca tctgtcagga tggatcaaac ctactgcaa gtcagagggc ctatgcgcct	3600
gagcagctgc agaaacagga taatctaacc attgacaccc agtactacct ggcccagcag	3660
atccaccag tcgtggctcg gatctgtgaa ccaatagacg gaattgatgc tgtcctcatt	3720
gcaacgtggt tgggacttga cccacccaa tttagagttc atcattatca taaagatgaa	3780
gagaatgatg ctctacttgg tggcccagca cagctcactg atgaagagaa atacagggac	3840
tgtgaaagat tcaaatgtcc atgccctaca tgtggaactg agaatattta tgataatgtc	3900
tttgatggtt cgggaacaga tatggagccc agcttgtatc gttgcagtaa catcgattgt	3960
aaggcttcac ctctgacctt tacagtacaa ctgagcaaca aattgatcat ggacattaga	4020

cgtttcatta aaaagtacta tgatggctgg ttgatatgtg aagagccaac ctgtcgcaat	4080
cgaactcgtc accttcccct tcaattctcc cgaactgggc ctctttgccc agcctgcatg	4140
aaagctacac ttcaaccaga gtattctgac aagtccctgt acaccagct gtgcttttac	4200
cggtagacatt ttgatgcgga gtgtgcactg gagaaactta ctaccgatca tgagaaagat	4260
aaattgaaga agcaatTTTT tcccccaaa gttctgcagg actacagaaa actcaagaac	4320
acagcagagc aattcttgtc ccgaagtggc tactccgaag tgaatctgag caaactcttc	4380
gctggttgtg ccgtgaaatc ctaagggaat cccaggagta accaaggagg gggtagttga	4440
aaaatcccag cttcctctgt gcctccactc tggccctaaa tgctcctcca gcatctgttt	4500
ctcccttggg actgtgtctc atgtttgtgt gaatgtagac caggaaaggg ggctgcaaaa	4560
atgttgagtc taatgttcgt aagcatcata gaaattcctg tcttcatatt aagatgtact	4620
gctttaaAAC acaactccag agcccctccc caagctcccc tcccaagct cctgaagacc	4680
cggtttctga gggagggaaa ttgctacttg gattgagagt agctggaatg taagtgaccc	4740
caggctttgc ctgaggcct ttagcctatg tccccccac ataaagagag cttctcagag	4800
cctgactgaa gagctgacgt tttgcttttt catatgccaa ttaaaccgg tctaaatcca	4860
aatgcttctc cagccatcca ggagtggctg tccttttcag tcttgtcttt tatataggta	4920
gctgaggggg aagatttaga agccttgac tcactaaata gattaaacag agcaggcttg	4980
ttgtttgaat tgctccaaag tccaacagac acacactgag caggtgtttt acactcacat	5040
tccctttttg ccccttaaat agaaagtgca ggtaaagggt tatacaaca gaaagcacat	5100
tgaaaataat ttgatactct aacaatccat taacatgtgt aggggttacg gtgaggatca	5160
ctgtgttgta ttcagaaaaa cggggagagg gatgcttaat tggccctggc gcttgctatt	5220
tttttctcat ttcttcaca taggaccgtc tttggcagca gcaaaatgta tttcagtatg	5280
gcagtctttc ctctcttaca ttattggtaa gattatacta acaaaatgtt tccccttgta	5340
caattatgct gtgtttttta aaaacattga cctgtgtgtt ttataaaaag aaaaagtatg	5400
ttgtgccttc ttcttaagaa taaagttttc taaaggga	5440

<210> 72

<211> 1329

<212> DNA

<213> Homo sapiens

<220>

<223> >POLB|ENSG00000070501|ENST00000265421|1329

<400> 72

aacaagcgcc gtgcgccccg ccccccacg ggggcaacca ttgttccgcc ggtcgcgccg	60
gagctgggtt gtcctgctc ccgtctccaa gtcctggtac ctccttcaag ctgggagagg	120
gctctagtcc ctggttctga acactctggg gttctcgggt gcaggccgcc atgagcaaac	180
ggaaggcgcc gcaggagact ctcaacgggg gaatcaccga catgctcaca gaactcgcaa	240
actttgagaa gaacgtgagc caagctatcc acaagtacaa tgcttacaga aaagcagcat	300
ctgttatagc aaaataccca cacaaaataa agagtggagc tgaagctaag aaattgcctg	360
gagtaggaac aaaaattgct gaaaagattg atgagttttt agcaactgga aaattacgta	420
aactggaaaa gattcggcag gatgatacga gttcatccat caatttcctg actcgagtta	480
gtggcattgg tccatctgct gcaaggaagt ttgtagatga aggaattaaa aactagaag	540
atctcagaaa aaatgaagat aaattgaacc atcatcagcg aattgggctg aaatattttg	600
gggactttga aaaaagaatt cctcgtgaag agatgttaca aatgcaagat attgtactaa	660
atgaagttaa aaaagtggat tctgaataca ttgctacagt ctgtggcagt ttcagaagag	720
gtgcagagtc cagtggtagc atggatgttc tcctgacca tcccagcttc acttcagaat	780
caaccaaaca gccaaaactg ttacatcagg ttgtggagca gttacaaaag gttcatttta	840
tcacagatac cctgtcaaag ggtgagacaa agttcatggg tgtttgccag cttcccagta	900
aaaatgatga aaaagaatat ccacacagaa gaattgatat caggttgata ccaaagatc	960
agtattactg tgggtgttctc tatttcactg ggagtgatat tttcaataag aatatgaggg	1020
ctcatgccct agaaaagggt ttcacaatca atgagtacac catccgtccc ttgggagtc	1080
ctggagttgc aggagaaccc ctgccagtgg atagtgaata agacatcttt gattacatcc	1140
agtggaaata ccgggaaccc aaggaccgga gcgaatgagg cctgtatcct ccctggcaga	1200
cacaacccaa taggagtctt aatttatctt ttaacctttg ctatgtaagg gtctttggtg	1260
tttttaaatg attgtttctt cttcatgctt ttgcttgcaa tgtagtcaat aaaacctcat	1320
gtactatta	1329

<210> 73

<211> 3540

<212> DNA

<213> Homo sapiens

<220>

<223> >POLH|ENSG00000170734|ENST00000372236|3540

<400> 73

aacggccctt cgcagcgggc gcgctgtcag acctcagtct ggcggctgca ttgctgggcg	60
cgccgctctc gtctgatccc tgctggggac ggttgcccgg gcaggatcct ttacgatccc	120
ttctcggttt ctccgtcgtc acaggaata aatctcgtct gaaactcact ggaccgctcc	180
tagaaaggcg aaaagatatt caggagccct tccattttcc ttccagtagg caccgaaccc	240
agcattttcg gcaaccgctg ctggcagttt tgccaggtgt ttgttacctt gaaaaatggc	300
tactggacag gatcgagtgg ttgctctcgt ggacatggac tgtttttttg ttcaagtgga	360
gcagcggcaa aatcctcatt tgaggaataa accttgtgca gttgtacagt acaaatcatg	420
gaagggtggg ggaataattg cagtgagtta tgaagctcgt gcatttgagg tcactagaag	480
tatgtgggca gatgatgcta agaagttatg tccagatctt ctactggcac aagtctgtga	540
gtcccgtggg aaagctaacc tcaccaagta ccgggaagcc agtggtgaag tgatggagat	600
aatgtctcgt tttgctgtga ttgaacgtgc cagcattgat gaggcttacg tagatctgac	660
cagtgtgtga caagagagac tacaaaagct acaaggctcag cctatctcgg cagacttggt	720
gccaagcact tacattgaag ggttgcccca aggccctaca acggcagaag agactgttca	780
gaaagagggg atgcgaaaac aaggcttatt tcaatggctc gattctcttc agattgataa	840
cctcacctct ccagacctgc agctcaccgt gggagcagtg attgtggagg aaatgagagc	900
agccatagag agggagactg gttttcagtg ttcagctgga atttcacaca ataaggtcct	960
ggcaaaactg gcctgtggac taaacaagcc caaccgcaa accctgggtt cacatgggtc	1020
agtcccacag ctcttcagcc aaatgcccac tcgcaaaatc cgtagtcttg gaggaaagct	1080
aggggcctct gtcattgaga tcctagggat agaatacatg ggtgaactga cccagttcac	1140
tgaatcccag ctccagagtc attttgggga gaagaatggg tcttggctat atgccatgtg	1200
ccgaggggatt gaacatgatc cagttaaacc caggcaacta cccaaaacca ttggctgtag	1260
taagaacttc ccaggaaaaa cagctcttgc tactcgggaa caggtacaat ggtggctgtt	1320
gcaattagcc caggaactag aggagagact gactaaagac cgaaatgata atgacagggt	1380

agccaccag ctggttgga gcattcgtgt acaaggagac aaacgcctca gcagcctgcg	1440
ccgctgctgt gcccttacc gctatgatgc tcacaagatg agccatgatg catttactgt	1500
catcaagaac tgtaatactt ctggaatcca gacagaatgg tctcctcctc tcacaatgct	1560
tttcctctgt gctacaaaat tttctgcctc tgccccttca tcttctacag acatcaccag	1620
cttcttgagc agtgaccaa gttctctgcc aaaggtgcca gttaccagct cagaagctaa	1680
gaccagggga agtggcccag cggtgacagc cactaagaaa gcaaccacgt ctctggaatc	1740
attcttccaa aaagctgcag aaaggcagaa agttaagaa gcttcgcttt catctcttac	1800
tgctcccact caggctccca tgagcaattc accatccaag ccctcattac cttttcaaac	1860
cagtcaaagt acaggaactg agcccttctt taagcagaaa agtctgcttc taaagcagaa	1920
acagcttaat aattcttcag tttcttcccc ccaacaaaac ccatgggtcca actgtaaagc	1980
attaccaaac tctttacca cagagtatcc aggggtgtgtc cctgtttgtg aaggggtgtc	2040
gaagctagaa gaatcctcta aagcaactcc tgcagagatg gatttggtccc acaacagcca	2100
aagcatgcac gcctcttcag cttccaaatc tgtgtctggag gtgactcaga aagcaacccc	2160
aatccaagt cttctagctg ctgaggacca agtgccctgt gagaagtgtg gctccctggt	2220
accggtatgg gatatgccag aacacatgga ctatcatttt gcattggagt tgcagaaatc	2280
ctttttgcag cccactctt caaaccccca ggttgtttct gccgtatctc atcaaggcaa	2340
aagaaatccc aagagccctt tggcctgcac taataaacgc ccaggcctg agggcatgca	2400
aacattggaa tcatttttta agccattaac acattagtgc tgccctcagg cttgcctgta	2460
ggatttaata ttttttatct ttacagatct ttatctttaa ttttttatct ttacagattt	2520
ccctgagaaa gggaattatg aaatttttaa taaaaaaat aatccattta ggtgctgagt	2580
tacggtccca tctcttcaca ggcatggatt ctaatccac tgctgacaga gatgtaaaaa	2640
ttcatcctac cagagttttt aatctttagc atttagggag gcagtgtcat aaagtaaaaa	2700
gtgtgtgggc cttggagtct aagagacgtg gttgcaaact tagctctggt tattgcaatg	2760
agggccttga acaagtcatt ttcttcacat tctcatctgt aaaatggaga taatacctta	2820
cagattattg cagattaata acaatgtatt caaattatgt aactcggccg ggtacaatgg	2880
ctcacgcctg taatcctaac actttgggag gccgaggcag acagatcacc tgaggtcagg	2940
agtttgagac cagcctggcc aacatggcaa aaccatctct actaaaaata gaaaaattag	3000

ccaggcacgt tccaggcacc tgtgatccca gctacttaga ggctgaggca gaagaattgc	3060
tttaaccttg gaggcggagg ttgcattgag ctgagatcat gctagtgcgc tccagcctgg	3120
gcaacagagc gagacttcat ctcaaaaaat aaaaaatagg ggccaggcac agtggctcat	3180
acctgtaatg ccagcacttt gggaggccaa ggcgggcaga tcacgaggtc aggagtttca	3240
gaccaatatg gtgaaacccc atctctacta aaattacaaa aaaaattatc caggcgtggt	3300
ggtgcacgcc tgtaatccca gctactcagg aggctaaggc aggagaatca cttgaaccca	3360
ggaggcagag gttggagtga gctgagatcg cgccaccgca ctccagcctg ggcaacagag	3420
cgagactcca tctcaaaaaa aaacaagaac aaaaacaaac ataaagttgg cacagaaaag	3480
ggaccaagtt taaaaaaggg ttttaaattgt aatgagactt gcatagttaa aaaaaaaaaa	3540

<210> 74
 <211> 2792
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >POLL|ENSG00000166169|ENST00000370162|2792

<400> 74	
tgactcgga gctattctgg ccatttgccc tccttcccc cttcgtccgc tctcattggc	60
tctgctggta agtggcttat tcctgcccac ccccggtga ctagcttggc cagtagtcga	120
ccccacccgg ggaccgactc tgggggttgg agagactctt ggggccgggg tcgggcactc	180
cagctttctt ctagccccga gctgggattc cctggcctgc gccagctgcg tacacggcga	240
gtacaccgca cctgcccggg acttcacccg cagctgcgag actcctccat tcccggaggg	300
ctccccacac ctgctgcggc cgtgccccat ctccccgcag ctgcggcgct gagcaccccc	360
atgtcgagag gctgagacca ggacaggtgc agggcgttcc cactcacccc cgaaagtcct	420
cctccttcct ctgcgagtct gtgttgagc agccctgacc aacgctcaa taggccggga	480
tccagccata cttcaatgga tcccaggggt atcttgaagg catttcccaa gcggcagaaa	540
attcatgctg atgcatcatc aaaagtactt gcaaagattc ctaggaggga agaggagaaa	600
gaagcagaag agtggctgag ctcccttcgg gcccatgttg tgcgcactgg cattggacga	660
gcccgggcag aactctttga gaagcagatt gttcagcatg gcggccagct atgccctgcc	720
cagggccag gtgtcactca cattgtggtg gatgaaggca tggactatga gcgagccctc	780

cgctttctca gactacccca gctgcccccg ggtgctcagc tggatgaagtc agcctggctg	840
agcttgtgcc ttcaggagag gaggctggtg gatgtagctg gattcagcat cttcatcccc	900
agtaggtact tggaccatcc acagcccagc aaggcagagc aggatgcttc tattcctcct	960
ggcaccatg aggccctgct tcagacagcc ctttctcctc ctctcctcc caccaggcct	1020
gtgtctcctc cccaaaaggc aaaagaggca ccaaacaccc aagcccagcc catctctgat	1080
gatgaagcca gtgatgggga agaaaccag gttagtgcag ctgatctgga agccctcatc	1140
agtggccact accccacctc ccttgaggga gattgtgagc ctagcccagc ccctgctgtc	1200
ctggataagt gggctctgtc acagccctca agccagaagg cgaccaatca caacctccat	1260
atcacagaga agctggaagt tctggccaaa gcctacagtg ttcagggaga caagtggagg	1320
gccctgggct atgccaaggc catcaatgcc ctcaagagct tccataagcc tgtcacctcg	1380
taccaggagg cctgcagtat ccctgggatt gggaagcgga tggctgagaa aatcatagag	1440
atcctggaga gcgggcattt gcggaagctg gaccatatca gtgagagcgt gcctgtcttg	1500
gagctcttct ccaacatctg gggagctggg accaagactg cccagatgtg gtaccaacag	1560
ggcttccgaa gtctggaaga catccgcagc caggcctccc tgacaacca gcaggccatc	1620
ggcctgaagc attacagtga cttcctggaa cgtatgccc gggaggaggc tacagagatt	1680
gagcagacag tccagaaagc agcccaggcc tttaactctg ggctgctgtg tgtggcatgt	1740
ggttcatacc gacggggaaa ggcgacctgt ggtgatgtcg acgtgctcat cactcaccca	1800
gatggccggt cccaccgggg tatcttcagc cgctccttg acagtcttcg gcaggaaggg	1860
ttcctcacag atgacttggg gagccaagag gagaatggtc agcaacagaa gtacttgggg	1920
gtgtgccggc tcccagggcc agggcgggcg caccggcgcc tggacatcat cgtggtgccc	1980
tatagcgagt ttgcctgtgc cctgctctac ttaccggct ctgcacactt caaccgctcc	2040
atgcgagccc tggccaaaac caaggcatg agtctgtcag aacatgccct cagcactgct	2100
gtggtccgga acacccatgg ctgcaagggt gggcctggcc gagtgtgcc cactcccact	2160
gagaaggatg tcttcaggct cttaggcctc ccctaccgag aacctgctga gcgggactgg	2220
tgacccatgg ctgggggtgc tgaggagagc cgagttggac tggctacccc tcctggccac	2280
ccagtactcc ctccagcctc agctggctga acctcgccgc tccaaccacc agcttctca	2340
gcgagcaggg cccagggtc tgggcctgaa gcaagagcca gcccggtcc cagtgtctgc	2400

ccggctccca gtgtctgccc agccctctcc cagacaggag caggctgcca ccccttctac	2460
ctcaccactg cccctcgaag aattttgcaa atggcccctt gccccatttt aagcaggagc	2520
aggtggctgg tttgaagccc caggtatccc ctttcctgc tatgggaaag gccaaagctgc	2580
tgggtgggga cagaagctgc aggggagagg gaagcagccg tgctgtcaac atcatccggc	2640
accctctggg gtaggagaac agccattcca catgtgttcc ctctatccgt cctgcttcct	2700
gggcagctgg tggtgctggg aatggggctgc cccagccttg gtgaggacag tgttgggagg	2760
cccaggggcc cagtaaagtg catttgacat tg	2792

<210> 75
 <211> 3253
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >POLN|ENSG00000130997|ENST00000511885|3253

<400> 75	
agtgcgaggc gcgcggggca cggagggcgg tggcggcggg ctcttgcgag aagcaagcgg	60
aacttcctga gcgtgcttag gtgtcacagc tgggtgacct tccgaaggaa ggttttgccg	120
ctggcgcgtc tgggtgtatcg cggccttggg gccgcggatc cctggggctt gccgctgcct	180
ctgggggttcg cgcggccctg gcgtcggggc tgcctcagcg ctggctgaag gccggcccga	240
agcaccctta attgtggccg cgcgcttccc tgctgtcctt gctgttctcc gcgtcgcggt	300
gggtggactct ggaggtggag ccttgcccgg cgagacgttg aggattttgt gaaaatggaa	360
aattatgagg cattggtagg ctttgccttc tgtaatacac cgctctccag tgttgctcag	420
aagattatgt ctgctatgca ttcaggtgat ttagtggatt ctaagacttg gggaaagagt	480
acagagacta tggaagtgat aaacaagtcc agtggttaagt attcagtaca acttgaagac	540
aggaagactc aatcaccaga aaaaaaggat cttaaactct taagaagtca gacatcaaga	600
ggttctgcca agctgtctcc tcagtccttc agtgtcaggc tcacagatca gctgtctgct	660
gacaaaaaac agaagagcat cagctcattg actctttcaa gttgtttaat tccacagtat	720
aatcaagagg cttcagttct acagaaaaag gggcataaaa gaaagcattt cctaattggag	780
aatataaata atgaaaataa aggaagcatt aatcttaaaa gaaaacatat tacatataat	840

aatttgtcag agaaaacaag taaacaaatg gcattggaag aagatactga tgacgccgaa	900
ggctacctaa attctgggaa ctcaggagca ttgaaaaaac atttttgtga tattaggcat	960
ttggatgatt gggcaaaaag ccagctgatt gaaatgctca aacaggcagc agccctgggtg	1020
ataactgtga tgtatactga tggttccacc cagctaggag ctgaccagac ccccgtttct	1080
tctgttagag gaattgtggt gttagtataa cgccaagcag aggggtggcca tggctgtcca	1140
gatgccccgg cctgtgggtc tgttctggag ggctttgtgt cagatgatcc atgcatctac	1200
attcaaatag agcactctgc tatctgggac caagaacagg aggcacatca acaatttgcc	1260
cggaacgtgc tatttcaaac aatgaaatgt aaatgtcctg ttatttgttt taatgctaag	1320
gattttgtga gaatagtgtc gcagtttttt ggcaatgatg gcagttggaa gcatgttgct	1380
gattttatag ggctagatcc cagaattgct gcatggctta tagatcctag tgatgccaca	1440
ccctcttttg aagatttagt agaaaaatac tgtgaaaaat ccattacagt taaagtgaac	1500
agcacatatg gaaattcctc aagaaatatt gtgaatcaga atgtacgtga gaacctgaag	1560
acactctaca gacttacaat ggacctttgc tctaaactga aggattatgg tttatggcaa	1620
ctatttcgta ctttggagct tcctctgata ccaattttgg cagtgatgga aagccatgcc	1680
attcaggtga acaaagagga gatggagaag acgtcagcac ttcttggggc tcgtctcaag	1740
gaattggagc aagaagctca ttttgttgca ggagaacggt ttcttataac gagcaataac	1800
cagcttcgag agatcctctt tggcaagtta aagctgcacc tgctgagtca aaggaacagt	1860
ctccccagaa cgggggttgca gaaatacccg tctacatcag aagcagtgtt aaatgctctg	1920
cgagaccttc atccattacc caagataatt ttggaatata ggcaggttca caagatcaag	1980
tcaacctttg tagatggatt actagcttgc atgaaaaagg gctccatttc ctctacatgg	2040
aatcagactg gaactgtgac tggaagactt tcagccaagc atcctaatat ccaaggtatc	2100
tccaagcacc caattcagat tactacacct aagaatttta aaggtaaaga agacaagatt	2160
ctcacgatct ccccgagggc catgtttgtt tcatccaaag gccacacctt tctagcagca	2220
gacttttcac agattgaatt gcgcattctt acacatttat ctggagatcc ggaacttctg	2280
aagttattcc aggaatctga aagagatgat gtattttcta ctctgacttc acagtggaag	2340
gatgtgcccc tggaacaggt gacacacgca gacagagagc aaaccaagaa ggtggtgtac	2400
gcggtggtct atggagcagg gaaggagcgg ctggctgctt gccttggagt tcctattcag	2460

gaagctgccc agtttttggga gagtttttttg cagaagtaca agaaaatcaa ggacttgcgc	2520
cgagcagcta ttgccagtg tcaccagaca ggctgtgtgg tgtccatcat gggcagaagg	2580
agaccctgc caaggattca cgctcatgac cagcaactcc gggcacaagc agagcgacag	2640
gcagtgaact tcgtgggtgca aggctccgct gctgacctct gcaagctggc catgatccat	2700
gtcttcactg cagtgggtgc ttcccacacc ttgacggcca ggctgggtggc ccagatccat	2760
gatgagctgc tgtttgaagt ggaagatccg cagatcccgg agtgtgcagc tctcgtcagg	2820
aggacatgg agtccttggga acaggtgcag gcattggagc tgcagcttca ggtaccctc	2880
aaggtgagcc tgagtgccgg ccgctcatgg ggacacctgg tgccactgca ggaggcctgg	2940
ggccctccgc caggcccatg tcgcactgag tctcccagca acagcctggc tgccctggg	3000
tcccctgcca gcaccagcc cccaccctg catttttcgc cttcattttg tctgtagccc	3060
caggcaacag tgggaggaga gaactggttt ccagcagtcc atttgtgtggc cttccccaag	3120
gtcaccagct ctgtacgcc caggacgcat taaccctttg gggctggggg ggcccgcct	3180
cccctggagt aaatgcctgt gcaaagccct cctcatcctt gtctgagtga attcttcctg	3240
tgagtgggaa gtg	3253

<210> 76

<211> 8775

<212> DNA

<213> Homo sapiens

<220>

<223> >POLQ|ENSG00000051341|ENST00000264233|8775

<400> 76

gggcgggccc ggaggtttga gtttgaagac tggcgggaag atgtccgcag ctgttgccag	60
gccagggttc tcccgagagg gaggacgctg ggactgtggc ttgccctgat cggccgagaa	120
gagtttgcca tgaatcttct gcgtcggagt gggaaacggc ggcgttcaga atcaggctca	180
gattcgttct cgggaagcgg cggtgacagc agtgccagcc cccagttcct ctccgggtcc	240
gtgctgagcc cgccgcccgg ctttggtcgc tgcctgaagg ccgcagctgc aggagaatgc	300
aagcctacag ttcctgacta cgaaagagac aagctactat tggcaaactg gggacttcct	360
aaagcagttc tggaaaaata ccacagtttt ggtgtaaaaa agatgtttga atggcaggca	420
gagtgccttt tgcttggaca agtcctggaa ggaaagaatt tagttttattc agctcctaca	480

agtgctggga agactcttgt ggcagaatta cttattttga agcgggtttt ggaaatgcgg	540
aagaaagctt tgtttattct tccctttgtt tctgtggcta aagagaagaa atactacctc	600
cagagtctgt ttcaggaagt aggaataaaa gtagacgggtt atatgggcag cacctctcca	660
tcaaggcatt tctcttcatt ggatattgca gtctgcacaa ttgagagagc caatggtctg	720
atcaatcgcc tcatagagga aaataagatg gatctgttag gaatggtggt tgtggatgaa	780
ttacatatgc tgggagactc tcaccgaggg tatctgctgg aacttttgct gaccaagatt	840
tgctatatta ctcggaaatc agcatcttgt caggcagatc tagccagttc tctgtcta	900
gctgtgcaaa tcgttggcat gagtgtacc cttcctaatt tggagcttgt ggcttcctgg	960
ttgaatgctg aactctacca taccgacttt cgccctgtac cgcttttgga gtcagtaaaa	1020
gttggaaatt ccatatatga ctcttcaatg aaacttgtga ggggaatttga gcccatgcta	1080
caagtgaagg gagatgagga ccatgttgtt agtttatgtt atgagacgat ttgtgataac	1140
cattcagtat tacttttttg tccatcaaag aaatggtgtg agaagctggc agatatcatt	1200
gctcgagagt ttataatct acatcatcaa gctgagggat tggtgaaacc ctctgaatgc	1260
ccaccagtaa ttctggaaca aaaagaactc ctggaagtga tggatcagtt aagacgtttg	1320
ccttcaggac tggactctgt attacagaaa actgtaccat ggggagtagc atttcatcat	1380
gcaggtctta cttttgagga gagggatatc attgaaggag cttttcgtca aggtctcatt	1440
cgggtcttgg cggcaacttc tactctttct tctgggggtga atttacctgc acgtcgtgtg	1500
attattcgaa cccctatttt tgggtggtcga cctctagata ttcttactta taagcagatg	1560
gttggccgtg ctggcaggaa aggagtggac acagtaggcg agagtatctt aatttgtaag	1620
aactctgaga aatcaaaagg catagctctc cttcagggtt ctctaaagcc tgttcgcagc	1680
tgtctgcaaa gacgagaagg agaagaagta actggcagca tgatacgagc tattctggag	1740
ataatagttg gtggagtggc aagtacatca caagatatgc atacttatgc tgcctgcaca	1800
tttttggtg caagtatgaa agaagggaag caaggaattc agagaaatca agagtctgtt	1860
cagcttggag cgattgaggc ctgtgtgatg tggctactag aaaatgaatt catccagagt	1920
acagaagcca gtgatggaac agaaggaaag gtgtatcatc caacacatct tggttcggcc	1980
actctttctt cttcactttc tccagctgat acttttagata tttttgctga cctgcaaaga	2040
gcaatgaagg gctttgtttt agagaatgat cttcatattc tctatctggt tacacctatg	2100

tttgaggatt ggactactat tgattggtat cgatTTTTct gtttatggga gaagttgcca	2160
acttcaatga aaaggggtggc agagctagtg ggagttgaag aggggttctt ggcccgttgt	2220
gtgaaaggaa aagtagtagc cagaactgag agacagcatc gacaaatggc catccataaa	2280
aggTTTTca ccagtcttgt gctattagat ttaatcagtg aagttccctt aagggaata	2340
aatcagaaat atggatgcaa tcgtgggcag attcaatctt tgcaacagtc agctgctgtt	2400
tatgcaggga tgattacagt attttccaac cgtctgggct ggcaacaat ggaactacta	2460
ctttcccaat ttcagaagcg tcttacgttt ggcatccaga gggagctgtg tgacctggtt	2520
cgggtatcct tactaaatgc tcagagagcc agggttctct atgcttctgg ctttcatact	2580
gtggcagacc ttgctagagc aaatattgtg gaggtggagg tgattctgaa aaatgctgtg	2640
cctttcaaaa gtgcccggaa ggcagtggat gaggaagagg aagcagttga agaacgtcgc	2700
aatatgcgaa ctatctgggt gactggcaga aaaggtttaa ctgaaaggga agcagcagcc	2760
cttatagtgg aagaagccag aatgattctg cagcaggact tagttgaaat gggagtgcaa	2820
tggaatccat gtgccctgtt acattctagt acatgctcat tgactcatag tgagtccgaa	2880
gtaaaggaa acacatttat atcccaaact aagagttctt ataaaaatt aacatcaaag	2940
aacaaaagta acacaatatt tagtgattct tatattaagc attcaccaa tatagtgcaa	3000
gacttaaata aaagtagaga gcatacaagt tcctttaatt gtaatttcca gaatgggaat	3060
caagaacatc agacatgttc cttttcaga gcaagaaaac gggcctcttt agatataaat	3120
aaagagaagc caggagcctc tcagaatgag gggaaaacaa gtgataagaa agttgttcag	3180
actttttcac agaaaacaaa aaaggcacct ttgaatttca attcagaaaa gatgagcaga	3240
agttttcgat cttggaaacg tagaaagcat ctaaagcgat ctagggacag cagccccctg	3300
aaagactctg gagcgtgtag aatccattta caaggacaga ctctgtctaa tcctagtctt	3360
tgtgaagacc cgtttacctt agatgagaag aaaacggaat ttagaaattc agggccattt	3420
gctaaaaatg tatctttgag tggttaaggaa aaagataata aaacatcatt cccattacaa	3480
ataaagcaaa attgttcatg gaacataaca ctaactaatg ataattttgt ggagcatatt	3540
gtcacaggat ctcagagtaa aaatgtgact tgtcaggcca ctagtgtggt tagtgaaaag	3600
ggcagaggag tagctgttga ggcagaaaaa ataaatgaag tgctgataca aaatggttca	3660
aaaaaccaga atgtttatat gaaacaccat gacatccatc caattaacca gtacctgcga	3720

aagcaatctc atgaacagac aagcactatt accaaacaga aaaatataat agagagacaa	3780
atgccctgtg aagcagtcag tagttacata aatagagact caaatgttac tatcaattgt	3840
gaaaggataa agcttaatac agaggaaaat aaaccaagtc attttcaggc attaggagat	3900
gatataagca gaactgtgat acccagtga gtactttccat cagctggagc atttagcaaa	3960
tcagaaggcc agcatgagaa ttttctaaat atttctagac tacaagaaaa aacaggtact	4020
tatacaacaa acaaaaactaa aaataatcat gtttctgact taggttttagt cctctgtgat	4080
tttgaagata gtttctatct ggatactcag tcagagaaaa taatacaaca gatggcaact	4140
gaaaatgcc aactaggagc aaaggacacc aacctggcag cagggataat gcagaagagc	4200
ttagtccaac agaactcaat gaactctttt cagaaggagt gtcacattcc ttttcctgct	4260
gaacagcacc ctctaggagc gactaagata gatcatttgg accttaagac tgtaggtact	4320
atgaaacaaa gcagtgattc acatgggggt gatatcctga ctccagaaag cccgattttc	4380
cattctccaa tactattgga ggaaaatggc ctttttttaa aaaagaatga agtttctgtt	4440
actgattcac aattaaatag ttttcttcaa ggttatcaaa cacaagaaac tgtgaaacca	4500
gttatacttc tgattcctca aaagagaact cccactgggtg tagaaggaga atgtcttcca	4560
gttcctgaaa caagtttgaa tatgagtgat agtttactat ttgatagctt cagtgatgac	4620
tatctagtaa aagaacaatt acctgatatg caaatgaaag aacccttcc ttcagaagta	4680
acatcaaacc attttagtga ttctctgtgt ctacaagaag acctaattaa aaaatcaaat	4740
gtaaatagaga atcaagatac ccaccagcag ttgacttggt ccaatgatga atctattata	4800
ttttcagaaa tggattctgt tcagatgggt gaagctttgg acaatgtgga tatatttcct	4860
gtccaagaga agaatcatac tgtagtatct cctagagcat tagaactaag tgatccagta	4920
cttgatgagc accaccaagg tgatcaagat ggaggagatc aagatgaaag ggctgaaaaa	4980
tcaaaattaa ctgggaccag gcaaaatcat tcattcatat ggtcaggggc atcatttgat	5040
ctaagtccag gactgcaaag gatttttagat aaagtatcca gtcctctaga aaatgaaaag	5100
ctaaaatcaa tgactataaa cttttccagt ttgaatagaa aaaatacaga gttaaataaa	5160
gaacaagaag ttatttcaaa cttggagaca aaacaagtc agggattttc attttcttct	5220
aataatgaag taaaaagcaa gattgagatg ctagaaaaca atgccaatca tgatgaaacc	5280
tcatcccctc tacctcgtaa agaaagtaat atagttgatg ataatggctc cattcctcct	5340

acacccattc caacatctgc ttctaagctg acatttccag ggattcttga aacacctgta	5400
aacccttgga aaactaataa tgttttaciaa cctggtgaaa gttattttatt tggctcacct	5460
tcagatatta aaaaccacga ttttaagtcca gggagtagaa atggggttcaa agacaacagc	5520
cctattagtg acacaagctt ttcacttcag ttatcacagg atggattaca gtttaactcca	5580
gcctcaagca gttcagaaag tttgtccata attgatgtag caagtgacca aaatcttttc	5640
caaacattca ttaaggagtg gcggtgcaaa aagcgatttt ccatctcact ggcttgtgaa	5700
aagattagaa gtttgacatc ttctaaaact gctactattg gcagtagggt taagcaagct	5760
agctcacctc aggaaattcc tattagagat gatggatttc ccattaaagg ttgtgatgac	5820
accttggtgg ttggactggc agtatgctgg ggtggaaggg atgcctatta tttttcactg	5880
cagaaggaac aaaagcattc tgaaattagt gccagtttgg ttccaccttc tttagatcca	5940
agcctgactt tgaaagacag gatgtggtac cttcaatctt gcttgcgaaa ggaatctgat	6000
aaagaatgtt ctgttgtcat ctatgacttc atccagagct ataaaaattct tcttctttct	6060
tgtggcatct ccttggagca aagttatgaa gatcctaagg tggcatgctg gttactagat	6120
ccagattctc aggagccgac tcttcatagc atagttacca gttttcttcc tcatgagctt	6180
ccactcctag aagggatgga gaccagccaa gggattcaaa gcctggggct aaatgctggc	6240
agtgagcatt ctgggcgata cagagcatct gtggagtcca ttctcatctt caactctatg	6300
aatcagctca actctttgtt gcagaaggaa aaccttcaag atgttttccg taaggtggaa	6360
atgccctctc agtactgctt ggccttgcta gaactaaatg gaattggctt tagtactgca	6420
gaatgtgaaa gtcagaaaca tataatgcaa gccaaagtgg atgcaattga gaccagggc	6480
tatcaactag ctggccacag tttttctttc accagttcag atgacatcgc tgaggtttta	6540
tttttggaat tgaagttgcc cccaaataga gagatgaaaa accaaggcag caagaaaact	6600
ctgggttcta ccagaagagg gattgacaat ggacgcaagc taaggctggg aagacagttc	6660
agcactagta aggacgtttt aaataaatta aaggcattac atcctttacc aggcttgata	6720
ttagaatgga gaagaatcac taatgctatt accaaagtgg tctttcccct tcagcgggaa	6780
aagtgtctta atccttttct tggaatggaa agaatctatc ctgtatcaca gtcgcacact	6840
gctacaggac gaataacctt tacagaacca aatattcaga atgtgccaag agattttgaa	6900
atcaaaatgc caacactagt aggagaaagc ccaccttctc aagctgtagg caaaggccta	6960

cttcccatgg gcagaggaaa atataagaag ggtttcagcg tgaatcctag atgccaggca	7020
cagatggagg agagagctgc agacagagga atgccatttt caattagcat gcgacatgcc	7080
tttgtgcctt tcccagggtg ttcaatactg gctgctgact actctcagct tgaactgagg	7140
atcttggtc atttatccca tgatcgtcgt ctcatccaag tgttaaacaac tggagctgat	7200
gttttcagga gcattgcagc agagtgggaag atgattgagc cagagtctgt tggggatgat	7260
ctgaggcagc aggcaaaaca gatttgctat gggatcattt atggaatggg agctaaatct	7320
ttgggagagc agatgggcat taaagaaaat gatgctgcat gctatatatga ctccttcaaa	7380
tccagataca cagggattaa tcaattcatg acagagacag tgaagaattg taaaagagac	7440
ggatttggtc agaccatttt gggaaggcgt agatatttgc caggaatcaa agacaacaac	7500
ccttatcgta aagctcacgc tgagcgtcaa gctatcaaca caatagtcca aggatcagca	7560
gctgatattg tcaaaatagc cacagttaac attcagaagc aattagagac cttccactca	7620
accttcaaat cccatggtca tcgagagggt atgctccaaa gtgaccaaac aggattgtca	7680
cgaaagagaa aactgcaagg gatgttctgc ccaatcagag gaggcttctt catccttcaa	7740
ctccatgatg aactcctata tgaagtggca gaagaagatg ttgttcaggt agctcagatt	7800
gtcaagaatg aaatggaaag tgctgtaaaa ctgtctgtga aattgaaagt gaaagtgaaa	7860
ataggcgcca gctggggaga gctaaaggac tttgatgtgt aactgtgctg ttgatgaagt	7920
cctcccaggg aagcctgtgc agatgcagtc acctggaaag aacagagatt accctttcac	7980
ctacctcagc aaaacaaact ttcaagtctt gatagactta gcctagtaat tttatagtga	8040
gagtttcaaa ctatatatca gtgtctatag catcaaaaac ttctgggggc gtgggggaag	8100
tagaatacca agtataatag ttacattcac tttcaaagag catctatgaa tttgcctttt	8160
gtaacttact gtggctttta acatattcag aacagatgct tgaaatatgc acttagcact	8220
ttggttccac atctgtctgg gtaaaccatg aagaaaatga agctgctgcc tcaatcgacc	8280
cagacagcag ccataggcag ataaagattt ggtttcaccc tgggtggtggt aggcacgtg	8340
tgtgactttt tttcctctaa tatcaatttt acagtacgga aatagtattt taaaatagta	8400
ttggctaata aattatgaat tctataaagt agtaagactt ggtatggttg gagtgtagga	8460
atgaatattc atgaaatgtt tcttattgct tttccttccc taattcatac aatgaatgta	8520
tttggataac ttacatatta taaaataaac tatacctctt caagaggat cctgttctgt	8580

aagatcagat gtttttattg caggtcaata taatactgcc agagacagaa aataccccct	8640
tatcagtccc ttagtgcctc tttctgtttg tggcatggtg agaaaacca tgctgaaaag	8700
attgtacttt gtgatcccaa tcagagggat ggagctaatc tttttgctgt tgaaataaaa	8760
tgaatttatg agaaa	8775

<210> 77
 <211> 13506
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PRKDC|ENSG00000253729|ENST00000314191|13506

<400> 77	
ggggcatttc cgggtccggg ccgagcgggc gcacgcgcgg gagcgggact cggcggcatg	60
gcgggctccg gagccggtgt gcgttgctcc ctgctgcggc tgcaggagac cttgtccgct	120
gcggaccgct gcggtgctgc cctggccggt catcaactga tccgcggcct ggggcaggaa	180
tgcgtcctga gcagcagccc cgcggtgctg gcattacaga catctttagt tttttccaga	240
gatttcggtt tgcttgtatt tgtccggaag tactcaaca gtattgaatt tcgtgaatgt	300
agagaagaaa tcctaaagtt tttatgtatt ttcttagaaa aaatgggcca gaagatcgca	360
ccttactctg ttgaaattaa gaacacttgt accagtgttt atacaaaaga tagagctgct	420
aaatgtaaaa ttccagccct ggaccttctt attaagttac ttcagacttt tagaagttct	480
agactcatgg atgaatttaa aattggagaa ttatttagta aattctatgg agaacttgca	540
ttgaaaaaaaa aaataccaga tacagtttta gaaaaagtat atgagctcct aggattattg	600
ggtgaagttc atcctagtga gatgataaat aatgcagaaa acctgttccg cgcttttctg	660
ggtgaactta agaccagat gacatcagca gtaagagagc ccaaactacc tgttctggca	720
ggatgtctga aggggttgct ctcaactctg tgcaacttca ctaagtccat ggaagaagat	780
ccccagactt caaggagat ttttaatttt gtactaaagg caattcgtcc tcagattgat	840
ctgaagagat atgctgtgcc ctgagctggc ttgcgcttat ttgccctgca tgcattctag	900
tttagcacct gccttctgga caactacgtg tctctatttg aagtcttggt aaagtgggtg	960
gcccacacaa atgtagaatt gaaaaaagct gcactttcag ccctggaatc ctttctgaaa	1020

caggtttcta atatggtggc gaaaaatgca gaaatgcata aaaataaact gcagtacttt	1080
atggagcagt tttatggaat catcagaaat gtggattcga acaacaagga gttatctatt	1140
gctatccgtg gatatggact ttttgcagga ccgtgcaagg ttataaacgc aaaagatggt	1200
gacttcatgt acgttgagct cattcagcgc tgcaagcaga tgttcctcac ccagacagac	1260
actggtgacg accgtgttta tcagatgcca agcttcctcc agtctgttgc aagcgtcttg	1320
ctgtaccttg acacagttcc tgaggtgtat actccagttc tggagcacct cgtggtgatg	1380
cagatagaca gtttcccaca gtacagtcca aaaatgcagc tgggtgtgttg cagagccata	1440
gtgaagggtg tcctagcttt ggagcaaaa gggccagttc tcaggaattg cattagtact	1500
gtggtgcatc agggtttaat cagaatatgt tctaaaccag tggtccttcc aaagggccct	1560
gagtctgaat ctgaagacca ccgtgcttca ggggaagtca gaactggcaa atggaagggtg	1620
cccacataca aagactacgt ggatctcttc agacatctcc tgagctctga ccagatgatg	1680
gattctatth tagcagatga agcatttttc tctgtgaatt cctccagtga aagtctgaat	1740
catttactth atgatgaatt tgtaaaatcc gttttgaaga ttgttgagaa attggatctt	1800
acacttgaaa tacagactgt tggggaacaa gagaatggag atgaggcgcc tgggtgttgg	1860
atgatcccaa cttcagatcc agcggctaac ttgcatccag ctaaacctaa agatthttcg	1920
gctthcatta acctggtgga atthttgcaga gagattctcc ctgagaaaca agcagaatth	1980
thtgaacct gggtgtactc atthttcatat gaattaatth tgcaatctac aaggttgccc	2040
ctcatcagtg gthttctaaa attgctthct attacagtaa gaaatgccaa gaaaataaaa	2100
tatttcgagg gaggtagtcc aaagagtctg aaacactctc ctgaagaccc agaaaagtat	2160
tcttgctthg cthttatthgt gaaatthggc aaagaggtgg cagthtaaat gaagcagtac	2220
aaagatgaac thttggcctc ttgthtgacc thtctthctgt ccttgccaca caacatcatt	2280
gaactcgatg ttagagccta cgttcctgca ctgcagatgg cthtcaaact gggcctgagc	2340
tataccccct tggcagaagt aggcctgaat gctctagaag aatggtcaat ttatattgac	2400
agacatgtaa tgcagcctta ttacaaagac atthtccccct gcctggatgg atacctgaag	2460
acttcagcct tgtcagatga gaccaagaat aactgggaag tgtcagctct thtctgggct	2520
gccagaaaag gatttaataa agtggtgtta aagcatctga agaagacaaa gaacctthca	2580
tcaaacgaag caatatcctt agaagaaata agaattagag tagtacaaat gcttgatct	2640

ctaggaggac aaataaaca	aaatcttctg acagtcacgt	cctcagatga gatgatgaag	2700
agctatgtgg cctgggacag	agagaagcgg ctgagctttg	cagtgccctt tagagagatg	2760
aaacctgtca ttttcctgga	tgtgttcctg cctcgagtca	cagaattagc gctcacagcc	2820
agtgacagac aaactaaagt	tgcagcctgt gaacttttac	atagcatggg tatgtttatg	2880
ttgggcaaag ccacgcagat	gccagaaggg ggacaggag	ccccacccat gtaccagctc	2940
tataagcggg cgtttcctgt	gctgcttcga cttgcgtgtg	atgttgatca ggtgacaagg	3000
caactgtatg agccactagt	tatgcagctg attcactggg	tcactaaca caagaaattt	3060
gaaagtcagg atactgttgc	cttactagaa gctatatattg	atggaattgt ggaccctgtt	3120
gacagtactt taagagattt	ttgtggtcgg tgtattcgag	aattccttaa atgggtccatt	3180
aagcaaataa caccacagca	gcaggagaag agtccagtaa	acaccaaata gcttttcaag	3240
cgactttata gccttgcgct	tcacccaat gctttcaaga	ggctgggagc atcacttgcc	3300
tttaataata tctacaggga	attcaggga gaagagtctc	tggtggaaca gtttgtgttt	3360
gaagccttgg tgatatacat	ggagagtctg gccttagcac	atgcagatga gaagtcctta	3420
ggtacaattc aacagtgttg	tgatgccatt gatcacctat	gccgcatcat tgaaaagaag	3480
catgtttctt taaataaagc	aaagaaacga cgtttgccgc	gaggatttcc accttccgca	3540
tcattgtgtt tattggatct	ggtcaagtgg cttttagctc	attgtgggag gccccagaca	3600
gaatgtcgac acaaattcat	tgaactcttt tataaattcg	ttcctttatt gccaggcaac	3660
agatccccta atttgtggct	gaaagatgtt ctcaaggaag	aaggtgtctc ttttctcatc	3720
aacacctttg aggggggtgg	ctgtggccag ccctcgggca	tcctggcca gccaccctc	3780
ttgtaccggg ggccattcag	cctgcaggcc acgctatgct	ggctggacct gtcctggcc	3840
gcgttggagt gctacaacac	gttcattggc gagagaactg	taggagcgct ccaggtccta	3900
ggtactgaag cccagtcttc	acttttgaaa gcagtggctt	tcttcttaga aagcattgcc	3960
atgcatgaca ttatagcagc	agaaaagtgc tttggcactg	gggcagcagg taacagaaca	4020
agcccacaag agggagaaag	gtacaactac agcaaataca	ccgttgtggg ccggattatg	4080
gagtttacca cgactctgct	aaacacctcc ccggaaggat	ggaagctcct gaagaaggac	4140
ttgtgtaata cacacctgat	gagagtcctg gtgcagacgc	tgtgtgagcc cgcaagcata	4200
ggtttcaaca tcggagacgt	ccaggttatg gctcatcttc	ctgatgtttg tgtgaatctg	4260

atgaaagctc taaagatgtc cccatacaaa gatatcctag agacccatct gagagagaaa	4320
ataacagcac agagcattga ggagctttgt gccgtcaact tgtatggccc tgacgcgcaa	4380
gtggacagga gcaggctggc tgctgtttgt tctgcctgta aacagcttca cagagctggg	4440
cttctgcata atatattacc gtctcagtcc acagatttgc atcattctgt tggcacagaa	4500
cttctttccc tggtttataa aggcattgcc cctggagatg agagacagtg tctgccttct	4560
ctagacctca gttgtaagca gctggccagc ggacttctgg agttagcctt tgcttttgga	4620
ggactgtgtg agcgccttgt gagtcttctc ctgaaccag cggtgctgtc cacggcgtcc	4680
ttgggcagct cacagggcag cgtcatccac ttctcccatg gggagtattt ctatagcttg	4740
ttctcagaaa cgatcaacac ggaattattg aaaaatctgg atcttgctgt attggagctc	4800
atgcagtctt cagtggataa taccaaatg gtgagtgccg ttttgaacgg catgttagac	4860
cagagcttca gggagcgagc aaaccagaaa caccaaggac tgaaacttgc gactacaatt	4920
ctgcaacact ggaagaagtg tgattcatgg tgggccaaag attcccctct cgaaactaaa	4980
atggcagtgc tggccttact ggcaaaaatt ttacagattg attcatctgt atctttta	5040
acaagtcattg gttcattccc tgaagtctt acaacatata ttagtctact tgctgacaca	5100
aagctggatc tacatttaaa gggccaagct gtcactcttc ttccattctt caccagcctc	5160
actggaggca gtctggagga acttagacgt gttctggagc agctcatcgt tgctcacttc	5220
cccatgcagt ccagggaatt tcctccagga actccgcggt tcaataatta tgtggactgc	5280
atgaaaaagt ttctagatgc attggaatta tctcaaagcc ctatgttggtt ggaattgatg	5340
acagaagttc tttgtcggga acagcagcat gtcattggaag aattatttca atccagtttc	5400
aggaggattg ccagaagggg ttcatgtgtc acacaagtag gccttctgga aagcgtgtat	5460
gaaatgttca ggaaggatga ccccgcccta agtttcacac gccagtcctt tgtggaccgc	5520
tccctcctca ctctgctgtg gcactgtagc ctggatgctt tgagagaatt cttcagcaca	5580
attgtgggtg atgccattga tgtgttgaag tccaggttta caaagctaaa tgaatctacc	5640
tttgatactc aaatcaccaa gaagatgggc tactataaga ttctagacgt gatgtattct	5700
cgccttccca aagatgatgt tcatgctaag gaatcaaaaa ttaatcaagt tttccatggc	5760
tcgtgtatta cagaaggaaa tgaacttaca aagacattga ttaaattgtg ctacgatgca	5820
tttacagaga acatggcagg agagaatcag ctgctggaga ggagaagact ttaccattgt	5880

gcagcataca actgcgccat atctgtcatc tgctgtgtct tcaatgagtt aaaattttac	5940
caaggttttc tgtttagtga aaaaccagaa aagaacttgc ttatttttga aaatctgac	6000
gacctgaagc gccgctataa ttttcctgta gaagttgagg ttcctatgga aagaaagaaa	6060
aagtacattg aaattaggaa agaagccaga gaagcagcaa atggggattc agatggtcct	6120
tcctatatgt cttccctgtc atatttggca gacagtaccc tgagtgagga aatgagtcaa	6180
tttgatttct caaccggagt tcagagctat tcatacagct cccaagaccc tagacctgcc	6240
actggtcgtt ttcggagacg ggagcagcgg gacccacagg tgcatgatga tgtgctggag	6300
ctggagatgg acgagctcaa tcggcatgag tgcatggcgc ccctgacggc cctggtcaag	6360
cacatgcaca gaagcctggg ccgcctcaa ggagaagagg attcagtgcc aagagatctt	6420
ccttcttgga tgaaattcct ccatggcaaa ctgggaaatc caatagtacc attaaatac	6480
cgtctcttct tagccaagct tgttattaat acagaagagg tctttcgccc ttacgcgaag	6540
cactggctta gccccttgct gcagctggct gcttctgaaa acaatggagg agaaggaatt	6600
cactacatgg tggttgagat agtggccact attctttcat ggacaggctt ggccactcca	6660
acaggggtcc ctaaagatga agtgttagca aatcgattgc ttaatttcct aatgaaacat	6720
gtctttcatc caaaaagagc tgtgtttaga cacaacctg aaattataaa gacccttgct	6780
gagtgctgga aggattgttt atccatccct tataggttaa tatttgaaaa gttttccggt	6840
aaagatccta attctaaaga caactcagta gggattcaat tgctaggcat cgtgatggcc	6900
aatgacctgc ctccctatga ccacagtggt ggcattcaga gtagcgaata cttccaggct	6960
ttggtgaata atatgtcctt tgtaagatat aaagaagtgt atgccgctgc agcagaagtt	7020
ctaggactta tacttcgata tgttatggag agaaaaaaca tactggagga gtctctgtgt	7080
gaactggttg cgaaacaatt gaagcaacat cagaatacta tggaggacaa gtttattgtg	7140
tgcttgaaca aagtgaccaa gagcttcct cctcttcag acaggttcat gaatgctgtg	7200
ttctttctgc tgccaaaatt tcatggagtg ttgaaaacac tctgtctgga ggtggtactt	7260
tgtcgtgtgg agggaatgac agagctgtac ttccagttaa agagcaagga cttcgttcaa	7320
gtcatgagac atagagatga tgaaagacaa aaagtatgtt tggacataat ttataagatg	7380
atgccaaagt taaaaccagt agaactccga gaacttctga accccgttgt ggaattcgtt	7440
tcccatcctt ctacaacatg tagggaacaa atgtataata ttctcatgtg gattcatgat	7500

aattacagag atccagaaag tgagacagat aatgactccc aggaaatatt taagttggca	7560
aaagatgtgc tgattcaagg attgatcgat gagaaccctg gacttcaatt aattattcga	7620
aatttctgga gccatgaaac taggttacct tcaaatacct tggaccgggt gctggcacta	7680
aattccttat attctcctaa gatagaagtg cactttttta gtttagcaac aaattttctg	7740
ctcgaaatga ccagcatgag cccagattat ccaaacccca tgttcgagca tcctctgtca	7800
gaatgcgaat ttcaggaata taccattgat tctgattggc gtttccgaag tactgttctc	7860
actccgatgt ttgtggagac ccaggcctcc cagggcactc tccagaccg taccaggaa	7920
gggtccctct cagctcgctg gccagtggca gggcagataa gggccacca gcagcagcat	7980
gacttcacac tgacacagac tgcagatgga agaagctcat ttgattggct gaccgggagc	8040
agcactgacc cgctggctga ccacaccagt ccctcatctg actccttgct gtttgccac	8100
aagaggagtg aaaggttaca gagagcacc ttgaagtcag tggggcctga ttttgggaaa	8160
aaaaggctgg gccttccagg ggacgaggtg gataacaaag tgaaagggtgc ggccggccgg	8220
acggacctac tacgactgcg cagacggttt atgagggacc aggagaagct cagtttgatg	8280
tatgccagaa aaggcgttgc tgagcaaaaa cgagagaagg aaatcaagag tgagttaaaa	8340
atgaagcagg atgcccaggt cgttctgtac agaagctacc ggcacggaga cttcctgac	8400
attcagatca agcacagcag cctcatcacc ccgttacagg ccgtggccca gagggacca	8460
ataattgcaa aacagctctt tagcagcttg ttttctggaa ttttgaaaga gatggataaa	8520
tttaagacac tgtctgaaaa aaacaacatc actcaaaagt tgcttcaaga cttcaatcgt	8580
tttcttaata ccaccttctc tttctttcca cccttgtct cttgtattca ggacattagc	8640
tgtcagcacg cagccctgct gagcctcgac ccagcggctg ttagcgctgg ttgcctggcc	8700
agcctacagc agcccgtggg catccgcctg ctagaggagg ctctgctccg cctgctgcct	8760
gctgagctgc ctgccaagcg agtccgtggg aaggcccgc tccctcctga tgtcctcaga	8820
tgggtggagc ttgctaagct gtatagatca attggagaat acgacgtcct ccgtgggatt	8880
tttaccagtg agataggaac aaagcaaadc actcagagtg cattattagc agaagccaga	8940
agtgattatt ctgaagctgc taagcagtat gatgaggctc tcaataaaca agactgggta	9000
gatggtgagc ccacagaagc cgagaaggat ttttgggaac ttgcatccct tgactgttac	9060
aaccaccttg ctgagtggaa atcacttgaa tactgttcta cagccagtat agacagtgag	9120

aacccccag acctaaataa aatctggagt gaaccatttt atcaggaaac atatctacct	9180
tacatgatcc gcagcaagct gaagctgctg ctccaggagg aggctgacca gtccctgctg	9240
acatttattg acaaagctat gcacggggag ctccagaagg cgattctaga gcttcattac	9300
agtcaagagc tgagtctgct ttacctcctg caagatgatg ttgacagagc caaatattac	9360
attcaaaatg gcattcagag ttttatgcag aattattcta gtattgatgt cctcttacac	9420
caaagtagac tcaccaaatt gcagtctgta caggctttta cagaaattca ggagttcatc	9480
agctttataa gcaaacaagg caatttatca tctcaagttc cccttaagag acttctgaac	9540
acctggacaa acagatatcc agatgctaaa atggacccaa tgaacatctg ggatgacatc	9600
atcacaaatc gatgtttctt tctcagcaaa atagaggaga agcttaccct tcttcagaa	9660
gataatagta tgaatgtgga tcaagatgga gacccagtg acaggatgga agtgcaagag	9720
caggaagaag atatcagctc cctgatcagg agttgcaagt tttccatgaa aatgaagatg	9780
atagacagtg cccggaagca gaacaatttc tcacttgcta tgaaactact gaaggagctg	9840
cataaagagt caaaaaccag agacgattgg ctggtgagct ggggtgcagag ctactgccgc	9900
ctgagccact gccggagccg gtcccagggc tgctctgagc aggtgctcac tgtgctgaaa	9960
acagtctctt tgttggatga gaacaacgtg tcaagctact taagcaaaaa tattctggct	10020
ttccgtgacc agaacattct cttgggtaca acttacagga tcatagcgaa tgctctcagc	10080
agtgagccag cctgccttgc tgaaatcgag gaggacaagg ctagaagaat cttagagctt	10140
tctggatcca gttcagagga ttcagagaag gtgatcgagg gtctgtacca gagagcattc	10200
cagcacctct ctgaggctgt gcaggcggct gaggaggagg ccagcctcc ctcttgagc	10260
tgtgggcctg cagctggggg gattgatgct tacatgacgc tggcagattt ctgtgaccaa	10320
cagctgcgca aggaggaaga gaatgcatca gttattgatt ctgcagaact gcaggcgtat	10380
ccagcacttg tggaggagaa aatggtgaaa gctttaaaat taaattccaa tgaagccaga	10440
ttgaagtttc ctagattact tcagattata gaacggatatc cagaggagac tttgagcctc	10500
atgacaaaag agatctcttc cgttccctgc tggcagttca tcagctggat cagccacatg	10560
gtggccttac tggacaaaga ccaagccgtt gctgttcagc actctgtgga agaaatcact	10620
gataactacc cgaggctat tgtttatccc ttcatacata gcagcgaaag ctattccttc	10680
aaggatactt ctactggtca taagaataag gagtttgtgg caaggattaa aagtaagttg	10740

gatcaaggag	gagtgattca	agatttttatt	aatgccttag	atcagctctc	taatcctgaa	10800
ctgctcttta	aggattggag	caatgatgta	agagctgaac	tagcaaaaac	ccctgtaaat	10860
aaaaaaaaaca	ttgaaaaaat	gtatgaaaga	atgtatgcag	ccttggggtga	cccaaaggct	10920
ccaggcctgg	gggccttttag	aaggaagttt	attcagactt	ttggaaaaga	atttgataaa	10980
cattttggga	aaggagggttc	taaactactg	agaatgaagc	tcagtgactt	caacgacatt	11040
accaacatgc	tactttttaa	aatgaacaaa	gactcaaagc	cccctgggaa	tctgaaagaa	11100
tgttcaccct	ggatgagcga	cttcaaagtg	gagttcctga	gaaatgagct	ggagattccc	11160
ggtcagtatg	acggtagggg	aaagccattg	ccagagtacc	acgtgcgaat	cgccgggttt	11220
gatgagcggg	tgacagtcac	ggcgtctctg	cgaaggccca	agcgcacatc	catccgtggc	11280
catgacgaga	gggaacaccc	tttcctgggtg	aagggtggcg	aggacctgcg	gcaggaccag	11340
cgcgtggagc	agctcttcca	ggtcacgaat	gggatcctgg	ccaagactc	cgcctgcagc	11400
cagaggggccc	tgacagctgag	gacctatagc	gttgtgcccc	tgacctccag	gttaggatta	11460
attgagtggc	ttgaaaatac	tgttaccttg	aaggaccttc	ttttgaacac	catgtcccaa	11520
gaggagaagg	cggcttacct	gagtgatccc	agggcaccgc	cgtgtgaata	taaagattgg	11580
ctgacaaaaa	tgtcaggaaa	acatgatgtt	ggagcttaca	tgctaatagt	taagggcgct	11640
aatcgtactg	aaacagtcac	gtctttttag	aaacgagaaa	gtaaagtgcc	tgctgatctc	11700
ttaaagcggg	ccttcgtgag	gatgagtaca	agccctgagg	ctttcctggc	gctccgctcc	11760
cacttcgcca	gctctcacgc	tctgatatgc	atcagccact	ggatcctcgg	gattggagac	11820
agacatctga	acaactttat	ggtggccatg	gagactggcg	gcgtgatcgg	gatcgacttt	11880
gggcatgcgt	ttggatccgc	tacacagttt	ctgccagtcc	ctgagttgat	gccttttcgg	11940
ctaactcgcc	agtttatcaa	tctgatgtta	ccaatgaaag	aaacgggcct	tatgtacagc	12000
atcatggtac	acgcactccg	ggccttccgc	tcagaccctg	gcctgctcac	caacaccatg	12060
gatgtgtttg	tcaaggagcc	ctcctttgat	tggaataatt	ttgaacagaa	aatgctgaaa	12120
aaaggagggt	catggattca	agaaataaat	gttgctgaaa	aaaattggta	ccccgcacag	12180
aaaatatgtt	acgctaagag	aaagttagca	ggtgccaatc	cagcagtcac	tacttgtgat	12240
gagctactcc	tgggtcatga	gaaggcccct	gccttcagag	actatgtggc	tgtggcacga	12300
ggaagcaaag	atcacaacat	tcgtgccccaa	gaaccagaga	gtgggctttc	agaagagact	12360

caagtgaagt gcctgatgga ccaggcaaca gacccaaca tccttggcag aacctgggaa	12420
ggatgggagc cctggatgtg aggtctgtgg gagtctgcag atagaaagca ttacattgtt	12480
taaagaatct actatacttt ggttggcagc attccatgag ctgattttcc tgaacacta	12540
aagagaaatg tcttttgtgc tacagtttcg tagcatgagt ttaaatacaag attatgatga	12600
gtaaagtgtg atgggttaaa tcaaagataa ggttatagta acatcaaaga ttaggtgagg	12660
tttatagaaa gatagatata caggcttacc aaagtattaa gtcaagaata taatatgtga	12720
tcagctttca aagcatttac aagtgtgca agttagtga acagctgtct ccgtaaattg	12780
aggaaatgtg gggaagcctt ggaatgccct tctggttctg gcacattgga aagcacactc	12840
agaaggcttc atcaccaaga ttttgggaga gtaaagctaa gtatagttga tgtaacattg	12900
tagaagcagc ataggaacaa taagaacaat aggtaaagct ataattatgg cttatatatta	12960
gaaatgactg catttgatat tttaggatat ttttctaggt tttttccttt cattttattc	13020
tcttctagtt ttgacatttt atgatagatt tgctctctag aaggaaacgt ctttatttag	13080
gagggcaaaa attttgggtca tagcattcac ttttgctatt ccaatctaca actggaagat	13140
acataaaagt gctttgcatt gaatttggga taacttcaaa aatcccatgg ttgttgtag	13200
ggatagtact aagcatttca gttccaggag aataaaagaa attcctattt gaaatgaatt	13260
cctcatttgg aggaaaaaaa gcatgcattc tagcacaaca agatgaaatt atggaataca	13320
aaagtggctc cttcccatgt gcagtccctg tcccccccg ccagtcctcc acacccaaac	13380
tgtttctgat tggcttttag ctttttgttg tttttttttt tccttctaac acttgtattt	13440
ggaggctctt ctgtgatttt gagaagtata ctcttgagtg ttttaataaag tttttttcca	13500
aaagta	13506

<210> 78
 <211> 9027
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >PTEN|ENSG00000171862|ENST00000371953|9027

<400> 78	
ggtaacctca gactcgagtc agtgacactg ctcaacgcac ccattctcagc tttcatcatc	60
agtcctccac ccccgcccca caacagccta ccctgcctcc ggctgggttt ctgggcagag	120

gccgaggctt agctcgttat cctcgcctcg cgttgctgca aaagccgcag caagtgcagc	180
tgcaggctgg cggctgggaa ccggcccgag caagccccag gcagctacac tgggcatgct	240
cagtagagcc tgcggcttgg ggactctgcg ctcgcaccca gagctaccgc tctgccccct	300
cctaccgccc cctgccctgc cctgccctcc cctcgcccgg cgcggtcccg tccgcctctc	360
gctcgcctcc cgcctcccct cggctcttcg aggcgcccgg gctcccggcg cggcggcggg	420
gggggcgggc aggccggcgg gcggtgatgt ggccgggactc tttatgcgct gcggcaggat	480
acgcgctcgg cgctgggacg cgactgcgct cagttctctc ctctcggaag ctgcagccat	540
gatggaagtt tgagagtga gccgctgtga ggcgaggccg ggctcaggcg agggagatga	600
gagacggcgg cggccgcggc ccggagcccc tctcagcgcc tgtgagcagc cgcgggggca	660
gcgccctcgg ggagccggcc ggctgcggc ggccggcagcg gcggcgtttc tcgcctcctc	720
ttcgtctttt ctaaccgtgc agcctcttcc tcggcttctc ctgaaaggga aggtggaagc	780
cgtgggctcg ggccgggagcc ggctgaggcg cggcggcggc ggccggcacct cccgctcctg	840
gagcgggggg gagaagcggc ggccggcggcg gcccgggcgg ctgcagctcc agggaggggg	900
tctgagtcgc ctgtcacat ttccagggt gggaacgccg gagagtgggt ctctcccctt	960
ctactgcctc caacacggcg gcggcggcgg ctggcacatc cagggacccg ggccggtttt	1020
aaacctcccg tgcgccgccg ccgcaccccc cgtggcccgg gctccggagg ccgccggcgg	1080
aggcagccgt tcggaggatt attcgtcttc tccccattcc gctgccgccg ctgccaggcc	1140
tctggctgct gaggagaagc aggcccagtc gctgcaacca tccagcagcc gccgcagcag	1200
ccattacccg gctgcggtcc agagccaagc ggccggcagag cgaggggcat cagctaccgc	1260
caagtccaga gccatttcca tcctgcagaa gaagccccgc caccagcagc ttctgccatc	1320
tctctcctcc tttttcttca gccacaggct cccagacatg acagccatca tcaaagagat	1380
cgtagcaga aaaaaagga gatatcaaga ggatggattc gacttagact tgacctatat	1440
ttatccaaac attattgcta tgggatttcc tgcagaaaga cttgaaggcg tatacaggaa	1500
caatattgat gatgtagtaa ggtttttggg ttcaaagcat aaaaaccatt acaagatata	1560
caatctttgt gctgaaagac attatgacac cgccaaattt aattgcagag ttgcacaata	1620
tccttttgaa gaccataacc caccacagct agaacttatc aaaccctttt gtgaagatct	1680
tgaccaatgg ctaagtgaag atgacaatca tgttgcagca attcactgta aagctggaaa	1740

gggacgaact ggtgtaatga tatgtgcata tttattacat cggggcaaat ttttaaaggc	1800
acaagaggcc ctagatttct atgggggaagt aaggaccaga gacaaaaagg gagtaactat	1860
tcccagtcag aggcgctatg tgtattatta tagctacctg ttaaagaatc atctggatta	1920
tagaccagtg gcaactgttg ttcacaagat gatgtttgaa actattccaa tgttcagtgg	1980
cggaacttgc aatcctcagt ttgtggtctg ccagctaaag gtgaagatat attcctccaa	2040
ttcaggaccc acacgacggg aagacaagtt catgtacttt gagttccctc agccgttacc	2100
tgtgtgtggt gatatcaaag tagagttctt ccacaaacag aacaagatgc taaaaaagga	2160
caaaatgttt cacttttggg taaatacatt cttcatacca ggaccagagg aaacctcaga	2220
aaaagtagaa aatggaagtc tatgtgatca agaaatcgat agcatttgca gtatagagcg	2280
tgcagataat gacaaggaat atctagtact tactttaaca aaaaatgatc ttgacaaagc	2340
aaataaagac aaagccaacc gatacttttc tccaaatttt aaggtgaagc tgtacttcac	2400
aaaaacagta gaggagccgt caaatccaga ggctagcagt tcaacttctg taacaccaga	2460
tgtagtgac aatgaacctg atcattatag atattctgac accactgact ctgatccaga	2520
gaatgaacct tttgatgaag atcagcatac acaaattaca aaagtctgaa ttttttttta	2580
tcaagaggga taaaacacca tgaaaataaa cttgaataaa ctgaaaatgg accttttttt	2640
ttttaatggc aataggacat tgtgtcagat taccagttat aggaacaatt ctcttttcct	2700
gaccaatctt gttttaccct atacatccac agggttttga cacttgttgt ccagttgaaa	2760
aaaggttgtg tagctgtgtc atgtatatac ctttttgtgt caaaaggaca tttaaaattc	2820
aattaggatt aataaagatg gcactttccc gttttattcc agttttataa aaagtggaga	2880
cagactgatg tgtatacgta ggaatttttt ctttttgtgt tctgtcacca actgaagtgg	2940
ctaaagagct ttgtgatata ctggttcaca tcctaccctt ttgcacttgt ggcaacagat	3000
aagtttgcag ttggctaaga gaggtttccg aagggttttg ctacattcta atgcatgtat	3060
tcgggtagg ggaatggagg gaatgctcag aaaggaaata attttatgct ggactctgga	3120
ccatatacca tctccagcta tttacacaca cttttcttta gcatgctaca gttattaatc	3180
tggacattcg aggaattggc cgctgtcact gcttgttgtt tgcgcatttt tttttaaagc	3240
atattggtgc tagaaaaggc agctaaagga agtgaatctg tattggggta caggaatgaa	3300
ccttctgcaa catcttaaga tccacaaatg aagggatata aaaataatgt cataggtaag	3360

aaacacagca acaatgactt aaccatataa atgtggaggc tatcaacaaa gaatgggctt	3420
gaaacattat aaaaattgac aatgatttat taaatatgtt ttctcaattg taacgacttc	3480
tccatctcct gtgtaatcaa ggccagtgct aaaattcaga tgctgttagt acctacatca	3540
gtcaacaact tacacttatt ttactagttt tcaatcataa tacctgctgt ggatgcttca	3600
tgtgctgcct gcaagcttct tttttctcat taaatataaa atattttgta atgctgcaca	3660
gaaattttca atttgagatt ctacagtaag cgtttttttt ctttgaagat ttatgatgca	3720
cttattcaat agctgtcagc cgttccaccc ttttgacctt acacattcta ttacaatgaa	3780
ttttgcagtt ttgcacattt tttaaatgtc attaactgtt agggaatttt acttgaatac	3840
tgaatacata taatgtttat attaaaaagg acatttgtgt taaaaaggaa attagagttg	3900
cagtaaaactt tcaatgctgc acacaaaaaa aagacatttg atttttcagt agaaattgtc	3960
ctacatgtgc tttattgatt tgctattgaa agaatagggt tttttttttt tttttttttt	4020
ttttttttta atgtgcagtg ttgaatcatt tcttcatagt gctccccga gttgggacta	4080
gggcttcaat ttcacttctt aaaaaaatc atcatatatt tgatatgccc agactgcata	4140
cgattttaag cggagtacaa ctactattgt aaagctaag tgaagatatt attaaaaagg	4200
tttttttttc cagaaatttg gtgtcttcaa attatacctt caccttgaca tttgaatatc	4260
cagccatttt gtttcttaat ggtataaaat tccattttca ataacttatt ggtgctgaaa	4320
ttgttcacta gctgtggctt gacctagtta atttacaat acagattgaa taggacctac	4380
tagagcagca tttatagagt ttgatggcaa atagattagg cagaacttca tctaaaatat	4440
tcttagtaaa taatgttgac acgttttcca taccttgta gtttcattca acaattttta	4500
aatttttaac aaagctctta ggatttacac atttatattt aaacattgat atatagagta	4560
ttgattgatt gtcataagt taaattggta aagttagaga caactattct aacacctcac	4620
cattgaaatt tatatgccac cttgtctttc ataaaagctg aaaattgtta cctaaaatga	4680
aaatcaactt catgttttga agatagttat aaatattgtt ctttgttaca atttcgggca	4740
ccgcatatta aaacgtaact ttattgttcc aatatgtaac atggagggcc aggtcataaa	4800
taatgacatt ataatgggct tttgcactgt tattatTTTT ctttgggaat gtgaaggctt	4860
gaatgagggt tttgattttg aatgtttcaa tgtttttgag aagccttgct tacattttat	4920
ggtgtagtca ttggaaatgg aaaaatggca ttatatatat tatatatata aatatatatt	4980

atacatactc tccttacttt atttcagtta ccatcccatc agaatttgac aagaattgct	5040
atgactgaaa ggtttttcgag tcctaattaa aactttatatt atggcagtat tcataattag	5100
cctgaaatgc attctgtagg taatctctga gtttctggaa tattttctta gactttttgg	5160
atgtgcagca gcttacatgt ctgaagttac ttgaaggcat cacttttaag aaagcttaca	5220
gttgggccct gtaccatccc aagtcctttg tagctcctct tgaacatgtt tgccatactt	5280
ttaaaagggt agttgaataa atagcatcac cattctttgc tgtggcacag gttataaact	5340
taagtggagt ttaccggcag catcaaagt ttccagcttta aaaaataaaa gtaggggtaca	5400
agtttaagt ttagttctag aaattttgtg caatatgttc ataacgatgg ctgtggttgc	5460
cacaaagtgc ctggtttacc tttaaatact gttaatgtgt catgcatgca gatggaaggg	5520
gtggaactgt gcactaaagt gggggcctta actgtagtat ttggcagagt tgccttctac	5580
ctgccagttc aaaagttcaa cctgttttca tatagaatat atatactaaa aaatttcagt	5640
ctgttaaaca gccttactct gattcagcct cttcagatac tcttgtgctg tgcagcagtg	5700
gctctgtgtg taaatgctat gcactgagga tacacaaaaa taccaatatg atgtgtacag	5760
gataatgcct catcccaatc agatgtccat ttgttattgt gtttgttaac aaccctttat	5820
ctcttagtgt tataaaactcc acttaaaact gattaaagtc tcattcttgt catttgtgtg	5880
gtgttttatt aaatgagagt ttataattca aattgcttaa gtccattgaa gttttaatta	5940
atgggcagcc aaatgtgaat acaaagtttt cagttttttt ttttcctgct gtccttcaaa	6000
gcctactgtt taaaaaaaaa aaaaaaaaaa aacatggcct gagagtagag tatctgtcta	6060
ctcatgttta attaaagaaa aacacttatt tttagggtt tagtcatcac ttcataaatt	6120
gtataagcac attaaatagc gttctagtcc tgaaaaagtc caagattctt agaaaattgt	6180
gcatattttt attatgacag atgtttgaag ataattcccc agaatggatt tgatacttta	6240
gatttcaatt ttgtggcttt tgtctattat tctgtactct gccatcagca tatggaaagc	6300
ttcatttact catcatgact tgtgccatat aaaaattgat atttcggaat agtctaaagg	6360
actttttgta cttgaattta atcatgttgt ttctaataat cttaaaagct tgaagactaa	6420
agcatatcct ttcaacaaag catagtaagg taataagaaa gtgtagtttg tacaagtgtt	6480
aaaaaaaaaa agtagacaat gttacagtgg gacttattat ttcaagttta cattttctcc	6540
atgtaatttt ttaaaaagta aatgaaaaaa tgtgcaataa tgtaaaatat gaagtgtatg	6600

tgtacacaca	ttttatTTTT	cggtatcttg	ggtatacgta	tggttgaaaa	ctatactgga	6660
gtctaaaagt	attctaattt	ataagaagac	atTTTggtga	tgTTTgaaaa	atagaaatgt	6720
gctagTTTTg	TTTTtatatc	atgtcctttg	tacgttgtaa	tatgagctgg	cttggttcag	6780
taaatgccat	caccatttcc	attgagaatt	taaaactcac	cagtgtttaa	tatgcaggct	6840
tccaaaggct	tatgaaaaaa	atcaagaccc	ttaaacttag	ttaatttgct	gctaacatga	6900
aactctttgg	ttcttttatt	tttgccagat	aattagacac	acatctaaag	cttagtctta	6960
aatggcttaa	gtgtagctat	tgattagtgc	tgTTgctagt	tcagaaagaa	atgtTTgtga	7020
atggaaacaa	gaatattcag	tccaaactgt	tgtaaggaca	gtacctgaaa	accaggaaac	7080
aggataatgg	aaaaagtctt	ttaaagatga	aatgttggag	ccaactttct	tatagaatta	7140
attgtatgtg	gctatagaaa	gcctaattgat	tgTTgcttat	TTTTgagagc	atattattct	7200
tttatgacca	taatcttgct	gtttttccat	cttccaaaag	atcttccttc	taatatgtat	7260
atcagaatgt	gggtagccag	tcagacaaat	tcatatTggt	tggtagcttt	aaaaagTTtg	7320
taatgtgaag	acaggaaagg	acaaaatagt	ttgctTTggt	ggtagtactc	tggttgTTaa	7380
gctaggtatt	ttgagactac	ttcccatca	caacaacaat	aaaataatca	ctcataatcc	7440
tatcacctgg	agacatagcc	atcgTTaata	tgTTagtgc	tatacaatca	tgTTTTcttc	7500
tgtatatcca	tgtatatctt	ttaaaaatga	aatttatact	gtacctgatc	tcaaagcttt	7560
ttagcttagt	atatctgtca	tgaatttgta	ggatgttcca	ttgcatcaga	aaacggacag	7620
tgatttgatt	actttcta	gccacagatg	cagattacat	gtagttattg	agaatccttt	7680
cgaattcagt	ggcttaatca	tgaatgtcta	aatattgttg	acattaggat	gatacatgta	7740
aattaaagtt	acatttgTTt	agcatagaca	agcttaacat	tgtagatgtt	tctcttcaaa	7800
aatcatctta	aacatttgca	tttggaattg	tgTTaaatag	aatgtgtgaa	acactgtatt	7860
agtaaaacttc	atcacctttc	tacttcctta	tagtttgaac	ttttcagttt	ttgtagttcc	7920
caaacagttg	ctcaatttag	agcaaattaa	tttaacacct	gccaaaaaaa	ggctgctgtt	7980
ggcttatcag	ttgtctttaa	attcaaagtc	tcatgtgact	tttatcacat	caaaaaatat	8040
ttcattaatg	attcaccttt	agctctgaaa	attaccgcgt	ttagtaatta	tagtgggctt	8100
ataaaaaacat	gcaactcttt	ttgatagtta	tttgagaatt	ttggtgaaaa	atatttagct	8160
gagggcagta	tagaacttat	aaaccaatat	attgatattt	ttaaaacatt	tttacaatata	8220

agtaaaactgc catcttttgag cataactaca tttaaaaata aagctgcata tttttaaatc	8280
aagtgtttta caagaattta ttttttttat tttttaaaat taaaaataat ttatatattcc	8340
tctgttgcat gaggattctc atctgtgctt ataatggtta gagattttat ttgtgtggaa	8400
tgaagtgagg cttgtagtca tggttctagt gtttcagttt gccaaagtctg tttactgcag	8460
tgaaattcat caaatgtttc agtgtggttt tctgtagcct atcatttact ggctatTTTT	8520
ttatgtacac ctttaggatt ttctgcctac tctatccagt tgtccaaatg atatacctaca	8580
ttttacaaat gccctttcag tttctatTTT ctttttccat taaattgccc tcatgtccta	8640
atgtgcagtt tgtaagtgtg tgtgtgtgtg tctgtgtgtg tgtgaatttg attttcaaga	8700
gtgctagact tccaatttga gagattaaat aattttaattc aggcaaacat ttttcattgg	8760
aatttcacag ttcatgttaa tgaaaatgtt aatcctggat gacctttgac atacagtaat	8820
gaatcttgga tattaatgaa tttgttagta gcatcttgat gtgtgtttta atgagttatt	8880
ttcaaagttg tgcattaaac caaagttggc atactggaag tgtttatatc aagttccatt	8940
tggctactga tggacaaaaa atagaaatgc cttcctatgg agagtatttt tcctttaaaa	9000
aattaaaaag gttaattatt ttgacta	9027

<210> 79
 <211> 3233
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RAD17|ENSG00000152942|ENST00000509734|3233

<400> 79	
ccccgggagg ccgtacctcc gagaggctcg gcgttgagcc cgggtagggc cagggtggctg	60
ccctttcacc tagggtagtc cctggctgcc tccgctcttc gcctaaaagg ggatgcagct	120
ccgggaaagt aaggccgccg cggttgcggc tatattatgt atatgtctta gagacgtgag	180
tctatctctg ctttcaagct ttcctgggct ctcgtcgctc ctctcccga cccgcccac	240
ccatctgggg atgagaagat tgagggtgca gagccctgtc ctgcagcggg gatttgcgag	300
ctcaaccgga caccactg attacaggat tacgttggac gaatatttga gcttagtatt	360
ccctgttcac tgtgtggggg ggtgggtggg cggctaggaa tagtcttgaa ggtctacctc	420

tgacatctca tttcagtaac ctcgcatctt cagggacagt tatctgcttt ttaaaggagc	480
aagtgaattc atgggtatctt acttttttgg gaaatactgg aaatgaagac ctgcaactgt	540
aatttgaaat aaggaaaact ttaattttca gtataaaaat tgctcaaata gaattgcctg	600
attttaatga caaaaggatga attatagttt aatgtactgc aagtcctaaa ctacggatgg	660
gaactattac agttttataat gtcaaaaact tttcttagac caaaggatatc ttccacaaag	720
gtaacagact ggggttgacct atcatttgat gattttctag agtgtagtgg cgtctctact	780
attactgcca catcattagg tgtgaataac tcaagtcata gaagaaaaaa tgggccttct	840
acattagaaa gcagcagatt tccagcgaga aaaagaggaa atctatcttc cttagaacag	900
atttatgggt tagaaaattc aaaagaatat ctgtctgaaa atgaaccatg ggtggataaa	960
tataaaccag aaactcagca tgaacttgct gtgcataaaa agaaaattga agaagtcgaa	1020
acctggttaa aagctcaagt tttagaaagg caacccaaac aggggtggatc tattttatta	1080
ataacaggtc ctctggatg tggaaagaca acgaccttaa aaatactatc aaaggagcat	1140
ggtattcaag tacaagagt gattaatcca gttttaccag acttccaaaa agatgatttc	1200
aaggggatgt ttaatactga atcaagcttc catatgtttc cctatcagtc tcagatagca	1260
gttttcaaag agtttctact aagagcgaca aagtataaca agttacaaat gcttggagat	1320
gatctgagaa ctgataagaa gataattctg gttgaagatt tacctaacca gttttatcgg	1380
gattctcata ctttacatga agttctaagg aagtatgtga ggattggctg atgtcctctt	1440
atatttataa tctcggacag tctcagtgga gataataatc aaaggttatt gtttcccaa	1500
gaaattcagg aagagtgttc tatctcaa attagtttca accctgtggc accaacaatt	1560
atgatgaaat ttcttaatcg aatagtact atagaagcta acaagaatgg aggaaaaatt	1620
actgtccctg acaaaacttc tctagagttg ctctgtcagg gatgttctgg tgatatcaga	1680
agtgcaataa acagcctcca gttttcttct tcaaaaggag aaaacaactt acggccaagg	1740
aaaaaaggaa tgtcttttaa atcagatgct gtgctgtcaa aatcaaacg aagaaaaaa	1800
cctgataggg tttttgaaaa tcaagaggtc caagctattg gtggcaaaga tgtttctctg	1860
tttctcttca gagctttggg gaaaattcta tattgtaaaa gagcatcttt aacagaatta	1920
gactcacctc ggttgccctc tcatttatca gaatatgaac gggatacatt acttggtgaa	1980
cctgaggagg tagtagaaat gtcacacatg cctggagact tatttaattt atatcttcac	2040

caaaactaca tagatttctt catggaaatt gatgatattg tgagagccag tgaatttctg	2100
agttttgcag atatacctcag tgggtgactgg aatacacgct ctttactcag ggaatatagc	2160
acatctatag ctacgagagg tgtgatgcat tccaacaaag cccgaggata tgctcattgc	2220
caaggaggag gatcaagttt tcgacccttg cacaaacctc agtggtttct aataaataaa	2280
aagtatcggg aaaattgcct ggagcaaaa gcactttttc ctgacttctg cctaccagct	2340
ttatgcctcc aaactcagct attgccatac cttgctctac taaccattcc aatgagaaat	2400
caagctcaga tttcttttat ccaagatatt ggaaggctcc ctctgaagcg acactttgga	2460
agattgaaaa tggaagccct gactgacagg gaacatggaa tgatagacc tgacagcgga	2520
gatgaagccc agcttaatgg aggacattct gcagaggaat ctctgggtga acccactcaa	2580
gccactgtgc cggaaacctg gtctcttcct ttgagtcaga atagtgccag tgaactgcct	2640
gctagccagc cccagccctt ttcagcccaa ggagacatgg aagaaaacat aataatagaa	2700
gactacgaga gtgatgggac atagaagcca gcctgctaata cagattgcta cttcacagct	2760
tcatttttgt ttcattcagt ggtacttcag cagagttaat atgcttttct gatgaattac	2820
acaacagttt gttaattctt cattcttgta gtatttcata acaagaaacc tactcttctg	2880
tcattctgaa gtaaatagaa gatcaagcct tcaaatctct taattttttc ggtatttatt	2940
aaatctgtga gtggtttaag gagcggtcag tgtgtataaa gtgtgtttga acattatgcc	3000
aaatatcaag atgtgaagga ctaattcagg atgcaaaaac gttattgggg ggttgtaaat	3060
atcaactatt caacagttta ggatgcaatt acgagtgtaa actgtgtgcc ttatttacac	3120
tttattgtct cccgcttctc agatagtttt gatgtgttgt acagtggaat atcttagata	3180
ctttttggaa agtattttaca taagttatat cacaattaaa atgttgaatt taa	3233

<210> 80

<211> 5886

<212> DNA

<213> Homo sapiens

<220>

<223> >RAD18|ENSG00000070950|ENST00000264926|5886

<400> 80

cccgctctg gcgcgtcgc ggctgctgac gtaatgcgg agcgcgggga atttcgagtg	60
--	----

gtgttgagc gccggaggct agtgggtggc tgacccccag catcctcggg agcgaccatg	120
--	-----

gactccctgg ccgagtctcg gtggcctccg ggcctggcag tcatgaagac aatagatgat	180
ttgctgcggt gtggaatttg cttcgagtat ttcaacattg caatgataat acctcagtgt	240
tcacataact actgctctct ctgtataaga aaatttctgt cctataaaac tcagtgtcca	300
acttgctgtg tgactgtcac agagccgat ctgaaaaata accgcatatt agatgaactg	360
gtaaaaagct tgaattttgc acggaatcat ctgctgcagt ttgctttaga gtcaccagcc	420
aaatctcctg cttcttcctc ttcaaagaat cttgctgtca aagtatatac tcctgtagcc	480
tccagacagt ctttaaagca ggggagcagg ttaatggata atttcttgat cagagaaatg	540
agtggttcta catcagagtt gttgataaaa gaaaataaaa gcaaattcag ccctcaaaaa	600
gaggcgagcc ctgctgcaaa gaccaaagag acacgttctg tagaagagat cgctccagat	660
ccctcagagg ctaagcgtcc tgagccaccc tcgacatcca ctttgaaaca agttactaaa	720
gtggattgtc ctgtttgcgg ggttaacatt ccagaaagtc acattaataa gcatttagac	780
agctgtttat cacgcgaaga gaagaaggaa agcctcagaa gttctgttca caaaaggaag	840
ccgctgccc aactgtata taatttgctc tctgatcgtg atttaaagaa aaagctaaaa	900
gagcatggat tatctattca aggaaataaa caacagctca ttaaaaggca ccaagaattt	960
gtacacatgt acaatgccc atgcgatgct ttgcatccta aatcagctgc tgaaatagtt	1020
cgagaaatcg aaaatataga gaagactagg atgcgtcttg aagctagtaa actcaatgaa	1080
agtgtaatgg tttttacaaa ggaccaaaca gaaaaggaaa tagatgaaat ccacagtaaa	1140
tatcgtaaaa aacataagag tgaatttcag cttctggtgg atcaggctag aaaaggatac	1200
aagaaaattg ctggaatgtc acaaaaaaca gtaacaataa caaaagaaga tgaatctaca	1260
gaaaagctat cttctgtatg catgggacag gaagataata tgacctcagt aacaaaccac	1320
ttttctcaat caaagctgga ctccccagag gaattggaac ctgacagaga agaggattct	1380
tctagctgta ttgatattca agaagttctt tcttcacag aatcagattc atgcaatagt	1440
tccagttcag acatcataag agatctttta gaagaagagg aagcctggga agcatcacat	1500
aaaaacgatc ttcaagacac agaaataagt ccaagacaga atcgccgcac aagagccgct	1560
gaaagtgtg agattgaacc aagaaacaag cgtaatagga attaatgtgg gcttttgctg	1620
acttttcaaa tgcattgatt agaataccgt acttttggtt gccacagata gattttctat	1680
ttataaatgc ccaaggaaag atgctaaatt ctaaataatta cggtttagctg atattcattc	1740

ttttctgctt ttccagaggg gaaaaatgtt acaaaatata cttacttggt cagttgcctc	1800
ctgcctctaa aacatctctc tctaaaaata ctgacatttc acacaggtac cagctttgca	1860
gaggaggtag actctttggc actttggcac agggatttgg tttggtttgg tttggtttga	1920
taattaaatt tcagaatgtc cttaggccat tctccttctc ttccatggag aatccagcct	1980
cacaaaagga ttctttcaac ttctttattg caagaaacag ttgagttaaa tgtttgTTTT	2040
tggaaaggcg ggatgtcagt tcacatcctt ggtgtgtgta agtaccgatg cacgccacca	2100
ccatgcctgt gttacagcag ctccatgatg gttgttgccc gaggttaatg tagttgtttg	2160
ttagacctgt gtcttacaca tttctcagag taggatatta gtattataat ttaaagctac	2220
gacagtcaca aagtcacaat aacttagaaa cattgcgcac attcttctga aatagcactt	2280
aaaaatgatt agtgtcagta ttttttact tgggtcaatc aaatctgtaa cactgaatcc	2340
aagctattaa acaaaaagta tgcaatgaat gaattttgta aatgaataga gagtatcagt	2400
ttacaataat gttcttaaaa cagtatcctc tggatagttg agtatgggtt agaaatcatt	2460
gaaatggatt ggtcagaaat tgctatctgt gtaaaatgtc taccagtagc cagatgcttc	2520
cagagtctt aatgcctctc tggcatttca gagccagcat cccccaactc ccacccctct	2580
gccatcacc aacccaaaca catcagcttt caaatgagat gatagtaaat gcggcaatgt	2640
taagacaaga aatttatgat ttgccagatt caacatttat gacctcccct tccaaagact	2700
gtctccgttg acctgtctt tttggtatgc cttggggttt ctgataatgt gtggagtctc	2760
attatggctg agagtttagt gttttcacag tgaagtgcag acatttgatt tctttatgag	2820
ttccctgtgt tagaaatggc tatagaaaaa tttgtcataa taatttcatt tgcatgaaat	2880
cctgaggggt gcattaagga aactaaaagc accacttacc aaatctatcg gcagaactga	2940
tgtgaggtaa gtgagcatgt caaacaaaat aggagctcac atggatatat ttatgtcact	3000
gagttgtcag aaattatgtc aaaatgaaaa ctgtttgttt catgacaaat tatatagtct	3060
ataaattaaa ctggaagtaa ttattacttt aattgcagca aaaggagttt gtgagggagc	3120
ggtgagacc aagattggga aagtaggcac atgagttcat tcagcaaata tttggttatc	3180
tatgtctgtc actgtgctga cactgggaat acaaaggtgg ccaaagatca tctagaacaa	3240
tggttcccag tggggtgggc agaaagattt tgccccccag gagacagctg gcaatgtgtg	3300
gagacacttt tggaggtgga ggggtgtgag ggggtactacc agtatcaatg ggtggaggcc	3360

agggatgagg ctaaacaccc aaccctcatg gattagtttg ccagggtgc cgtcacaaga	3420
tactgcagac ggggaagctt acacaacaga aatgtgtatt ctcagaattc tggaggctgg	3480
aagtccaaga tcagtaggggt ttcttcttcg gcctctctcc ttggcttcca ctcatggtgt	3540
ccttgcatgg tcttttctct gtgtgcacat gactgtgcag gactggtgtc ccgatttctt	3600
gtagggacag cagtcattgg atcagggtccc atgcgtatgg cctcatttta cttcagttac	3660
ctctttttta gatagagggc cctatctcta aatactgtca cattctgaga taccggcagt	3720
taagacttcc aacatattta gtggtagggg ttaaggggca cgctttagac tacaacactc	3780
cacagtaaag aattatccag tctaacatgt cagttgtgcc aagactgaga aacctgtaat	3840
gtaaaggaac tccaagtct acctgaaaag taaggtctaa agagcattta ttgagcacta	3900
tctgttagca cttagcatat gttaaatatt catacattga aagaaatgta tgttcatatc	3960
catctgacaa agattaagga acttagagta agtaatagat ccaaaattga aactgataaa	4020
cttctgggtc taaaattcac cttttagtga aaatacagtt tctcaccaaa cttacagtaa	4080
ttcagagtta ttaactcttt ctacctctc actcccctca aagttataac atctcccatc	4140
aacatcagct cagaatggca gtgatgctta tggtttgtgg gggtgcaggt gagctgctgg	4200
atgcacttta gatggcctag tatgccaggc cgtcctttca tcctctactc tcgtctttct	4260
ttttcgacct catgttgagg attactggtc tagttgaaaa aagaattcat acacataaga	4320
ccatgtatat aatagaggga ggtatgcctt taaaaacaaa attacatggt aaagactgat	4380
gtgtcaacag ggatagacaa catagagaag gctagaactg tttctggaaa aaaacccaaa	4440
aacaatggat ttcaaacttc ttagcatccg ataaaagaaa gtagcaatta ttcaaatgag	4500
aaacacttct gtctgttatg tacatattat gaaagtatga acataagagg gaaaaccaca	4560
accattacat tttttacat ttcagtaatt ttttttgggt ttttagtctg tttttagaca	4620
tccttaagaa caactatctg aggctgtaaa caagtattta cattaacacc catgatactg	4680
actttatact ccctacctgc attgagacct taaagggggc atttctccat tccttctgtg	4740
tatttctggg tatccccagg ttcagatatt cagaatacag atacaagtcc atgcttgcac	4800
atgttgctgc tcccccgac agacttctac cagcatctgc cctcctcacc gctgctctgt	4860
cacgtgcact gcggctggcc ctgctggaaa gtcctacctt gcctaggacg ccatccccga	4920
tatagcccca agatgcaaac ctgccaagtg ctagggcagc cctgccttcc tgtccatttc	4980

agtctcttgg ccttgctgtg ttccaagtgg aaatgggagt gtgctttcct ggtttgctag	5040
tgtactatca cccggaccac agcgtcctgt gcaaaaattt gcttcctctc tcttcccctt	5100
tcactctttc tcttttctct tccccacccc cttttttttc tctttcccct tacctttctt	5160
tcctttccac tcttcctgtc cttccttcct tcaactatta ggactatatg tcaagtaatg	5220
tgcaaaacaa agacagttcc tgccctgatg gagctttag atcagaaggg aagatgagca	5280
ttaagtaatt atatagatta atgtacaact gtgatgagtg caacaaaaga gaggttcatg	5340
gcactatcaa acatgcaatg gacggatttt gtttttgcct ataaactctt gccatagaaa	5400
gcttgaaaat cagtccaggg ggacagcatt aatcaccatg ttctttcctt atccctgtct	5460
tccccaaaat tcatgtgttg aaacttaaca agaggtggga actttaggat gtgattaagt	5520
cgtgagggca cagctcttat agatgggggtc acggtccttc taaaagggtc tgagggagt	5580
gttttgtttc cttccatccc ttccgccaca tgaggacaca gtcttcgttc cctctggagg	5640
actcagcaac aacacaccat cttggaagca gagagcagtc ctcaccagac acggaatctg	5700
ccagtgcctt gatcttggac ttaccagcct ccatacctgt gagaaataaa tttctattgt	5760
ctataagtta cccagactgt ggtattttct tctggcagca caaatggaca cccagaaact	5820
gatgcatgca tcttctgact cctttgattc caaaatacaa aatattagaa taaagttatg	5880
gaagtc	5886

<210> 81
 <211> 6022
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RAD50|ENSG00000113522|ENST00000265335|6022

<400> 81	
ccaggagagc ggcgtggacg cgtgcgggcc tagaggccca cgtgatccgc agggcggccg	60
aggcaggaag ctgtgagtgc gcggttgcgg ggtcgcattg tggctacggc tttgcgtccc	120
cggcgggcag cccaggctg gtccccgcct ccgctctccc caccggcggg gaaagcagct	180
ggtgtgggag gaaaggctcc atccccgcc ccctctctcc cgctgttggc tggcaggatc	240
ttttggcagt cctgtggcct cgctccccgc ccggatcctc ctgaccctga gattcgcggg	300

tctcacgtcc cgtgcacgcc ttgcttcggc ctacgttaag cctttgtggg ctccaggtcc	360
ctggtgagat tagaaacgtt tgcaaacatg tcccggatcg aaaagatgag cattctgggc	420
gtgcggagtt ttggaataga ggacaaagat aagcaaatta tcactttctt cagccccctt	480
acaattttgg ttggacccaa tggggcggga aagacgacca tcattgaatg tctaaaatat	540
atttgtactg gagatttccc tcctggaacc aaaggaaata ctttgtaca cgatcccaag	600
gttgctcaag aaacagatgt gagagcccag attcgtctgc aatttcgtga tgtcaatgga	660
gaacttatag ctgtgcaaag atctatggtg tgtactcaga aaagcaaaaa gacagaattt	720
aaaactctgg aaggagtcac tactagaaca aagcatggtg aaaaggtcag tctgagctct	780
aagtgtgcag aaattgaccg agaaatgac agttctcttg gggtttccaa ggctgtgcta	840
aataatgtca ttttctgtca tcaagaagat tctaattggc ctttaagtga aggaaaggct	900
ttgaagcaaa agtttcatga gattttttca gcaacaagat acattaaagc cttagaaaca	960
cttcggcagg tacgtcagac acaaggtcag aaagtaaaag aatatcaaag ggaactaaaa	1020
tatctgaagc aatataagga aaaagcttgt gagattcgtg atcagattac aagtaaggaa	1080
gcccagttaa catcttcaaa ggaaattgtc aaatcctatg agaatgaact tgatccattg	1140
aagaatcgtc taaaagaaat tgaacataat ctctctaaaa taatgaaact tgacaatgaa	1200
attaaagcct tggatagccg aaagaagcaa atggagaaag ataatagtga actggaagag	1260
aaaatggaaa aggtttttca agggactgat gagcaactaa atgacttata tcacaatcac	1320
cagagaacag taaggagaaa agaaaggaaa ttggtagact gtcacgtga actggaaaaa	1380
ctaaataaag aatctaggct tctcaatcag gaaaaatcag aactgcttgt tgaacagggt	1440
cgtctacagc tgcaagcaga tcgccatcaa gaacatatcc gagctagaga ttcattaatt	1500
cagtctttgg caacacagct agaattggat ggctttgagc gtggaccatt cagtgaagaa	1560
cagattaaaa attttcacia acttgtgaga gagagacaag aagggaagc aaaaactgcc	1620
aaccaactga tgaatgactt tgcagaaaaa gagactctga aacaaaaaca gatagatgag	1680
ataagagata agaaaactgg actgggaaga ataattgagt taaaatcaga aatcctaagt	1740
aagaagcaga atgagctgaa aaatgtgaag tatgaattac agcagttgga aggatcttca	1800
gacaggattc ttgaactgga ccaggagctc ataaaagctg aacgtgagtt aagcaaggct	1860
gagaaaaaca gcaatgtaga aaccttaaaa atggaagtaa taagtctcca aatgaaaaa	1920

gcagacttag acaggaccct gcgtaaactt gaccaggaga tggagcagtt aaaccatcat	1980
acaacaacac gtacccaaat ggagatgctg accaaagaca aagctgacaa agatgaacaa	2040
atcagaaaaa taaaatctag gcacagtgat gaattaacct cactgttggg atattttccc	2100
aacaaaaaac agcttgaaga ctggctacat agtaaataca aagaaattaa tcagaccagg	2160
gacagacttg ccaaattgaa caaggaacta gcttcatctg agcagaataa aaatcatata	2220
aataatgaac taaaaagaaa ggaagagcag ttgtccagtt acgaagacaa gctgtttgat	2280
gtttgtggta gccaggattt tgaaagtgat ttagacaggc tttaaagagga aattgaaaaa	2340
tcatcaaaac agcgagccat gctggctgga gccacagcag ttactcca gttcattact	2400
cagctaacag acgaaaacca gtcattgttc cccgtttgtc agagagtttt tcagacagag	2460
gctgagttac aagaagtcac cagtgtttg cagtctaaac tgcgacttgc tccagataaa	2520
ctcaagtcaa cagaatcaga gctaaaaaaa aaggaaaagc ggcgtgatga aatgctggga	2580
cttgtgcca tgaggcaaag cataattgat ttgaaggaga aggaaatacc agaattaaga	2640
aacaaactgc agaattgtcaa tagagacata cagcgcctaa agaacgacat agaagaacaa	2700
gaaacactct tgggtacaat aatgcctgaa gaagaaagtg ccaaagtatg cctgacagat	2760
gttacaatta tggagagggt ccagatggaa cttaaagatg ttgaaagaaa aattgcacaa	2820
caagcagcta agctacaagg aatagactta gatcgaactg tccaacaagt caaccaggag	2880
aaacaagaga aacagcacia gttagacaca gtttctagta agattgaatt gaatcgtaag	2940
cttatacagg accagcagga acagattcaa catctaaaaa gtacaacaaa tgagctaaaa	3000
tctgagaaac ttcagatatc cactaatttg caacgtcgtc agcaactgga ggagcagact	3060
gtggaattat ccactgaagt tcagtctttg tacagagaga taaaggatgc taaagagcag	3120
gtaagccctt tggaaacaac attggaaaag ttccagcaag aaaaagaaga attaatcaac	3180
aaaaaaaaata caagcaacaa aatagcacag gataaactga atgatattaa agagaagggt	3240
aaaaatattc atggctatat gaaagacatt gagaattata ttcaagatgg gaaagacgac	3300
tataagaagc aaaaagaaac tgaacttaat aaagtaatag ctcaactaag tgaatgcgag	3360
aaacacaaag aaaagataaa tgaagatatg agactcatga gacaagatat tgatacacag	3420
aagatacaag aaagggtggct acaagataac cttactttta gaaaaagaaa tgaggaacta	3480
aaagaagttg aagaagaaag aaaacaacat ttgaaggaaa tgggtcaa at gcaggttttg	3540

caaatgaaaa gtgaacatca gaagttggaa gagaacatag acaatataaa aagaaatcat	3600
aatttggcat tagggcgaca gaaaggttat gaagaagaaa ttattcattt taagaaagaa	3660
cttcgagaac cacaatttcg ggatgctgag gaaaagtata gagaaatgat gattgttatg	3720
aggacaacag aacttgtgaa caaggatctg gatattttatt ataagactct tgaccaagca	3780
ataatgaaat ttcacagtat gaaaatggaa gaaatcaata aaattatacg tgacctgtgg	3840
cgaagtacct atcgtggaca agatattgaa tacatagaaa tacgggtctga tgccgatgaa	3900
aatgtatcag cttctgataa aaggcggaat tataactacc gagtggtgat gctgaaggga	3960
gacacagcct tggatatgcg aggacgatgc agtgctggac aaaaggtatt agcctcactc	4020
atcattcgcc tggccctggc tgaaacgttc tgcctcaact gtggcatcat tgccttggat	4080
gagccaacaa caaatcttga ccgagaaaac attgaatctc ttgcacatgc tctggttgag	4140
ataataaaaa gtcgctcaca gcagcgtaac ttccagcttc tggtaatcac tcatgatgaa	4200
gattttgtgg agcttttagg acgttctgaa tatgtggaga aattctacag gattaaaaag	4260
aacatcgatc agtgctcaga gattgtgaaa tgcagtgtta gctccctggg attcaatgtt	4320
cattaaaaat atccaagatt taaatgccat agaaatgtag gtcctcagaa agtgtataat	4380
aagaaactta tttctcatat caacttagtc aataagaaaa tatattcttt caaaggaaca	4440
tttgttctag gattttggat gttgagaggt tctaaaatca tgaaacttgt ttcactgaaa	4500
attggacaga ttgcctgttt ctgatttgct gctcttcac c cattccagg cagcctctgt	4560
caggccttca gggttcagca gtacagccga gactcgactc tgtgcctccc tcccagtg	4620
aaatgcatgc ttcttctcaa agcactgttg agaaggagat aattactgcc ttgaaaattt	4680
atggtttttg tattttttta aatcatagtt aaatgttacc tctgaattta cttccttgca	4740
tgtggtttga aaaactgagt attaatatct gaggatgacc agaaatggtg agatgtatgt	4800
ttggctctgc ttttaacttt ataaatccag tgacctctct ctctgggact tggtttcccc	4860
aactaaaatt tgaagtagtt gaatggggtc tcaaagtttg acaggaacct taagtaatca	4920
tctaagtcag taccaccac cttcttctcc tacatatccc ttccagatgg tcatccagac	4980
tcagagctct ctctacagag aggaaattct cactgtgca caccacctt tggaaagctc	5040
tgaccacttg aggctgatc tgcccatcgt gaagaagcct gtaacactcc tctgctgcta	5100
tcctgtgtag catactggct tcaccatcaa tcctgattcc tctctaagtg ggcattgcca	5160

tgtggaaggc aagccaggct cactcacaga gtcaaggcct gtcacctgta gggccaacc	5220
agacctggaa gaacaggcct ctccatttgc tcttcagatg ccacttctaa gaaaagccta	5280
atcacagttt ttcctggaat tgccagctga catcttgaat ccttccattc cacacagaat	5340
gcaaccaagt cacacgcttt tgaattatgc tttgtagagt tttgtcattc agagtcagcc	5400
aggaccatac cgggtcttga ttcagtcaca tggcatgggtt ttgtgccatc ttagctata	5460
atgagcatgt ttgcctagac agcttttctc aactgggtcc agaagagaat taagccctaa	5520
ggtcctaagg catctatctg tgctagggtta aatgggtggc ccccaaagat agacagggtcc	5580
tgatttctag aaccctgac tgttacttta tacagcaaag gaaactttgc agatgtgatt	5640
aaagctaagg accttaagac agagtatcct ggggggtggg gtgggggtggg ggggggtcct	5700
aaatgtaatc acgagtaaga ttaagagcaa atcaattcta gtcatatatt aaacatccac	5760
aataaccaag atatttttat cccaagaatg caagatttca gaaaatgaaa aatctgttga	5820
taaatccatc actataataa aaccgaagggt gaaaaaaatt ctgaaaaaat tctagcagct	5880
atatttgata aaattcaaca tctcctagct ttagcaaact cacagttttg caaataatat	5940
tttcttaatg ttatctgttg ctaaataaaa attaaacagt catcttaact gcaaaataaa	6000
acatttctca gtaaatatta aa	6022

<210> 82

<211> 1588

<212> DNA

<213> Homo sapiens

<220>

<223> >RAD51|ENSG0000051180|ENST00000382643|1588

<400> 82

aattctgaaa gccgctggcg gaccgcgcgc agcggccaga gaccgagccc taaggagagt	60
gcggcgcttc ccgaggcgtg cagctgggaa ctgcaactca tctgggttgt gcgcagaagg	120
ctggggcaag cgagtagaga agtggagcta atggcaatgc agatgcagct tgaagcaaat	180
gcagatactt cagtggaaga agaaagcttt ggcccacaac ccatttcacg gttagagcag	240
tgtggcataa atgccaacga tgtgaagaaa ttggaagaag ctggattcca tactgtggag	300
gctgttgctt atgcgcaaaa gaaggagcta ataaatatta agggaattag tgaagccaaa	360
gctgataaaa ttctgacgga gtctcgctct gttgccaggc tggagtgcaa tagcgtgatc	420

ttggtctact gcaccctccg cctctcaggt tcaagtgatt ctctgcctc agcctcccga	480
gtagttggga ctacaggtgg aattgagact ggatctatca cagaaatggt tggagaattc	540
cgaactggga agaccagat ctgtcatagc ctagctgtca cctgccagct tcccattgac	600
cggggtggag gtgaaggaaa ggccatgtac attgacactg agggtagctt taggccagaa	660
cggctgctgg cagtggctga gaggtatggg ctctctggca gtgatgtcct ggataatgta	720
gcatatgctc gagcgttcaa cacagaccac cagaccagc tcctttatca agcatcagcc	780
atgatggtag aatctaggta tgcactgctt attgtagaca gtgccaccgc cttttacaga	840
acagactact cgggtcgagg tgagctttca gccaggcaga tgcacttggc caggtttctg	900
cggatgcttc tgcgactcgc tgatgagttt ggtgtagcag tggtaatcac taatcaggtg	960
gtagctcaag tggatggagc agcgatgttt gctgctgac ccaaaaaacc tattggagga	1020
aatatcatcg cccatgcatc aacaaccaga ttgtatctga ggaaaggaag aggggaaacc	1080
agaatctgca aaatctacga ctctccctgt cttcctgaag ctgaagctat gttcgccatt	1140
aatgcagatg gagtgggaga tgccaaagac tgaatcattg ggtttttcct ctgttaaaaa	1200
ccttaagtgc tgcagcctaa tgagagtgc ctgctccctg gggttctcta caggcctctt	1260
cctgttgatga ctgccaggat aaagcttccg ggaaaacagc tattatatca gcttttctga	1320
tggtataaac aggagacagg tcagtagtca caaactgatc taaaatgttt attccttctg	1380
tagtgtatta atctctgtgt gttttctttg gttttggagg aggggtatga agtatctttg	1440
acatggtgcc ttaggaatga cttgggttta acaagctgtc tactggacaa tcttatgttt	1500
ccaagagaac taaagctgga gagacctgac ctttctctca cttctaaatt aatggtaaaa	1560
taaaatgcct cagctatgta gcaaaggg	1588

<210> 83

<211> 2710

<212> DNA

<213> Homo sapiens

<220>

<223> >RAD52|ENSG0000002016|ENST00000358495|2710

<400> 83

ttcccagaga cgctcgcga cccttcccat tctcctctgc gcggcctcca tctaagatct	60
--	----

cttcccccttg tccatagcct agatcgagct ccctgtgtgc accgcgcgct gcccgaggcg	120
cagggtcaacc agaatacaaga tgtctgggac tgaggaagca attcttggag gacgtgacag	180
ccatcctgct gctggcggcg gctcagtgtt atgctttgga cagtgccagt acacagcaga	240
agagtaccag gccatccaga aggccctgag gcagaggctg ggcccagaat acataagtag	300
ccgcatggct ggcggaggcc agaaggtgtg ctacattgag ggtcatcggg taattaatct	360
ggccaatgag atgtttgggtt acaatggctg ggcacactcc atcacgcagc agaatgtgga	420
ttttgttgac ctcaacaatg gcaagttcta cgtgggagtc tgtgcatttg tgagggtcca	480
gctgaaggat ggttcatatc atgaagatgt tggttatggt gttagtgagg gcctcaagtc	540
caaggcttta tctttggaga aggcaaggaa ggaggcgggtg acagacgggc tgaagcgagc	600
cctcaggagt tttgggaatg cacttggaat ctgtattctg gacaaagact acctgagatc	660
actaaataag cttccacgcc agttgcctct tgaagtggat ttaactaaag cgaagagaca	720
agatcttgaa ccgtctgtgg aggaggcaag atacaacagc tgccgaccga acatggccct	780
gggacacca cagctgcagc aggtgacctc cccttccaga cccagccatg ctgtgatacc	840
ggcggaccag gactgcagct cccgaagcct gagctcatcc gccgtggaga gcgaggccac	900
gcaccagcgg aagctccggc agaagcagct gcagcagcag ttccgggagc ggatggagaa	960
gcagcaggtt cgagtctcca cgccgtcagc tgagaagagt gaggcagcgc ctccggcccc	1020
tcctgtgacg cacagcactc ctgtaactgt ctcagaacca ctcttgagaa aagacttcct	1080
tgcaggagtg actcaagaat taatcaagac tcttgaagac aactctgaaa agtgggctgt	1140
gactcccgat gcaggggatg gtgtgggtcaa gccctcgtct agagcagacc cagcccagac	1200
ctctgacaca ttagccttga acaaccagat ggtgaccag aacaggactc cacacagcgt	1260
ttgccaccag aaaccacaag caaatctgg atcttgggac ctccaaactt atagcgtga	1320
ccaacgcaca acaggaaact gggaatctca taggaagagc caggacatga agaaaaggaa	1380
atatgatcca tcttaactga ggctcaggcc acataattgg actctgtcac aaagggactt	1440
tggaaaacta ctttttgggtc atgaaattgt tcatcgctgc tggagaatga acgtcattgc	1500
gatttatctt gcttcattct gaaccttatc aagaggatct gactgagagc cactgcagt	1560
tagagctgag cacttttgaa aagcttgtcc atcactctag tagggagagg ctctggacag	1620
atgaatacct tttcttcggc ttgtgaggct tcccactatt tattactgaa ctattatgtt	1680

aatgaagatg gacatttttag gaatcaccaa tggctccttg ccctcaagca atataggcca	1740
gacttgggtcc taagcacctg cctcagcaat tgtctacatt cagttgtttt gcataacgtc	1800
tgcctttcttt cttttacggt ccatgccttt aatgttgccc acattaagca ctgtggatca	1860
cgacaggaaa aagggtggag cagtgccttt cactactttg tatcaatcca ggctacaatc	1920
ttcatttaat ataaataatt tatggattta tgacattaca atcctgcatt gtttcaagac	1980
tgacattttt tcctaaggaa ggaaataatc atctaagacc acgaaaaaag gctgtttttt	2040
gttttttttt tttttttttt ttttgagacg gggctctggct gtgttgccct gactggagtt	2100
cagtggtgca aacacagctc tctccacaac ctcttgggcc caagtgatac tcccacctct	2160
gccttacaaa atacagggat tactgggtgtg agccactgtg tctggccaga aaaggcattt	2220
ttgagaaagc aaatcgtata ccttattaac aaaatagaat atatatatat tgcttatctg	2280
aaatgcttga aaccagaatt gttttgcatt ttttgaatat ttgtatacac ataatgagac	2340
cttgggggatg ggaccaagt ctgaacgtgg aattcacctg tgtttcgtgt atatgcctca	2400
tacacataat tttgtgcatg aaacagagtt tttgtataag aagatacact gcagctgaag	2460
agggctgggt ttttttttct cttagggtcg ctgcataaac tgttgtatgc ctgggtgcttt	2520
gcgacttgct acacgaggtc acgtgtggaa ttttccactt ctggcatcac gtcagtgtc	2580
agaaattttc tgatctcaga gcatttcaat tagggatgct caaacgcaac tgtttctact	2640
tccccatttc aggtgtgaga tgtaaccac cttgaccata aattggcttt tcatagtgt	2700
cagatgtttc	2710

<210> 84
 <211> 3068
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RAD54B|ENSG00000197275|ENST00000336148|3068

<400> 84	
gattggcttc gccgaggcgg gatccttgag cttctccggc ggcgagggga tagctggtta	60
ccagaaggac ttctttgcag ggccagtggg ttctgtcaga ttttcgccgg tcacgactgc	120
tgaatatgag acgatctgca gcaccaagtc agttgcaggg gaattccttc aaaaaaccaa	180
aatttatacc tccaggaaga agtaatccag gtctgaatga agagattaca aaactgaatc	240

cagatataaa attatttgag ggtgttgcaa ttaataacac ctttctcccg tcacaaaatg	300
atcttagaat atgcagttta aatctgccta gtgaagaaag tactagagaa atcaataaca	360
gagataattg cagtggaaaa tattgttttg aagcacctac actggcaaca ttagatccac	420
ctcatacagt tcattcggct cctaaagaag tagcagtgtc caaggaacaa gaagagaaat	480
ctgatagcct agttaaatat ttcagtgttg tttggtgtaa gccttcaaag aaaaaacata	540
aaaagtggga aggtgatgct gttcttattg taaaaggaaa gtcatttata ttaaagaatt	600
tggaaggcaa agacattgga agaggcattg gttataaatt caaagagctt gaaaagattg	660
aagagggcca aacactgatg atttgtggaa aagaaataga agtcatgggt gtaatctctc	720
cagatgactt cagcagtggc aggtgttttc agcttggagg aggaagtact gctatctcgc	780
attctttctca ggttgccagg aaatgtttct ctaacccttt caaaagtgtt tgtaaaccaa	840
gttcaaagga aaatagacag aatgatttcc aaaattgcaa accacgccat gacctatata	900
cgccaaattc cctcgttatg ccacgaccag ataagaatca ccagtgggta ttcaataaga	960
actgtttccc tcttgtggat gtagtgattg atccttacct tgtatatcat cttcgaccac	1020
atcagaaaga aggaatcata ttcctttatg aatgtgtaat gggaatgaga atgaatggca	1080
gatgtggagc tattcttgct gatgaaatgg gtttagggaa gacattgcaa tgtatttcgc	1140
tcatctggac cctgcagtgt cagggaccct atggaggcaa gccagtaata aagaagacac	1200
taattgtcac acctggaagc ttggtgaata attggaagaa agaatttcaa aaatggctag	1260
gaagtgaaag gatcaagata tttactgttg atcaggacca caaagttgaa gaattcatca	1320
agtctatatt ttattctgtt cttattatca gttatgaaat gttacttcgt tccctggatc	1380
aaattaagaa tataaaattt gatcttctaa tctgtgacga ggggcatcgt ttgaagaaca	1440
gtgccattaa gacaactaca gccctcatta gcctttcttg tgagaaaaga ataattctaa	1500
ctgggtactcc aattcagaat gatctgcaag aatTTTTTgc attaattgat tttgtaaadc	1560
caggaatatt aggctctttg tcactttata ggaaaatata tgaagaaccc atcattttat	1620
cgagagaacc ttctgcttct gaggaagaaa aggagttagg agaaagaaga gcagctgaac	1680
ttacttgcct cactggactc tttatcctta gaagaacca agaaattata aataaatatc	1740
tcccacctaa aatagagaat gttgtctttt gccgaccagg agcactacag attgagcttt	1800
atcgaaagct gttaaattct caggttgtca ggttctgcct tcaagggttg ttggaaaata	1860

gtcccatct aatatgtata ggagctctta aaaaactgtg caatcacccc tgccttttgt	1920
tcaactctat aaaggaaaag gaatgtagct caacttgtga taaaaatgaa gaaaagagtc	1980
tatacaaagg cttgctaagt gtgtttcctg ctgactacaa ccctctcctg tttactgaaa	2040
aggagtcagg aaaactacag gtgttgtcca agctcttagc ggttatccac gaacttcgac	2100
ctactgaaaa ggtgggtgtg gtatccaact atacacaaac cttgaacatt ttacaagaag	2160
tatgtaagcg tcatggatat gcttatacaa gacttgatgg acaaacacca atctctcaaa	2220
ggcagcagat tgttgatggc tttaacagtc aacactcttc tttttttatt tttttgttaa	2280
gttcaaaagc tgggtggtgta ggacttaacc tcattggagg atctcactta attctctatg	2340
acattgattg gaatccagcc actgacattc aggcaatgtc tagagtatgg agagatggtc	2400
agaaatatcc tgtacatatt tacagactcc taactacagg tacaatagaa gaaaagatct	2460
atcaaaggca gatcagtaag caaggtcttt gtggggcagt tgtcgacctc accaagacat	2520
ctgaacatat tcagttttca gtagaagaac ttaaaaattt gttcacatta catgaaagtt	2580
cagattgtgt tactcatgat ctgcttgact gtgagtgtac aggagaagaa gttcatacag	2640
gtgattcgtt ggaaaaattc attgtctcta gagattgtca gcttggtcca catcaccaga	2700
aatctaactc cctgaaacct ctttctatgt cccagctgaa gcaatggaaa catttttctg	2760
gagatcattt aaatcttaca gatccttttc ttgaaagaat aacagaaaaat gtgtcattca	2820
tttttcagaa tataaccact caagctactg gcacatagtg aaagattact tctgacattc	2880
cattgctctt cttttgaaaa ttagtatggg aattaaatgt actttttgaa aattaaataga	2940
attatttaaa ttacagtata tgttgcaaaa tatatcactt ttgatacaat agtcaaaatt	3000
gagtggttta atgttttgta aatattaagt gtttaaatga aaaataaaga tgtgcttata	3060
tcattgta	3068

<210> 85
 <211> 3108
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RAD54L|ENSG00000085999|ENST00000371975|3108

<400> 85

ggctcttggcg ggtcgggtgag tcttggcggc tgtaaacgcg cgctttggga acaggaaggt	60
tgagagagag gtgctgggggt ctgcgtctat ctctgtcgtc cttttcagcc cctcctggta	120
ttccccctcct aacctggggtt ttttacacgc ccgcgtggct tcctgctcga cctccctgag	180
tctgatcctg gtttccacct ccagccctgg gaaatttcct ttctccagac tcgccctccc	240
cacccgggcc tcggactttc accccagctt ctctctcctg gccagtgatt acccaccccc	300
aatcccaccc cgccccgccg cgcaactacc tcctcccttc acccggactg ggaccatcat	360
ccccactcca ctccgccag tctgggactc cacctgcctc ctccccaatc ccacactaat	420
ctctgcttgg tctcttcctc tttggcctaa tctctcgtct cggcttattg gggacggcca	480
ctctcacagt ttggttccaa acaccagttc ctggatggat tcccgccatc catgccccct	540
ctttaattag ccggtcctct caataatgta gcagccccct ctacagatta gaccctggtc	600
ctacactctt agccgctgcc tgcttttgac ctttggctca tgggtacttg acgttttaaa	660
ctcctaggcc caggatgagg aggagcttgg ctcccagcca gctggccaag agaaaacctg	720
aaggcaggtc ctgtgatgat gaagactggc aacctggcct agtgactcct aggaaacgga	780
aatccagcag tgagaccag atccaggagt gtttcctgtc tccttttcgg aaacctttga	840
gtcagctaac caatcaacca ccttgtcttg acagcagtca gcatgaagca tttattcgaa	900
gcattttgtc aaagcctttc aaagtcccca ttccaaatta tcaaggtcct ctgggctctc	960
gagcattggg cctgaaaagg gctgggggtcc gccgggcctt ccatgacccc ctggaaaaag	1020
atgccttgggt tctgtatgag cctccccgc tgagcgtca tgaccagctg aagcttgaca	1080
aggagaaact ccctgtccat gtggttgttg accctattct cagtaagggt ttgcggcctc	1140
atcagagaga gggagtgaat ttcctgtggg agtgtgtcac cagtcggcgc atccctggca	1200
gccatggctg catcatggct gatgagatgg gcctaggaaa gacgctgcag tgcatacat	1260
tgatgtggac acttttacgc cagagtccag agtgcaagcc agaaattgac aaggcagtgg	1320
tggtgtcgcc ttccagcctg gtgaagaact ggtacaatga gggttgggaaa tggctcggag	1380
ggaggatcca acctctggcc atcgatggag gatctaagga tgaaatagac caaaagctgg	1440
aaggattcat gaaccagcgt ggagccaggg tgtcttctcc catcctcatc atttcctatg	1500
agaccttccg ctttcatgtt ggagtcctcc agaaaggaag tgttgggtctg gtcatatgtg	1560
acgagggaca caggctcaag aactctgaga atcagactta ccaagccctg gacagcttga	1620

acaccagccg gcgggtgctc atctccggaa ctcccatcca gaatgatctg cttgagtatt	1680
tcagcttggg acattttggt aattccggca tcctaggagc tgcccatgaa ttcaagaagc	1740
attttgaatt gccaatTTTg aagggtcgag acgctgctgc tagtgaggca gacaggcagc	1800
taggagagga gcggctgcgg gagctcacca gcattgtgaa tagatgcctg atacggagga	1860
cttctgatat cttttctaaa tatctgcctg tgaagattga gcaggtcgtt tgTTgtaggc	1920
tgacaccctt tcagactgag ttatacaaga ggtttctgag acaagccaaa ccggcagaag	1980
aattgcttga gggcaagatg agtgtgtctt ccctttcttc catcacctcg ctaaagaagc	2040
tttgtaatca tccagctcta atctatgata agtgtgtgga agaggaggat ggctttgtgg	2100
gtgccttggc cctcttccct cctggttaca gctctaaggc cctggagccc cagctgtcag	2160
gtaagatgct ggtcctggat tatattctgg cggtgacccg aagccgtagc agtgacaaag	2220
tagtgctggg gtcgaattac acccagactt tggatctctt tgagaagctg tgccgtgccc	2280
gaagg tactt atacgtccgc ctggatggca cgatgtccat taagaagcga gccaaagttg	2340
tagaacgctt caatagtcca tcgagccctg actttgtctt catgctgagc agcaaagctg	2400
ggggctgtgg cctcaatctc attggggcta accggctggg catgtttgac cctgactgga	2460
accagccaa tgatgaacaa gccatggccc gggctctggcg agatgggtcaa aagaagactt	2520
gctatatcta ccgcctgctg tctgcaggga ccattgagga gaagatcttc cagcgtcaga	2580
gccacaagaa ggactgagc agctgtgtgg tggatgagga gcaggatgta gagcgccact	2640
tctctctggg cgagttgaag gagctgttta tcctggatga agctagcctc agtgacacac	2700
atgacagggt gcactgccga cgttgtgtca acagccgtca gatccggcca cccctgatg	2760
gttctgactg cacttcagac ctggcagggt ggaaccactg cactgataag tgggggctcc	2820
gggatgagggt actccaggct gcctgggatg ctgcctccac tgccatcacc ttcgtcttcc	2880
accagcgttc tcatgaggag cagcggggcc tccgctgata accagctggg ctgggtgtag	2940
ctcttagagg aaggagatag ggaaaagggg ctccttgctc cacagggcc tgttgaattt	3000
tgttctctgg gagaaaatca tcaagaaggg ctgcatgatg ttgccccaaa atttatttta	3060
taagaaaaac ttttttgggt aaaaaaaga ataaagggtat gaaagggt	3108

<210> 86

<211> 2119

<212> DNA

<213> Homo sapiens

<220>

<223> >RAD9A|ENSG00000172613|ENST00000307980|2119

<400> 86

gggccggcag gggcgggtgcg cgggaaggga ccccggaacc ggaggtcgcg gagagctggg	60
cagtgttggc cgctggcgga gcgctggggc agcatgaagt gcctggtcac gggcggcaac	120
gtgaagggtgc tcggcaaggc cgtccactcc ctgtcccga tcggggacga gctctacctg	180
gaacccttgg aggacgggct ctccctccgg acggtgaact cctcccgctc tgcctatgcc	240
tgctttctct ttgccccgct cttcttccag caataccagg cagccacccc tggtcaggac	300
ctgctgcgct gtaagatcct gatgaagtct ttcctgtctg tcttcgctc actggcgatg	360
ctggagaaga cgggtgaaaa atgctgcatc tccctgaatg gccggagcag ccgcctgggtg	420
gtccagctgc attgcaagtt cgggggtgcgg aagactcaca acctgtcctt ccaggactgt	480
gagtccttgc aggccgtctt cgaccagcc tcgtgcccc acatgctccg cggcccagca	540
cgggttcttg gggaggctgt tctgcccttc tctcctgcac tggctgaagt gacgctgggc	600
attggccgtg gccgcagggt catcctgcgc agctaccacg aggaggaggc agacagcact	660
gccaaagcca tggtgactga gatgtgcctt ggagaggagg atttccagca gctgcaggcc	720
caggaagggg tggccatcac tttctgcctc aaggaattcc gggggctcct gagctttgca	780
gagtcagcaa acttgaatct tagcattcat tttgatgctc caggcaggcc cgccatcttc	840
accatcaagg actctttgct ggacggccac tttgtcttgg ccacactctc agacaccgac	900
tcgactccc aggacctggg ctccccagag cgtcaccagc cagtgcctca gctccaggct	960
cacagcacac cccaccgga cgactttgcc aatgacgaca ttgactctta catgatcgcc	1020
atggaaacca ctataggcaa tgagggtctg cgggtgctgc cctccatttc cttttcacct	1080
ggccccagc ccccaagag ccccggtccc cactccgagg aggaagatga ggctgagccc	1140
agtacagtgc ctgggactcc cccaccaag aagttccgct cactgttctt cggctccatc	1200
ctggcccctg tacgctcccc ccagggcccc agccctgtgc tggcggaaga cagtgagggt	1260
gaaggctgaa ccaagaacct gaagcctgta cccagaggcc ttggactaga cgaagcccca	1320
gccagtggca gaactgggtc tctcagccct ggggatcaga aaggtgggct tgctggagct	1380
gagctgtttc actgcctctc gcaggcccca gctggctgtc actgtaaagc tgtcccacag	1440

cggtcggggcc tggggcgtta tctccccaca acccccagcc aatcaggact ttccagactt	1500
ggccctgaac tactgacgtt cctacctctt atttctcatt gaggctcagg ctatactcca	1560
gctggccaag gctggaaacc tgtctccctc aggctcacct tcctaaggaa aatgtcatag	1620
taggtgctgc tggcccctgg tgatccagct tctctgcaa tcatgacctg ttccttcctg	1680
aagtcctggg catgcatctg ggacccccgt ggagctgaca agttttcctt gctttcctga	1740
tactcttttg cgctgacttg gaattctaag agccttggac ccgagtgtgt ggctagggtt	1800
gccctggctg gggcccgggtg ccgagactcc caagcggctc tgtgcagaag agctgccagg	1860
cagtgtctta gatgtgagac ggaggccatg gcgagaatcc agctttgacc tttattcaag	1920
agaccagatg ggttgcccca ggatccggct gccagccctg aggccaagca cggctggaga	1980
cccacgacct ggctgccgt tgccctgagc tgcagcctcg gccccaggat cctgctcaca	2040
gtcaccgcag gtgcaggcag gaagcagccc tgggggactg gacgctgcta ttgattcatt	2100
aaaaaaagaa aagaaaaat	2119

<210> 87
 <211> 4840
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RB1|ENSG00000139687|ENST00000267163|4840

<400> 87	
tccggttttt ctcaggggac gttgaaatta tttttgtaac gggagtcggg agaggacggg	60
gcgtgccccg acgtgcgcgc gcgtcgtcct ccccggcgct cctccacagc tcgctggctc	120
ccgccgcgga aaggcgtcat gccgccc aaa acccccga aaacggccgc caccgccgcc	180
gctgccgccg cggaaccccc ggcaccgccg ccgccgcccc ctctgagga ggaccagag	240
caggacagcg gcccggagga cctgcctctc gtcaggcttg agtttgaaga aacagaagaa	300
cctgatttta ctgcattatg tcagaaatta aagataccag atcatgtcag agagagagct	360
tggttaactt gggagaaagt ttcatctgtg gatggagtat tgggaggtta tattcaaaag	420
aaaaaggaac tgtggggaat ctgtatcttt attgcagcag ttgacctaga tgagatgtcg	480
ttcactttta ctgagctaca gaaaaacata gaaatcagtg tccataaatt ctttaactta	540

ctaaaagaaa ttgataccag taccaaagtt gataatgcta tgtcaagact gttgaagaag	600
tatgatgtat tgtttgcact cttcagcaaa ttggaaagga catgtgaact tatatatattg	660
acacaaccca gcagttcgat atctactgaa ataaattctg cattgggtgct aaaagtttct	720
tggatcacat ttttattagc taaaggggaa gtattacaaa tggaagatga tctggtgatt	780
tcatttcagt taatgctatg tgtccttgac tattttatta aactctcacc tcccatgttg	840
ctcaaagaac catataaaac agctgttata cccattaatg gttcacctcg aacacccagg	900
cgaggtcaga acaggagtgc acggatagca aaacaactag aaaatgatac aagaattatt	960
gaagttctct gtaaagaaca tgaatgtaat atagatgagg tgaaaaatgt ttattttcaa	1020
aattttatac cttttatgaa ttctcttgga cttgtaacat ctaatggact tccagaggtt	1080
gaaaatcttt ctaaacgata cgaagaaatt tatcttaaaa ataaagatct agatgcaaga	1140
ttattttttg atcatgataa aactcttcag actgattcta tagacagttt tgaaacacag	1200
agaacaccac gaaaaagtaa ccttgatgaa gaggtgaatg taattcctcc acacactcca	1260
gtaggactg ttatgaacac tatccaacaa ttaatgatga ttttaaattc agcaagtgat	1320
caaccttcag aaaatctgat ttcctatttt aacaactgca cagtgaatcc aaaagaaagt	1380
atactgaaaa gagtgaagga tataggatac atctttaaag agaaatttgc taaagctgtg	1440
ggacaggggt gtgtcgaaat tggatcacag cgatacaaac ttggagttcg cttgtattac	1500
cgagtaatgg aatccatgct taaatcagaa gaagaacgat tatccattca aaattttagc	1560
aaacttctga atgacaacat ttttcatatg tctttattgg cgtgcgctct tgaggttgta	1620
atggccacat atagcagaag tacatctcag aatcttgatt ctggaacaga tttgtctttc	1680
ccatggattc tgaatgtgct taatttaaaa gcctttgatt tttacaaagt gatcgaaagt	1740
tttatcaaag cagaaggcaa cttgacaaga gaaatgataa aacattttaga acgatgtgaa	1800
catcgaatca tggaatccct tgcattggctc tcagattcac ctttatttga tcttattaaa	1860
caatcaaagg accgagaagg accaactgat caccttgaat ctgcttgtcc tcttaatctt	1920
cctctccaga ataatcacac tgcagcagat atgtatcttt ctctgtgtaag atctccaaag	1980
aaaaaagggt caactacgcg tgtaaatctt actgcaaagt cagagacaca agcaacctca	2040
gccttccaga cccagaagcc attgaaatct acctctcttt cactgtttta taaaaaagtg	2100
tatcggctag cctatctccg gctaaataca ctttgtgaac gccttctgtc tgagcaccca	2160

gaattagaac	atatcatctg	gacccttttc	cagcacaccc	tgcagaatga	gtatgaactc	2220
atgagagaca	ggcatttgga	ccaaattatg	atgtgttcca	tgtatggcat	atgcaaagtg	2280
aagaatatag	accttaaatt	caaaatcatt	gtaacagcat	acaaggatct	tcctcatgct	2340
gttcaggaga	cattcaaacg	tgttttgatc	aaagaagagg	agtatgattc	tattatagta	2400
ttctataact	cggctttcat	gcagagactg	aaaacaaata	ttttgcagta	tgcttccacc	2460
aggcccccta	ccttgtcacc	aatacctcac	attcctcgaa	gcccttacia	gtttcctagt	2520
tcacccttac	ggattcctgg	agggaacatc	tatatattcac	ccctgaagag	tccatataaa	2580
atttcagaag	gtctgccaac	accaacaaaa	atgactccaa	gatcaagaat	cttagtatca	2640
attggtgaat	cattcgggac	ttctgagaag	ttccagaaaa	taaatcagat	ggtatgtaac	2700
agcgaccgtg	tgctcaaaag	aagtgtgtaa	ggaagcaacc	ctcctaaacc	actgaaaaaa	2760
ctacgctttg	atattgaagg	atcagatgaa	gcagatggaa	gtaaacatct	cccaggagag	2820
tccaaatttc	agcagaaact	ggcagaaatg	acttctactc	gaacacgaat	gcaaaagcag	2880
aaaatgaatg	atagcatgga	tacctcaaac	aaggaagaga	aatgaggatc	tcaggacctt	2940
ggtggacact	gtgtacacct	ctggattcat	tgtctctcac	agatgtgact	gtataacttt	3000
cccaggttct	gtttatggcc	acatttaata	tcttcagctc	tttttgtgga	tataaaatgt	3060
gcagatgcaa	ttgtttgggt	gattcctaag	ccacttgaaa	tgttagtcac	tgttatttat	3120
acaagattga	aaatcttgtg	taaatcctgc	catttaaaaa	gttgtagcag	attgtttcct	3180
cttccaaagt	aaaattgctg	tgctttatgg	atagtaagaa	tggccctaga	gtgggagtcc	3240
tgataacca	ggcctgtctg	actactttgc	cttcttttgt	agcatatagg	tgatgtttgc	3300
tcttgttttt	attaatttat	atgtatatatt	ttttaattta	acatgaacac	ccttagaaaa	3360
tgtgtcctat	ctatcttcca	aatgcaattt	gattgactgc	ccattcacca	aaattatcct	3420
gaactcttct	gcaaaaaatgg	atattattag	aaattagaaa	aaaattacta	attttacaca	3480
ttagatttta	ttttactatt	ggaatctgat	atactgtgtg	cttgttttat	aaaattttgc	3540
ttttaattaa	ataaaaagctg	gaagcaaagt	ataaccatat	gatactatca	tactactgaa	3600
acagatttca	tacctcagaa	tgtaaaaagaa	cttactgatt	attttcttca	tccaacttat	3660
gtttttaaat	gaggattatt	gatagtactc	ttggttttta	taccattcag	atcactgaat	3720
ttataaagta	cccatctagt	acttgaaaaa	gtaaagtgtt	ctgccagatc	ttaggtatag	3780

aggaccctaa cacagtatat cccaagtgca ctttctaatag tttctgggtc ctgaagaatt	3840
aagatacaaa ttaatttttac tccataaaca gactgttaat tataggagcc ttaatttttt	3900
tttcatagag atttgtctaa ttgcatctca aaattattct gccctcctta atttgggaag	3960
gtttgtgttt tctctggaat ggtacatgtc ttccatgtat cttttgaact ggcaattgtc	4020
tatttatctt ttattttttt aagtcagtat ggtctaacac tggcatgttc aaagccacat	4080
tatttctagt ccaaaattac aagtaatcaa gggtcattat gggttaggca ttaatgtttc	4140
tatctgattt tgtgcaaaag cttcaaatta aaacagctgc attagaaaaa gaggcgcttc	4200
tccccctccc tacacctaaa ggtgtattta aactatcttg tgtgattaac ttatttagag	4260
atgctgtaac ttaaaatagg ggatatttaa ggtagcttca gctagctttt aggaaaatca	4320
ctttgtctaa ctcagaatta tttttaaaaa gaaatctggg cttgttagaa aacaaaattt	4380
tattttgtgc tcatttaagt ttcaaactta ctattttgac agttattttg ataacaatga	4440
cactagaaaa cttgactcca tttcatcatt gtttctgcat gaatatcata caaatcagtt	4500
agtttttagg tcaagggctt actatttctg ggtcttttgc tactaagttc acattagaat	4560
tagtgccaga attttaggaa cttcagagat cgtgtattga gatttcttaa ataatgcttc	4620
agatattatt gctttattgc ttttttgtat tggttaaaac tgtacattta aaattgctat	4680
gttactatth tctacaatta atagtttgtc tattttaaaa taaattagtt gtttaagagtc	4740
ttaatggctt gatgttgtgt tctttgtatt aagtacacta atgttctctt ttctgtctag	4800
gagaagatag atagaagata actctcctag tatctcatcc	4840

<210> 88
 <211> 10789
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >REV3L|ENSG00000009413|ENST00000358835|10789

<400> 88	
cgagatctgc gcctcgctc cctcctcacg cgggcggatt ttcttacctt gtgaaaaacc	60
gtctcagcgg ctgccccgc ctcctctgct ggatgtgaaa tggcggtcgg cgagtggttg	120
cgtccgtgac atcatcatca tggcaacaag agctgcagcc tgggaccgag gagcccgtgt	180
gattccccggc ggtggcggca gtggcggcag caccagcacc gacgaaagct cgagggcttc	240

tctcctgcgg ccccttgccg ggtgctcctg aggaggcggc ggcagcagcg cctacaccgc	300
cccgcccgcc gtcctcag gtgcctctgt gtgaggggag ggggccgtgc cgagaagggg	360
agggggcgcc gccgccgtg cggaggggagc cgccgccgt gctgctgccg ctgccgggtc	420
gccagtgaag ggaggcagt gcggcggcgg cgaacatgtt ttcagtaagg atagtgactg	480
cagactacta catggccagc ccgctgcagg ggctggatac ctgccaatcc cccctcacc	540
aggcccctgt caagaagggt cgggtgggtgc gaggcttcgg agcgaccccg gcaggtcaga	600
agacatgtct tcatctacat ggcatctttc cttacctta tgtgccatac gatggttatg	660
gacagcagcc agaaagctat ctttctcaga tggcattcag tatcgacaga gcacttaatg	720
tggctttagg caatccatct tccactgctc agcatgtgtt caaagtgtca ttagtatcag	780
gaatgccttt ttatggttat catgagaagg aaagacactt tatgaagatc tatctttaca	840
atcctacaat ggtgaaaagg atatgtgaac ttttgcaaag cggagccata atgaataaat	900
tttaccagcc tcatgaagcg catattccct acctcctaca gctcttcatt gactacaatc	960
tttatggcat gaatttaata aatctggctg ctgtcaagtt ccgaaaagca agaaggaaaa	1020
gtaatacatt gcatgcaact ggatcctgca agaatcattt atcaggaaat tctcttgctg	1080
atactttatt tcggtgggaa caagatgaaa taccaagctc tttaatattg gaaggtgttg	1140
aaccacagag tacatgtgaa ttagaagtgg atgctgtagc tgctgatatc ttaaatacgtc	1200
tggacattga agctcaaatt ggtggaaacc ctggtctaca ggccatatgg gaagatgaaa	1260
agcaacggcg aagaaacaga aatgaaactt ctcaaagag ccaacctgag tcacaagatc	1320
acaggtttgt gccagcaaca gaaagtgaaa aaaaatttca gaagagactt caggaaattc	1380
tcaaacagaa tgatttctct gtaacattat caggatctgt ggactacagc gatggatccc	1440
aggagtctc tgctgagtta acattgcact ctgaggttct gtctcctgaa atgcttcagt	1500
gtacaccagc caatatggta gaagttcaca aagacaaaga gtcaagcaaa ggtcacacta	1560
gacacaaagt ggaagaagct cttattaatg aagaagcaat tttgaacctt atggaaaata	1620
gtcagacttt tcagcctttg acccaaagac tgagtgagtc acctgttttc atggacagta	1680
gtcctgatga ggctctggta catcttcttg ctggtttgga aagtgatgga tatcgggggg	1740
aaagaaatag gatgcatca ccatgtcgct cttttgaaa taataaatat ccacaaaata	1800
gtgatgatga agaaaatgaa ccacagattg aaaaagagga aatggagctt agtttggatga	1860

tgtcccagag atgggacagc aatattgaag aacattgtgc caaaaagaga tcaactgtgca	1920
gaaataccca cagaagttca actgaagatg atgactcatc ttcaggagaa gaaatggaat	1980
ggagtataa cagtttgctt ctagccagtc tttctatacc tcagttagat ggaactgcag	2040
atgaaaatag tgacaatcca ttgaacaatg aaaattctag aaccactct tctgtaattg	2100
caacaagcaa gctttcagtt aaaccctcca tctttcacia agatgctgct acattagaac	2160
cctcatcttc tgctaagatt acctttcagt gtaaacacac aagtgccctt tcttcccatg	2220
ttttgaacia ggaagattta attgaagacc tttcacagac aaacaaaaat acagaaaaag	2280
gtctagataa ctcactcact tctttttacia acgaaagcac ttattctatg aaataccctg	2340
gatctttaag cagtactgtt cattcagaaa attctcataa agagaatagt aagaaagaga	2400
tcctcccagt atcttcctgt gaaagtagta tttttgatta tgaagaagat attccatctg	2460
ttacaagaca agtaccaagt agaaaatata caaacattag aaaaatcgaa aaggattccc	2520
cttttataca tatgcaccgt caccctaacg agaatacatt gggcaaaaat tctttcaact	2580
tttctgactt aaatcattca aaaaataaag tatcctctga aggaaatgaa aaaggaaaca	2640
gcacagctct gagtagttta ttcccttcat catttactga aaattgtgaa ttactgtcat	2700
gctcagggga gaatagaact atggtgcatt ctcttaatag cactgctgat gaaagtggac	2760
taaataaact taaaattagg tatgaagaat ttcaagaaca taaaacagaa aagccaagcc	2820
tcagccagca agcagcacac tatatgtttt ttcccagtggt tgttctttct aactgtctta	2880
ctagaccaca gaaactatct cctgtcacat ataaattaca acctggcaat aaaccatccc	2940
ggttaaaatt gaataaaagg aaacttgcag gtcacagga gacttctacc aaaagtagtg	3000
agactggatc caaaaagat aattttatac aaaataatcc ttgtaatagt aatcctgaga	3060
aggataatgc attggctagt gatttaacta aaaccactcg tggagctttt gaaaataaaa	3120
caccacaga tggttttata gactgtcact ttggagatgg aacgttagaa actgagcagt	3180
cctttggact atatggaaat aaatacacac ttagagccaa acgcaaggta aattatgaga	3240
ctgaagacag tgagtcaagt tttgtaactc acaactcaaa aattagtcta cctcatccca	3300
tggaaattgg tgaaagtta gatggaactc tcaaatcccg aaaacgaaga aaaatgtcta	3360
aaaagctgcc ccctgtcatc ataaagtata ttattattaa tagatttaga gggagaaaaa	3420
atatgcttgt gaagctagga aaaatagact ctaaagaaaa acaagtaata ttaacagaag	3480

aaaaaatgga actatataaa aagcttgcac ctttgaagga cttttggcca aaagttcccg	3540
actcccctgc aaccaaatat cccatttatc cactaacacc aaagaaaagt cacagaagaa	3600
agtcaaaaca taaatctgct aagaaaaaaa ctggtaaaca acaaaggaca aataatgaaa	3660
atattaaaag aactttgtct ttcaggaaaa aacggtcaca tgctattctt tctcctcct	3720
caccatctta caatgctgaa accgaagatt gtgacctgaa ttatagtgat gttatgtcta	3780
aactaggttt tctttctgag agaagcacia gtcccataaa ttcttctcca cctcgctgct	3840
ggtctccac agatccaaga gctgaagaaa tcatggctgc tgcagaaaaa gaggcaatgc	3900
ttttaaggg tcctaataga tataagaaga ctgttaattc tcgtatagga aaaactagtc	3960
gcgcaagagc acagattaag aatcaaaaag caaagcttgc taatccctct atagttacta	4020
agaaaaggaa caaacgaaat cagacaaata aactagtaga tgatggaaaa aagaaaccaa	4080
gagcaaaaca aaaaacaaat gagaaaggta catcgagaaa gcatacaaca cttaggatg	4140
aaaaaataaa atctcagtct ggtgctgagg ttaagtttgt actgaaacac cagaatgtgt	4200
ctgaatttgc aagtagttct ggaggctctc aactactttt taaacagaaa gatatgccac	4260
taatgggctc tgctgtagat catcccctt ctgcttcctt acccactgga attaatgcac	4320
aacagaagtt atctggctgc ttttcttctt tcttagaaag caagaagtct gtagatttgc	4380
agacattccc cagttcacga gatgatttgc atccatcagt tgtttgtaat tctataggac	4440
ctggagtctc aaaaattaat gttcaaaggc ctcataatca aagtgctatg tttactctaa	4500
aggaatcaac gttaattcaa aaaaatatat ttgaccttct caatcattta tctcaggtag	4560
cacagaatac acagatatct tctggtatgt cctcaaagat agaagataat gcaaataata	4620
tacaaagaaa ctatttgtca tcaatcgga agttaagtga atatcgcaat tccctagaat	4680
caaagctgga ccaagcatat acccctaatt ttttgcatg caaagacagt cagcagcaga	4740
ttgtgtgcat agcggaacag tcaaagcaca gtgaaacttg ttctccggga aatacagctt	4800
cagaggaaag ccaaatgcct aataattgct ttgtaacttc cttgagaagt ccaatcaaac	4860
aaatagcatg ggagcaaaag caaaggggct ttattttaga tatgtcaaat tttaaacctg	4920
aaagagtaaa accgaggtcg ttatcagaag caatttcaca aaccaaagca ctttctcagt	4980
gtaaaaatcg aaatgtgtca acaccttcag catttggtga aggacagtct ggactggcag	5040
ttctaaaaga attgttacia aaaagacagc agaaagcaca aaatgcaaact actacacaag	5100

acccattatc caataaacat caaccaaata aaaatatttc tggttccctt gagcataaca	5160
aagcaaataa acggacacga tcggtaacgt cccaagaaa acctcgaact cccagaagta	5220
caaaacaaaa agaaaaaatc cccaaacttc tcaaagtaga ctctttaaat ttacaaaact	5280
ctagccagtt ggataactct gtatcagatg atagtcccat ctttttttca gatccaggct	5340
ttgaaagttg ttactcactt gaagatagtt tatctcctga acataattat aattttgata	5400
ttaacacaat aggtcagact ggattttgta gcttttattc tggaagtcag tttgtcccag	5460
ctgatcagaa tttgcctcag aagttcctaa gtgatgctgt tcaggatctt tttccaggac	5520
aagctataga aaaaaatgag tttttaagtc atgacaacca gaaatgtgat gaagacaagc	5580
atcataccac agactcagcc tcatggatta gatctggtac ttttaagtcct gaaatttttg	5640
agaagtcaac catagatagc aatgagaatc gtcgccacaa ccagtggaaa aatagctttc	5700
atcctctaac aactcggctt aactcaataa tggattcttt ctgtgttcag caggcagaag	5760
actgtctaag tgaaaaatct agattgaata ggagttcagt aagcaaagaa gtgtttctta	5820
gcctcccaca gccaaacaat tcagactgga ttcaaggtca caccagaaaa gaaatgggac	5880
agtctcttga ctcagccaat acctctttta ctgcaatact ctctcccct gatggtgaac	5940
ttgtagacgt ggcctgtgaa gatttagaac tgtatgtttc aagaaacaat gatatgttga	6000
caccaactcc tgatagttca ccaagatcta ctagctctcc ttcacaatct aaaaatggca	6060
gcttcacccc tcgaactgct aacattctga aaccacttat gtcccccca agtagggaag	6120
aaattatggc aactttgttg gatcatgacc tgtctgagac tatttaccag gaaccatttt	6180
gcagtaatcc ttctgatgta ccagaaaagc ccaggagat tgggtggacgg ctctcatgg	6240
tagaaactcg acttgcaaat gatctggctg agtttgaggg agacttttcc ttggaaggac	6300
ttcgtctttg gaaaacagca ttctcagcaa tgactcagaa tccaaggcca gggtcacccc	6360
ttcgcagtgg ccaaggagtt gtcaataaag ggtcaagtaa tagccctaag atggttgaag	6420
ataaaaaaat tgtgattatg ctttgcaaat gtgccccaa gtcgacaactg gttcaagtgt	6480
ggcttcaagc caaagaagaa tacgaacgtt ccaagaaact gcctaaaacc aagccaactg	6540
gagttgtaaa atctgctgag aactttagct cttcagttaa cccagatgac aaacctgtag	6600
tgcttccaaa aatggatgta agtccatgta tactccccac tacagcacat accaaggagg	6660
atgttgataa ttctcagatt gctttacaag caccaaccac gggatgtagt caaactgcaa	6720

gtgaaagtca gatgctgcca ccagttgcct ctgcaagtga tcccgaataa gatgaagatg	6780
atgatgataa ctattacatt agttatagct cccctgattc tccagtaatt ccccttggc	6840
aacaaccaat atccccagat tccaaagcat taaatggaga tgatagacc tcatcaccag	6900
tagaggagct gccttcattg gcttttgaga acttcttaaa gccataaaa gatggtatac	6960
aaaaagccc ctgcagtga cctcaagagc ctctagtga atctccaatt aatactaggg	7020
caagaactgg gaaatgtgaa tcactttgct ttcatagtac accaatcata cagagaaaac	7080
ttctggaaag gcttcctgaa gcacctggcc ttagccatt atcaacagaa ccaaaaacac	7140
agaagttgag taataagaaa ggaagtaata ctgacactct tagaagagta ctgttaacac	7200
aagcaaagaa tcaatttgca gcagtaaata cccacagaa agaaacttct cagattgatg	7260
gaccatcttt aaacaatact tacggtttca aagtcagcat aaaaaactta caggaggcaa	7320
aagctttaca tgagatacaa aatcttacc taatcagtgt ggagttgcat gctcgaacta	7380
gacgagactt agaaccggat cctgaatttg acccaatctg tgctctgttc tactgcatct	7440
catctgacac tccactgcca gatacagaaa aaacagaact cacaggtgta atagtgattg	7500
ataaagacaa gacagtttct agtcaagata tcagatatca gactccatta cttattagat	7560
ctggaattac aggactcgaa gtcacctatg ctgctgatga gaaggcactt tttcatgaaa	7620
ttgcaaatat aataaagagg tatgatcctg atattctgct aggatatgag attcagatgc	7680
attcctgggg ttacctctta caaagggtg ccgctttaag tattgactta tgtcggatga	7740
tctctcgggt gccagatgac aaaattgaga acagatttgc agctgaaaga gatgagtatg	7800
gatcatatac aatgagtga ataaatattg ttggccgaat tacactaaat ctttgagaa	7860
tcatgagaaa tgaggtggct ctaactaact acaccttga aaatgtgagc tttcatgttc	7920
ttcatcagcg ttttcccctc ttacctttc gagtctgtc agactgggtt gataacaaga	7980
cagatctata cagatggaaa atggttgatc attatgttag ccgtgtccgt ggaaatctcc	8040
aaatgttaga acagctggac ctgattggga aaaccagtga gatggctaga ctttttgga	8100
ttcagttttt acatgtactg acaaggggtt cacagtaccg tgtggaatca atgatgttgc	8160
gtattgctaa accaatgaac tatattcctg tgacacctag tgttcagcaa agatcccaga	8220
tgagagcccc acagtgtgtt cctctaatta tggagcctga atcccgcttc tatagcaact	8280
ctgttctcgt tttggatttc caatcacttt atccttctat tgtgattgca tataactact	8340

gcttttccac ctgccttggc catgtggaga acttgggaaa gtatgatgag ttcaaatttg	8400
gctgtacctc tctgagagta cctccagatt tactttacca agttaggcat gatatcacag	8460
tgtcccccac tggagtagct tttgtcaagc cttcagtaag aaaaggtgta ctaccaagaa	8520
tgcttgaaga aattttgaag actagattta tgggtgaagca gtcaatgaag gcttacaagc	8580
aagacagagc cctgtcacga atgcttgatg cgcgtcagtt gggacttaag ctgatagcaa	8640
atgtcacatt tggctataca tctgctaatt tttctgggag aatgccatgc attgaggttg	8700
gcgatagtat tgttcacaaa gccagagaga ccttggaacg agctattaaa ctggtgaatg	8760
ataccaagaa atggggggct aggggttgat atggcgatac tgacagtatg tttgtgctac	8820
tgaaaggagc cactaaggag cagtctttta agattgggtca ggaaattgcc gaagctgtaa	8880
ctgctaccaa tcctaaacca gtgaaattga agtttgaaaa ggtatatattg ccctgtgttt	8940
tacaaacaaa aaagaggtat gtgggttaca tgtatgaaac actggatcag aaggaccag	9000
tatttgatgc aaaaggaata gaaacagtca gaagagattc ctgccctgct gtttctaaga	9060
tacttgagcg ttctctaaag ctgctatttg aaacgagaga tataagtcta attaaacagt	9120
atgttcagcg acaatgtatg aagcttctgg aaggaaaggc cagcatacaa gactttatct	9180
ttgccaagga atacagagga agtttttctt ataaaccagg agcttggtgtg ccagcccttg	9240
aacttacaag gaaaatgctg acttatgacc ggcgctctga gcctcagggtt ggggagcgag	9300
tgccatacgt catcatttat gggacccccg gagtaccact tatccagctt gtaaggcgcc	9360
cagtggaagt cctgcaggac ccaactctga gactgaatgc tacttactat attaccaagc	9420
aaatccttcc acccttggca agaatcttct cacttatttg tattgatgtc ttcagctggt	9480
atcatgaatt accaaggatc cataaagcta ccagctcctc gcgaagtga cctgaagggc	9540
ggaaaggcac tatttcacaa tattttacta ccttacactg tcctgtgtgt gatgacctaa	9600
ctcagcatgg catctgtagt aaatgtcgga gccaacctca gcatgttgca gtcacctca	9660
accaagaaat ccgggagttg gaacgtcaac aggagcaact tgtaaagata tgcaagaact	9720
gtacaggttg ctttgatcga cacatcccat gtgtttctct gaactgccca gtacttttca	9780
aactctcccg agtaaataga gaattgtcca aggcaccata tctccggcag ttattagacc	9840
agttttaaat tgtcaatatc acagtattac aggtgctatt tttttcagtg cttaccacta	9900
aactgttgtg catggtgctt ttttaactttc atcgagtcaa ggatgttcac tgtctgttat	9960

ctgaagacta tgaagacttc tatgctaacc gaattaaaat gtacttgttg atctctgaat	10020
agctcacttc ttacaatgta caaatcctc attctgtcac cttttaaaca ttgttttata	10080
atgcagggtg tggatttgct ccagtatgtg taccatcttg taaattcatt tgagtagatc	10140
atgtttactt cccagtggaa ggagcactga aaacctctta aagaaaaagc atttgtgtgt	10200
tttccttgaa ctgtctgtat caagacgtgt tacttcgaga tatccattca ctttataatt	10260
ttgactgcaa aatattttgt aaatacactt ttttactttt caaacgagca aaataatgtg	10320
caatgatttt tatacaaag attttcaagt tgtttggtat atttcctcta ggttttgctt	10380
gactcaaagt agatcgttat tttgatcaaa ctgtgcaaac agtagtacca cgtgtagcat	10440
tttgaaacat tatttttttt taaaaaatgc tgtcttgctt tagctattaa tggggcattg	10500
tgagggaactg tgcaaagaca tttttgttac aaacctgtgg gcctgttgca atactttaaa	10560
aataaaaaat tttattccat ttgcttggtt tgtatagaca tttctattgc ttctaaatat	10620
acttaaaata ttttctttcc ttatgtactg tacagttaat cttatttgcc atcatcttga	10680
acacaaaatg tgtattttaga atatttgtat aactgtgtaa aataaaaaag gaattatgtg	10740
gtcagtgcac tgttttttaa actggaaatc attttgtttt aaaagttaa	10789

<210> 89
 <211> 4340
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >RPA1|ENSG00000132383|ENST00000254719|4340

<400> 89	
cgggagcggg gcgcgcaact tctcgggcca ataactgcgc agcgcgcggg acccggtgg	60
ggaagctgga gctgttgcgg ggtccgcggg gaagtcttgg cggaggagcc atggctcgcc	120
aactgagcga gggggccatt gcggccatca tgcagaaggg ggatacaaac ataaagccca	180
tcctccaagt catcaacatc cgtcccatta ctacggggaa tagtccgccg cgttatcgac	240
tgctcatgag tgatggattg aacactctat cctctttcat gttggcgaca cagttgaacc	300
ctctcgtgga ggaagaacaa ttgtccagca actgtgtatg ccagattcac agatttattg	360
tgaacactct gaaagacgga aggagagtag ttatcttgat ggaattagaa gttttgaagt	420

cagctgaagc agttggagtg aagattggca atccagtgcc ctataatgaa ggactcgggc	480
agccgcaagt agctcctcca gcgccagcag ccagcccagc agcaagcagc aggccccagc	540
cgcagaatgg aagctcggga atgggttcta ctgtttctaa ggcttatggt gcttcaaaga	600
catttgaaaa agctgcaggt cccagcctgt cacacacttc tgggggaaca cagtccaaag	660
tggtgccccat tgccagcctc actccttacc agtccaagtg gaccatttgt gctcgtgtta	720
ccaacaaaag tcagatccgt acctggagca actcccagg ggaagggaag cttttctccc	780
tagaactggt tgacgaaagt ggtgaaatcc gagctacagc tttcaatgag caagtggaca	840
agttctttcc tcttattgaa gtgaacaagg tgtattatct ctcgaaaggc accctgaaga	900
ttgctaacaa gcagttcaca gctgttaaaa atgactacga gatgacctc aataacgaga	960
cttccgtcat gccctgtgag gacgaccatc atttacctac gggttcagttt gatttcacgg	1020
ggattgatga cctcgagaac aagtcgaaag actcacttgt agacatcatc gggatctgca	1080
agagctatga agacgccact aaaatcacag tgaggtctaa caacagagaa gttgccaaga	1140
ggaatatcta cttgatggac acatccggga aggtggtgac tgctacactg tggggggaag	1200
atgctgataa atttgatggt tctagacagc ccgtgttggc tatcaaagga gcccgagtct	1260
ctgatttcgg tggacggagc ctctccgtgc tgtcttcaag cactatcatt gcgaatcctg	1320
acatcccaga ggcctataag cttcgtggat ggtttgacgc agaaggacaa gccttagatg	1380
gtgtttccat ctctgatcta aagagcggcg gagtcggagg gagtaacacc aactggaaaa	1440
ccttgatga ggtcaaatcc gagaacctgg gccaaaggca caagccggac tactttagtt	1500
ctgtggccac agtgggtgat cttcgcaaag agaactgcat gtaccaagcc tgcccgactc	1560
aggactgcaa taagaaagtg attgatcaac agaatggatt gtaccgctgt gagaagtgcg	1620
acaccgaatt tcccaatttc aagtaccgca tgatcctgtc agtaaatatt gcagattttc	1680
aagagaatca gtgggtgact tgtttccagg agtctgtga agctatcctt ggacaaaatg	1740
ctgcttatct tggggaatta aaagacaaga atgaacaggc atttgaagaa gttttccaga	1800
atgccaactt ccgatctttc atattcagag tcagggtcaa agtggagacc tacaacgacg	1860
agtctcgaat taaggccact gtgatggacg tgaagccgt ggactacaga gagtatggcc	1920
gaaggctggt catgagcatc aggagaagtg cattgatgtg agaggagcag tgccaatcgg	1980
gcagaagttt gcaaataggc agaatggaat cgatttcctc ccacctcgt gtgacgatcc	2040

catgttagct acacagtga gaggctcttg atggtggact aagcaatttc cccctcgtg	2100
cgcatctcag aacctatcgg taggcaaagg aaaatacgct caggtgggtg tgggttagac	2160
tgtgtcaggc ctacggagtc agccagtggc tagcgcaaga ccagtcactc cctctgcctt	2220
caggcttctg tcaatttcat tatcatcaag caggaattat gtcgtaagtc actgacccta	2280
actgcagacc atgaagtaaa ttatgtaact aggtttttgc ttctccagtg gtgaccaccc	2340
ccccccatcc ccgctcacia cttgggttct tctcagcggg gcgagctgag aagcggtcac	2400
gagcacctgg ggattttagt aagtgtgtct tcctagaatt cgaaggctct ctctttctag	2460
aggtgctaca tagttggtaa tgcttggaat ggcaataggg tagaatgatt aatcaaaggc	2520
atatcttcta tatctgaaga gtatccttcc ttcagggttt aatagactga gtcagatggg	2580
tctgatatta atcaaaattg tctcttctga ggaccgctga taagcattga cttgccgtcc	2640
cctaaggaaa tccgagcggc tacaaagcgt ttctttactt ctcaattcaa ttaatgctgc	2700
gcttcgcttg gtgagtgcgt actttttcta cctgtacaca ttcctgcatt catgtatttt	2760
gttttttttg actaaagcta tgttacatgg aaaggatttt gaagcctttt gtttcccttg	2820
ctttgtttta ataaacagta tattctttgg ttgtgaatcc tactttcttt gaatgcaaag	2880
agttcctata actggaaagc aattaattag ctttcataat aaatgttcac ttggggggga	2940
tgtaaactca ttataaacc gaagattagt ccaaggcatg gagatccttc ttcctagtgt	3000
ttgcagccct gaaatgcatc tttcaaagca tgaaaacact aaaaacaaaa agccattttg	3060
cctgaggatg ctgatgatct ggcaattggg attttattga tgtttacgca gcagtctaac	3120
accaaccacg ctttgaaatg tgtacagaca gtgagctggg aagaaaacag taattatgct	3180
agtgggcctt tcagtcagca aaagcatgct cgctctgtgt gttcctaatac atattaatta	3240
tctatccggg ggctgcaaca caccgcctgc cattggccgc acatctcgcc gtcgtacccc	3300
ggcagtgagg cggtcactct gcagccagag gacctgctgt tcatcactgc acatgccgcc	3360
tgaggaggct tttggatatg gggagtgatg gtgtcctgtc tgtctcccc ctcggtgtct	3420
gccgttgaca tagggggccag ccagccctag acgggatgac ttccgttcct gaggacagac	3480
acagagggac tcctgctcag cctcactaat tgtttagaca cattccttcc tacccttctc	3540
tagtctcagg agatggtaac tgggtcgcat ttcagtctct gactgaggcc tcagccattt	3600
ttacggaagt tcttttctgt ctgaccttgc ttataaagca tcgacgagaa aattacagtc	3660

ttcaaccctc ttctggattg acaaattgtg gctgggaagt ggcgatctag ctttcagccc	3720
agtaaccagt ctttcatgcc tactactccc agcattccct cctctcccca cgtgtctgtc	3780
cacacagtga agaggcctga ccagccgtgg taccaggaca ggacgtgtcc aggggaacgct	3840
gacacctgtc ctcgcgctt ctcaagtggc agcgtgatga aaccagcacg tctccgtgga	3900
tgtgattggg aaccagggg cagtgccagg gggagggtc cctcaggag gctgtttcta	3960
acagatttcc ccactcaaag atcagatcac cagcagagga gcatcagaaa ctggctccac	4020
tgttgggctt ttcagagatt ttggtccctg cgggttgctt aaatagattc tggccacag	4080
tttacctcga aaggctgttg atgttgttct gtttctctc tttcacttag agatcaatgt	4140
tgattttgcg tacaccatga catcagcgtt aggcaattag gagaaaaaa tctaatacatt	4200
tcgcctttat ttcaagtggc tcttgaattc cctccagtct cattgtgaaa ggggcaggga	4260
aaaaaataaa agggtaataa tctgatttct gtccatattt ccagtgtttt atgctttcca	4320
ttaaaacctt gctaattctt	4340

<210> 90

<211> 1752

<212> DNA

<213> Homo sapiens

<220>

<223> >RPA2|ENSG00000117748|ENST00000373912|1752

<400> 90

agtgcggagg gttttgccct tcgtaaagat ggccgcggag gcttttggag ccaactggga	60
gcgcagtacg cgttttctgg agcatgggca gaggagacag gaacaagcgt agcatccgtg	120
agcaccgatt ggctgaagcg agcaccccg gagctgactg gctccgcat tcgcgggaag	180
gcgtttgtgg tgccagagaa aagtagccag agcggcgagc tggcggccgc gttctgtggt	240
tttccgctat tccccagac ccgcacctt tcggcctctt tgcggagaat cgtgaccaag	300
atgtggaaca gtggattcga aagctatggc agctcctcat acgggggagc cggcggctac	360
acgcagtccc cgggggggctt tggatcgccc gcaccttctc aagccgaaaa gaaatcaaga	420
gcccagagccc agcacattgt gccctgtact atatctcagc tgctttctgc cactttggtt	480
gatgaagtgt tcagaattgg gaatgttgag atttcacagg tcactattgt ggggatcatc	540
agacatgcag agaaggctcc aaccaacatt gtttacaaaa tagatgacat gacagctgca	600

cccatggacg ttcgccagtg ggttgacaca gatgacacca gcagtgaaaa cactgtggtt	660
cctccagaaa catatgtgaa agtggcaggc cacctgagat cttttcagaa caaaaagagc	720
ctggtagcct ttaagatcat gcccctggag gatatgaatg agttcaccac acatattctg	780
gaagtgatca atgcacacat ggtactaagc aaagccaaca gccagccctc agcagggaga	840
gcacctatca gcaatccagg aatgagtga gagggaact ttggtgggaa tagcttcatg	900
ccagcaaatg gcctcactgt ggcccaaac caggtgttga atttgattaa ggcttgtcca	960
agacctgaag ggttgaactt tcaggatctc aagaaccagc tgaaacacat gtctgtatcc	1020
tcaatcaagc aagctgtgga ttttctgagc aatgaggggc acatctattc tactgtggat	1080
gatgaccatt ttaaattcac agatgcagaa taactggatc taactgggta cctgagatat	1140
tttacagctg gacctagttt cacaatctgt tgtctccagc tctgcatatg tctggccagg	1200
gggcttctag gaagtaggtt tcacttatca aatgtctcct ctgacttcct tttgaaactt	1260
actgctcttc tgttttatct tgttttgttt gaagctcaga gggagatggg caattgacag	1320
ggatgcaatc cagggtggga tttcttgagg aagttacaaa taagcttggtt acaacatcaa	1380
gatagatgga attggaagga tgctaccagg agagtactta catagtgtc aggagtttct	1440
cttcttaaaa tgtttactgc tgaaagatga gcaggaccag ggcgttatag gcagagccct	1500
agccgagaaa cctgctggcc tctgcctgtt ttcatttccc actttggttg tgtggcatta	1560
ctttcagaat tgcactttcc tgcttgtcat gactttttga cacacttgcc atgacgtgtg	1620
tttctgtgaa catgaagttc tgcggtagt cctccagggg cagaggaaaa gaagaagtgt	1680
tactgcattt tgtacaaaat aaatacagtc atatgtttta taaaacagtt ctattgtagt	1740
aacttgtaaa aa	1752

<210> 91
 <211> 7307
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >SLX4|ENSG00000188827|ENST00000294008|7307

gcgagctccc gttaagaagg tgccgcggcg gcgccggaga tgtgtaatta agtgaaccat	60
---	----

atatgtttca tcatcatgga gatcttggag aattatctga gcaccagggtt catatgtatt	120
cgatctcaga ggcattctatt ggacaacaaa acactctttc agttgtgaac tttattttatt	180
tattattatt attttttgag acagagtttt gctcttggtg cccagggttag agtgcagtgg	240
cacgatctcg gctcactgca atctccgcct cccagggttca agcgattctc ttgcctctgc	300
ctcccgagta gctgggatta caggcatctg ctaccacgcc tggccaattt tttgtatttt	360
cagttgaaac gaggtttcac catattggcc aggctggtct cgaacttctg acctcagggtg	420
atccaccccc cgcctcgtcc tccaaaagag ctgggattac aagtgtgagc caccgcgccc	480
ggcccagttg tggactttaa cagaggggaag ctttaaacad gtttaaccac aggcccaatt	540
tgaacaaaga tacttcaatc attatagaga ggaaaacagt actttttggt caattgtgca	600
aactctccaa gtatctaata gagaagtaga gaagaaccct aatgaaactg agtgtgaatg	660
aggctcagct aggcttctac ttgggttcac tttctcatct gtctgcctgt cctgggattg	720
accctcgctc ctctgaagac cagcctgaaa gccttaaaac tggtcagatg atggatgagt	780
ctgatgagga ctttaaagaa ctctgcgcta gctttttcca aagggtgaaa aaacatggaa	840
tcaaggaagt gtcaggagaa aggaagacac aaaaggctgc ctcaaacggc actcagataa	900
gaagcaaatt gaaaaggacc aaacaaactg ctaccaagac caaaaccctt caaggccctg	960
cagagaagaa acctccgtct ggcagccagg cccctaggac taaaaagcaa agggtaacca	1020
aatggcaagc aagtgaaccg gccactctg tgaatgggga ggggggtgtg cttgcctctg	1080
ctccagatcc acctgtgctc cgggaaacag caaaaaacac ccagacgggt aaccagcaag	1140
aaccatcgcc aaacctttcc agagagaaaa ccagagagaa tgtgccaac agcgactccc	1200
agcctcctcc ttcctgtttg acaacagcag tgccaagtcc ctcaaacc cgcacagcac	1260
aattggtcct acagcgaatg cagcagttca agagagcaga ccccgagcgt ttgagacacg	1320
cttcagaaga gtgctccctc gaggctgcgc gggaagaaaa tgtcccaaag gatcctcaag	1380
aggagatgat ggcggggaat gtgtatgggc ttgggcccc tgccccagag agcgacgctg	1440
cgggtggcctt gaccctgcag caggagtttg cacgggtagg agcatcggca catgatgata	1500
gcctggagga aaagggtttg ttcttctgcc agatttgtca aaagaacctc tcagccatga	1560
acgtgacccg aagggaacag catgtgaaca ggtgcttgga tgaagctgaa aagacactaa	1620
gaccttctgt gcctcagatc cctgagtgcc cgatttgtgg gaaaccgttt cttaccttaa	1680

agagcagaac cagtcacttg aagcagtgtg ctgtgaagat ggaggttggc cccagctcc	1740
tgcttcaggc tgtgcggtg cagacagcac agcctgaggg tagcagcagc ccacccatgt	1800
tcagcttcag tgatcacagt agaggtctga aacggagagg acccaccagc aagaaggagc	1860
cacggaagag gcggaaggtg gacgaggcac cgtccgagga cctgctggtg gccatggctc	1920
tgtcccggtc ggagatggag ccgggtgctg ctgtaccagc gctcaggctg gaaagtgcct	1980
tttctgagag gataagacca gaagcagaga ataaaagtcg caagaagaaa cccccgtat	2040
ccccccatt gttgttagtc caggactctg aaaccacagg ccgacagata gaggaccgtg	2100
tggccctgct cctctctgag gaagtggaat tgtctagcac gccaccactt cctgccagca	2160
ggattttaaa ggaagggtgg gaaagagcgg gccagtgtcc tcctccacct gaacgcaagc	2220
agagctttct gtgggagggc agcgcactga ctggggcctg ggccatggag gacttctaca	2280
cggccaggct ggtccctcct ctctgcccc agcggcctgc ccagggcctt atgcaggagc	2340
ccgtgccgcc tctggtgcca cctgagcact cagagctgag cgagcgaagg tcacccgctc	2400
tccacggcac cccactgca ggctgtggct ccaggggccc gtcgccttcg gccagccaga	2460
gggagcacca ggccctgcag gacctcgtgg acctggcgag ggagggactg agcgccagcc	2520
cgtggcccg cagtgggggc ctggctggct cggaaggagc tgcagggttg gacgtggtgc	2580
ccggcggcct tcctctgact gggtttgttg tgccatcgca ggacaagcac ccggacaggg	2640
gcggccgcac cttgctctcc ctccggctgc tggttgctga ctttggcgcc atggtcaata	2700
accacacct gagtgatgtc cagtttcaga cggacagcgg ggaggtgctt tacgccaca	2760
agttcgtgct ttatgcccga tgcccgtcc tcatccagta tgtgaacaat gaaggcttct	2820
ccgctgtaga ggacggggtt ctgaccagc gtgtcctgct gggtgacgtg agcaccgagg	2880
ccgcccgcac gttcctgcac tatctctaca ctgcggacac tggccttcct cctggcctta	2940
gctctgagct gagctccctg gccacaggt ttggcgtgag tgagctcggt cacctgtgcg	3000
aacaggtgcc tattgccact gactcagagg gcaaaccatg ggaggagaag gaagcagaga	3060
attgcgaaag cagggccgag aatttcagg aactcttgag gtcaatgtgg gcagatgaag	3120
aggaggaagc ggagactttg ttgaaatcca aggaccacga agaagatcaa gaaaacgtga	3180
atgaagcaga aatggaagaa atttatgaat ttgcagctac tcagcgaaag cttctccagg	3240
aagaaagggc agcgggtgcc ggcgaggacg ctgactggct ggagggtggc agtccggttt	3300

ctgggcaact cctagcaggt gtccaggtgc agaaacagtg ggacaaggtg gaggagatgg	3360
agccgttggga gccaggaaga gatgaggccg ccaccacctg ggagaagatg ggacagtgcg	3420
ctctcccgcc accccagggc cagcactcag gggcacgggg agcagaggcc cctgagcagg	3480
aggcgccaga ggaggcgctt ggccattcca gctgctccag cccttccagg gactgccagg	3540
cagagagaaa agaaggctct cttccgcact cagatgatgc cggggattac gaacagctct	3600
tctcatcaac tcagggagag atctcagagc cgtcccaaat aacaagtgag cccgaggaac	3660
aaagtggcgc tgtcagggaa agggggctgg aggttttctca tcgcctggct ccctggcagg	3720
catctccacc gcacccgtgc cgcttcctat tggggcctcc ccagggcggg agtccccgcg	3780
ggtctcatca cacaagtggg tcgtccctgt caacaccccg gtcccgtggc ggaacttccc	3840
aggtgggctc cccaaccttg ctgtctccag ctgtgccatc aaagcagaaa agggacagga	3900
gcatcctcac gctgtctaaa gagccagggc accagaaagg caaagagcgt cggtccgtgc	3960
tggagtgcag aaataagggg gtcctgatgt tcccagaaaa atctccgtct attgacctaa	4020
cccagtcaaa tcctgacat tcgagctcca gatctcagaa atcttcatcc aaactgaacg	4080
aagaagatga ggtcatcctc ttactggact cggatgagga gctggagcta gaacaaacca	4140
aaatgaagtc catttctagt gatcctctgg aagaaaagaa agctctagaa attagcccta	4200
ggtcctgtga gctgttttcc atcattgatg ttgatgcaga tcaggaacct tcccagagcc	4260
caccaagaag cgaagctgtg ctgcagcagg aggatgaggg ggcgctgccg gagaatcggg	4320
gctctttggg caggagaggg gctccctggc tgttctgtga ccgtgagagc agccccagcg	4380
aggccagcac cacagacacc tcgtggctgg tgcccgccac cccgctggcc agcagaagcc	4440
gtgactgttc ttcccagacc caaatcagca gcctcaggag cgggctggcc gtgcaggcgg	4500
tgactcagca cacgcccagg gcctcagtag gaaacaggga agggaacgaa gtcgcacaga	4560
agttttctgt catcaggccc cagacaccac cgccccagac accgtcctca tgcctcactc	4620
ccgtctctcc aggaacttct gacggcagaa ggcaaggcca cagaagccct tcccgtcccc	4680
accccggggg ccacccgcac tcctctccgc tggctccaca tcccattctca ggggaccgcg	4740
cccacttcag caggcggttc ctgaaacact cgccgcctgg gccaagcttc ctgaaccaga	4800
ccccagcggg tgaagtgggt gaagtcggag acagtgacga tgagcaggag gtggcctccc	4860
atcaggccaa cagaagcccc ccactggaca gtgaccccc aattccaatt gacgactgct	4920

gctggcacat ggagcccctc tcgccaattc ccattgacca ctggaacctg gagcggaccg	4980
gccccctgag caccagcagc cccagccgca ggatgaacga ggccgccgac agccgtgact	5040
gtcgtctccc gggactcctg gacaccaccc ccatccgagg aagctgcact acccagagga	5100
aattgcaaga gaagtcctcg ggcgcgggct ccctggggaa tagcaggccg agcttttctga	5160
attcggctct gtgggacgtt tgggacgggg aagagcagag gcctccagag acccctcctc	5220
cggcccagat gccaagcgct ggtggagctc agaagcccga agggttagag acacccaaag	5280
gtgctaactg gaagaagaac ttgccccca aagtgcccat aacgccgatg ccacagtatt	5340
ccattatgga gacgccggtg ctgaagaagg aactggatag gtttgagtc cgccctctgc	5400
ctaaacgcca gatggttctg aagctgaagg agatattcca gtacactcac cagaccctgg	5460
actcagactc cgaggacgag agccagtcct cacagccgct gttgcaggcg cctcactgcc	5520
agaccctcgc ctcccagacc tacaagcctt caagggcagg ggtccatgcc cagcaggagg	5580
ccaccacagg acctggggcc cataggccca agggacctgc taagaccaag ggcccccgac	5640
atcaaaggaa gcatcatgaa agcatcacac cccaagcag gtcgcccacc aaggaggcac	5700
ctccaggcct caatgatgac gcccagatcc cagcctctca agaatccgtg gccacctctg	5760
tggatggcag tgacagctcc ttgagctcac agagttcttc ctctgtgag tttggagcgg	5820
catttgagtc tgcaggtgaa gaggagggcg agggggaggt cagtgcctcg caggcagccg	5880
tgcaggcggc ggacacagac gaggcgctga ggtgctacat ccgctccaag ccggccctgt	5940
accagaaggt gctgctgtac cagccctttg agctgcggga gctgcaggca gagctgaggc	6000
agaacggcct ccgtgtgtcc tcgcgaggc tgttgactt cctggacacc cactgtatca	6060
ccttcaccac tgccgccacc cgcagggaga agctccaggg caggaggcgg cagcctcggg	6120
gcaagaagaa ggtggagcgg aactgatggg gccatcccga cccacccca acctgccatc	6180
agcagcccc acccccgcca tttgcaggga ggacctggga caccagcgt gggtcaggcc	6240
tccacaggca tttctgggcc tggggaccac atcagctctg cgctgtgatg atgaccacag	6300
cccaatccag ggcttctctc tctgggctct gctttctagg gtggcatttg gagcatgtca	6360
cccactggat ttacagactc cagccccttc ctctgtccgt gctcacagtg tgtctccctt	6420
tttggttttc tttttttttt tctttttgag acagtcgtgc tgtgtcacct aggctggagt	6480
gcagtggcac aatctcggct cactgcaacc tccgcctccc gggttcaagc aattctcctg	6540

cctcagcctc ccagatagct gggactacag gcacacgctg ccacgcccag ctgatttttt	6600
atatttttagt agaaacgggg tttcaccatg ttgctcaggc tggtcacaaa ctccagagct	6660
caggcaatcc gcctgctgcg gcctcccaaa gtgctgggat cacaggcgcc agccactgcg	6720
cccgggcccag tgtgtctccc ttaaccaag agggccctca gctgtcccag ggggcagtgg	6780
gccatcacca gctggccagg gcatggccta ttctgccaca tttgccaccc tctgagccca	6840
ccagtcctgg gcacagctgc cctacatgtc tgtcctgaga tggacgtcag gtccagcctg	6900
ccccggcagc ccggggcccgt cctcctcagc actcaggcca accccagcca ccgccagcct	6960
gagaccaggt gtcctgaggc tccctgcact gccacagccc agatgcagtt ctctgaccc	7020
agccgtgcta cccggacact tgtcattgtt accagcagtc tccaaactgg acagtgcaca	7080
agggcccaga acaactctga tgccaccaca aaacaaacat gttcactagc ggattccatt	7140
ctttgggtta aagctgcctc cagcctcagg agcagtgtgg aggaagatga gggccaggaa	7200
agaaggaaac cttggtttct ccatccttgt gaatgtcctc gtctgtttca aatacagtgc	7260
agtcagtttt atatgatgtg caataaacca aaaaggcttt attaaaa	7307

<210> 92
 <211> 3696
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >TDP1|ENSG0000042088|ENST00000335725|3696

<400> 92	
gttggttctg tgcgcctcag agttggagca cacagctgta ttaaaaaggc aaatcgaagg	60
ccgggcgcggtg tgactcacgc ctgtcatcct agcactttgg gaggccgagg cggctgaatc	120
acttgagggtt aggagtttga gatcagcccc ggcaacatgg tgaaaccccc tctctacaaa	180
aatagaaaaa ttagccgagc gtgatggtgg atgcctgtaa tcctagctcc tcgggaggct	240
aaggagtata atgtctcagg aaggcgatta tgggaggtgg accatatcta gtagtgatga	300
aagtgaggaa gaaaagccaa aaccagacaa gccatctacc tcttctcttc tctgtgccag	360
gcaaggagca gcaaatgagc ccagggtacac ctgttccgag gccagaaaag ctgcacacaa	420
gaggaaaata tcacctgtga aattcagcaa tacagattca gttttacctc ccaaaaggca	480
gaaaagcgggt tcccaggagg acctcggctg gtgtctgtcc agcagtgatg atgagctgca	540

accagaaatg ccgcagaagc aggctgagaa agtggatgac aaaaaggaga aagacatctc	600
tgctcccaat gacggcactg cccaaagaac tgaaaatcat ggcgctcccg cctgccacag	660
gctcaaagag gaggaagacg agtatgagac atcaggggag ggccaggaca tttgggacat	720
gctggataaa gggaaccctt tccagtttta cctcactaga gtctctggag ttaagccaaa	780
gtataactct ggagccctcc acatcaagga tattttatct cctttatttg ggacgcttgt	840
ttcttcagct cagtttaact actgctttga cgtggactgg ctctgtaaac agtatccacc	900
agagttcagg aagaagccaa tcctgcttgt gcatggatgag aagcgagagg ctaaggctca	960
cctccatgcc caggccaagc cttacgagaa catctctctc tgccaggcaa agttggatat	1020
tgcgttttga acacaccaca cgaaaatgat gctgctgctc tatgaagaag gcctccgggt	1080
tgtcatacac acctccaacc tcatccatgc tgactggcac cagaaaactc aaggaatatg	1140
gttgagcccc ttataccac gaattgctga tggaaccac aaatctggag agtcgccaac	1200
acattttaaa gctgatctca tcagttactt gatggcttat aatgcccctt ctctcaagga	1260
gtggatagat gtcattcaca agcacgatct ctctgaaaca aatgtttatc ttattggttc	1320
aaccccagga cgctttcaag gaagtcaaaa agataattgg ggacatttta gacttaagaa	1380
gcttctgaaa gaccatgcct catccatgcc taacgcagag tcctggcctg tcgtaggtca	1440
gttttcaagc gttggctcct tgggagccga tgaatcaaag tggttatgtt ctgagtttaa	1500
agagagcatg ctgacactgg ggaaggaaag caagactcca ggaaaaagct ctgttcctct	1560
ttacttgatc tatccttctg tggaaaatgt gcggaccagt ttagaaggat atcctgctgg	1620
gggctctctt ccctatagca tccagacagc tgaaaaacag aattggctgc attcctatct	1680
tcacaaatgg tcagctgaga cttctggccg cagcaatgcc atgccacata ttaagacata	1740
tatgaggcct tctccagact tcagtaaaat tgcttggttc cttgtcaciaa gcgcaaactc	1800
gtccaaggct gcctggggag cattggagaa gaatggcacc cagctgatga tccgctccta	1860
cgagctcggg gtccttttcc tcccttcagc atttggctta gacagtttca aagtgaacaa	1920
gaagttcttc gctggcagcc aggagccaat ggccaccttt cctgtgccat atgatttgcc	1980
tccagaactg tatggaagta aagatcggcc atggatatgg aacattcctt atgtcaaagc	2040
accggatacg catgggaaca tgtgggtgcc ctctgagaa tcttgaggca ctgtgaaatt	2100
taagtgtgaa acattgagcc acaaactggt aatctcttct ttgtactgga tgtccacttc	2160

ccttaaagtc ttatttgcac ccttacaaaa tctttccaaa ggtcactctt atgaatggat	2220
gttggttata cttttaatgg acattaacat tcctaataaa gtattagttt cttaattcac	2280
ttttatatgt tttggaaaga aaattagtgat acttctctat gttaaaaata cgtactgctt	2340
gagtatcccc tgtctgaaat gcttgggacc agaagtgttt cagcttttgg atttttttga	2400
attttggaat atttgcatag cataatgaga tatcttggga atgggaccca aatctaaaca	2460
caaaattcat ttatgtttca tatacacctt atatacaata acctaaaggt gattttatat	2520
gatattttga gtaattttat gcatgaaaca aagttttgac aggcttttga ccgtgattca	2580
tcacatgagt tcaggcatgg aaattttcat ttggagcatc atgtcagcac tcaaaaagtt	2640
ctggatcttg gagcagttca gattttcaga ttagggatgc tcaaactctat atagatatata	2700
aattatcctc acagtaacat agaactctctt ggtgctgtca gctgttggga attgaagatt	2760
gactttgtgc ttccaccctc catccagaaa ggcacccttc attccaccag aactttaccc	2820
aggaagaaca cgatcatttc ctttttcacc gatgccctct ctcagctttc tgagtacgtc	2880
tcttggggtc gctggaggtg atcctaggat ctgtctctga gaccaatgtg ctgtttcagc	2940
cccctgcagc taagaattgt attgactgtc ctcacagcgg cttttcatag ctttcagctt	3000
cagctttacg aggcttctcc tctctccctg gcaccctgct ggctgcctca ctgcttacag	3060
acaggtccca ccaaacccaa acacctgcct agggtaaagt ggtctctctt ctatccccag	3120
aaactttcag aggaagcagc tcatagaaac atacaaaagc acacaagtat tttgggaaaa	3180
aatcctaaaa ggtgacttaa tttgatgcct taaattcaca agtgaggaag ctaaggccta	3240
gaaggttaag gatgtcccca gggtcacaca gtgagcgggg ctcagagctt gagtgtcttt	3300
gtgctttgtg tacatttgtt tctccctagg gtgctttaga ccctgtttgt tttcttctgc	3360
atgaggctga tttccagttt gtcattcaacc tctttatctt ataatttagg atagagttga	3420
acgttagtct tgaaagattt tctaaagtag tctttcaaac tgttcctcag aggcctagga	3480
ttttccaaaa gtaccttagg aaccttgtag gctgcagtgg ggggtgtggcg atagagcagg	3540
aggcagggag acagggctgc agggcctccc accttccaac agacaggctc tgctgtatct	3600
gttgtagata ctgggattct gtaaaggaca ttatctgggg tgttgatatgt atttttgtgt	3660
gttctgcttt ttttaaataa acttgaaaag ctactg	3696

<210> 93
 <211> 2071
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >TDP2|ENSG00000111802|ENST00000378198|2071

<400> 93
 gcacccaccg ccccggcggc tcccttgagg cgcagctgca ccagttttcc gagagcagcg 60
 catttccccg ccgggcgggt gatgcgggag cgccatgaca caggcgcctg cgcagagccg 120
 cgcgttgagg tcctgttccg cttaaagggg cggcgcagag gcggcaggaa gatggagttg 180
 gggagttgcc tggagggcgg gagggaggcg gcggaggaag agggcgagcc tgaggtgaaa 240
 aagcggcgac ttctgtgtgt ggagtttgcc tcggctgcaa gctgcgatgc cgcagtggct 300
 cagtgccttc tggccgagaa cgactgggag atggaaaggg ctctgaactc ctacttcgag 360
 cctccggtgg aggagagcgc cttggaacgc cgacctgaaa ccatctctga gccaagacc 420
 tatgttgacc taaccaatga agaaacaact gattccacca cttctaaaat cagcccatct 480
 gaagatactc agcaagaaaa tggcagcatg ttctctctca ttacctggaa tattgatgga 540
 ttagatctaa acaatctgtc agagagggtc cgaggggtgt gttcctactt agctttgtac 600
 agcccagatg tgatatttct acaggaagtt attcccccat attatagcta cctaaagaag 660
 agatcaagta attatgagat tattacaggt catgaagaag gatatttcac agctataatg 720
 ttgaagaaat caagagtga aataaaaagc caagagatta ttccttttcc aagtaccaa 780
 atgatgagaa accttttatg tgtgcatgtg aacgtgtcag gaaatgagct ttgccttatg 840
 acatcccatt tggagagcac cagagggtc gctgcggaac gaatgaatca gttaaaaatg 900
 gttttaaaga aaatgcaaga ggctccagag tcagctacag ttatatttgc aggagataca 960
 aatctaaggg atcgagaggt taccagatgt ggtggtttac ccaacaacat tgtggatgtc 1020
 tgggagtttt tgggcaaacc taaacattgc cagtatacat gggatacaca aatgaactct 1080
 aatcttgga taactgctgc ttgtaaactt cgttttgatc gaatattttt cagagcagca 1140
 gcagaagagg gacacattat tccccgaagt ttggaccttc ttggattaga aaaactggac 1200
 tgtggtagat ttcctagtga tcaactgggt cttctgtgca acttagatat aatattgtaa 1260
 aatgcttttc aagtgtgggt tttgccctga ttgttgcaaa tacaatttcc accttctgga 1320

aaggtagggtg	tgctgtggag	gaaataatgt	actagatcat	tgtcacagaa	aaaccaacta	1380
tgatttatgg	ttgtgttttc	agaattcaac	attaaagatt	aatgtttatt	taaacgaaca	1440
cattcctgca	ttcaggatgt	gaggccattt	aataaaaagg	gcacaaagcc	tgtcagagtt	1500
ttcaacgggtg	cttatagctg	ccagctggat	tccaaacagg	taccacattg	tctctgagct	1560
aatgtttata	tttttccatt	caggcaccga	aatagttaat	atttgaaata	agtcttcaaa	1620
agaaaacata	agagattatt	gagttcttgg	gactggatcc	tttatttcat	aagttcagat	1680
catcttaaata	gaaaatgcc	tgattatctg	cagttaagta	gatgacagct	attctacatc	1740
agacttgatt	tttgtcagct	aattacataa	ttggtaagct	ataattgaaa	ccttatggct	1800
taaaattcct	taactccttt	ttgattcatg	tttgtagtca	tgttgtcaac	agaggcaaag	1860
ttaagcttga	tgatgggtta	aatcgggttg	atagcaccat	gggacatttt	tctaacaaaa	1920
ataaatgcat	gaagagacat	agccttttag	ttttgctaata	tgtgaaatgg	aaatgcttta	1980
caggaagtaa	atgcaaatta	cttttaagt	tgctttaaag	aaaaatattt	tccccacaag	2040
agaaatttaa	ataaagaatt	ttatttgttt	a			2071

<210> 94
 <211> 3240
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >TMPRSS2|ENSG00000184012|ENST00000398585|3240

<400> 94	
accagggtcc	cggtcgggg tccgggctgg ggagggaac ctgggcgcct gggacccgcc 60
gatgccccct	gccccgccg gaggtgaaag cgggtgtgag gagcgcggcg cggcaggtca 120
tattgaacat	tccagatacc tatcattact cgatgctgtt gataacagca agatggcttt 180
gaactcaggg	tcaccaccag ctattggacc ttactatgaa aacctggat accaaccgga 240
aaacccttat	cccgcacagc ccactgtggg cccactgtc tacgaggtgc atccggctca 300
gtactacccg	tccccgtgc ccagtacgc cccgagggtc ctgacgcagg cttccaacct 360
cgtcgtctgc	acgcagccca aatccccatc cgggacagtg tgcacctcaa agactaagaa 420
agcactgtgc	atcaccttga ccctggggac cttcctcgtg ggagctgcgc tggccgctgg 480
cctactctgg	aagttcatgg gcagcaagtg ctccaactct gggatagagt gcgactcctc 540

aggtacctgc atcaaccct ctaactggtg tgatggcgtg tcacactgcc ccggcgggga	600
ggacgagaat cggtgtgttc gcctctacgg accaaacttc atccttcagg tgtactcatc	660
tcagaggaag tcctggcacc ctgtgtgcca agacgactgg aacgagaact acgggcgggc	720
ggcctgcagg gacatgggct ataagaataa tttttactct agccaaggaa tagtggatga	780
cagcggatcc accagcttta tgaaactgaa cacaagtgcc ggcaatgtcg atatctataa	840
aaaactgtac cacagtgatg cctgttcttc aaaagcagtg gtttctttac gctgtatagc	900
ctgcggggtc aacttgaact caagccgcca gagcaggatt gtgggcggcg agagcgcgct	960
cccgggggcc tggccctggc aggtcagcct gcacgtccag aacgtccacg tgtgcggagg	1020
ctccatcatc acccccgagt ggatcgtgac agccgcccac tgcgtggaaa aacctcttaa	1080
caatccatgg cattggacgg catttgcggg gatthttgaga caatctttca tgttctatgg	1140
agccggatac caagtagaaa aagtgatttc tcatccaaat tatgactcca agaccaagaa	1200
caatgacatt gcgctgatga agctgcagaa gcctctgact ttcaacgacc tagtgaaacc	1260
agtgtgtctg cccaaccag gcatgatgct gcagccagaa cagctctgct ggatttccgg	1320
gtggggggcc accgaggaga aagggaagac ctcagaagtg ctgaacgctg ccaaggtgct	1380
tctcattgag acacagagat gcaacagcag atatgtctat gacaacctga tcacaccagc	1440
catgatctgt gccggcttcc tgcaggggaa cgtcgattct tgccagggtg acagtggagg	1500
gcctctggtc acttcgaaga acaatatctg gtggctgata ggggatacaa gctggggttc	1560
tggctgtgcc aaagcttaca gaccaggagt gtacgggaat gtgatggtat tcacggactg	1620
gatttatcga caaatgaggg cagacggcta atccacatgg tcttcgtcct tgacgtcgtt	1680
ttacaagaaa acaatggggc tggttttgct tccccgtgca tgatttactc ttagagatga	1740
ttcagaggtc acttcatttt tattaacag tgaacttgtc tggctttggc actctctgcc	1800
attctgtgca ggctgcagtg gctcccctgc ccagcctgct ctccctaacc cttgtccgc	1860
aaggggtgat ggccggctgg ttgtgggcac tggcgggtcaa gtgtggagga gaggggtgga	1920
ggctgccccca ttgagatctt cctgctgagt cttttccagg ggccaatttt ggatgagcat	1980
ggagctgtca cctctcagct gctggatgac ttgagatgaa aaaggagaga catggaaagg	2040
gagacagcca ggtggcacct gcagcggctg ccctctgggg ccacttggtg gtgtccccag	2100
cctacctctc cacaagggga ttttgctgat gggttcttag agccttagca gccctggatg	2160

gtggccagaa ataaaggac cagcccttca tgggtggtga cgtggtagtc acttgtaagg	2220
ggaacagaaa catttttgtt cttatgggggt gagaatatag acagtgcctt tgggtgcgagg	2280
gaagcaattg aaaaggaact tgccctgagc actcctggtg caggtctcca cctgcacatt	2340
gggtggggct cctgggaggg agactcagcc ttcctcctca tcctccctga ccctgctcct	2400
agcaccctgg agagtgcaca tgccccttgg tcctggcagg gcgccaagtc tggcaccatg	2460
ttggcctctt caggcctgct agtcactgga aattgaggtc catgggggaa atcaaggatg	2520
ctcagtttaa ggtacactgt ttccatgtta tgtttctaca cattgctacc tcagtgtctc	2580
tggaaactta gcttttgatg tctccaagta gtccaccttc atttaactct ttgaaactgt	2640
atcatctttg ccaagtaaga gtggtggcct atttcagctg ctttgacaaa atgactggct	2700
cctgacttaa cgttctataa atgaatgtgc tgaagcaaag tgcccatggt ggcggcgaag	2760
aagagaaaga tgtgttttgt tttggactct ctgtggtccc ttccaatgct gtgggtttcc	2820
aaccagggga agggtcctt ttgcattgcc aagtgccata accatgagca ctactctacc	2880
atggttctgc ctctggcca agcaggctgg tttgcaagaa tgaaatgaat gattctacag	2940
ctaggactta accttgaaat ggaaagtcat gcaatcccat ttgcaggatc tgtctgtgca	3000
catgcctctg tagagagcag cattcccagg gaccttgga acagttggca ctgtaagggtg	3060
cttgctcccc aagacacatc ctaaaagggt ttgtaatggt gaaaacgtct tccttcttta	3120
ttgccccttc ttatttatgt gaacaactgt ttgtcttttt ttgtatcttt tttaaactgt	3180
aaagttcaat tgtgaaaatg aatatcatgc aaataaatta tgcaattttt ttttcaaagt	3240

<210> 95

<211> 5758

<212> DNA

<213> Homo sapiens

<220>

<223> >TOP2A|ENSG00000131747|ENST00000423485|5758

<400> 95

gattggctgg tctgcttcgg gcgggctaaa ggaaggttca agtggagctc tcctaaccga	60
cgcgctctg tggagaagcg gcttggtcgg ggggtggtctc gtggggtcct gcctgtttag	120
tcgctttcag ggttcttgag ccccttcacg accgtcacca tggaagtgtc accattgcag	180

cctgtaaatg aaaatatgca agtcaacaaa ataaagaaaa atgaagatgc taagaaaaga	240
ctgtctgttg aaagaatcta tcaaaagaaa acacaattgg aacatatttt gctccgcccc	300
gacacctaca ttggttctgt ggaattagtg acccagcaaa tgtgggttta cgatgaagat	360
gttggcatta actatagga agtcactttt gttcctggtt tgtacaaaat ctttgatgag	420
attctagtta atgctgcgga caacaaacaa agggacccaa aaatgtcttg tattagagtc	480
acaattgatc cggaanaaaa ttttaattagt atatggaata atggaaaagg tattcctggt	540
gttgaacaca aagttgaaaa gatgtatgtc ccagctctca tatttgga gtcctaact	600
tctagtaact atgatgatga tgaaaagaaa gtgacaggtg gtcgaaatgg ctatggagcc	660
aaattgtgta acatattcag taccaaattt actgtggaaa cagccagtag agaatacaag	720
aaaatgttca aacagacatg gatggataat atgggaagag ctggtgagat ggaactcaag	780
cccttcaatg gagaagatta tacatgtatc acctttcagc ctgatttgtc taagtttaaa	840
atgcaaagcc tggacaaaga tattgttgca ctaatggtca gaagagcata tgatattgct	900
ggatccacca aagatgtcaa agtctttctt aatggaaata aactgccagt aaaaggattt	960
cgtagttagt tggacatgta tttgaaggac aagttggatg aaactggtaa ctccttgaaa	1020
gtaatacatg aacaagtaaa ccacaggtgg gaagtgtgtt taactatgag tgaaaaaggc	1080
tttcagcaaa ttagctttgt caacagcatt gctacatcca aggggtggcag acatgttgat	1140
tatgtagctg atcagattgt gactaaactt gttgatgttg tgaagaagaa gaacaagggt	1200
ggtgttgacg taaaagcaca tcagggtgaaa aatcacatgt ggatttttgt aaatgcctta	1260
attgaaaacc caacctttga ctctcagaca aaagaaaaca tgactttaca acccaagagc	1320
tttggatcaa catgccaatt gagtgaanaa tttatcaaag ctgccattgg ctgtggtatt	1380
gtagaaagca tactaaactg ggtgaagttt aaggcccaag tccagttaaa caagaagtgt	1440
tcagctgtaa aacataatag aatcaaggga attcccaaac tcgatgatgc caatgatgca	1500
gggggcccga actccactga gtgtacgctt atcctgactg agggagattc agccaaaact	1560
ttggctgttt caggccttgg tgtggttggg agagacaaat atggggtttt ccctcttaga	1620
ggaaaaatac tcaatgttcg agaagcttct cataagcaga tcatggaaaa tgctgagatt	1680
aacaatatca tcaagattgt gggctttcag tacaagaaaa actatgaaga tgaagattca	1740
ttgaagacgc ttcgttatgg gaagataatg attatgacag atcaggacca agatggttcc	1800

cacatcaaag gcttgctgat taattttatc catcacaact ggccctctct tctgcgacat	1860
cgttttctgg aggaatttat cactcccatg gtaaaggatg ctaaaaacaa gcaagaaatg	1920
gcattttaca gccttcctga atttgaagag tggaagagtt ctactccaaa tcataaaaaa	1980
tggaaagtca aatattacaa aggtttgggc accagcacat caaaggaagc taaagaatac	2040
tttgagata tgaaaagaca tcgtatccag ttcaaataat ctggctcctga agatgatgct	2100
gctatcagcc tggccttttag caaaaaacag atagatgatc gaaaggaatg gttaactaat	2160
ttcatggagg atagaagaca acgaaagtta cttgggcttc ctgaggatta cttgtatgga	2220
caaaactacca catatctgac atataatgac ttcatcaaca aggaacttat cttgtttctca	2280
aattctgata acgagagatc tatcccttct atgggtggatg gtttgaaacc aggtcagaga	2340
aagggtttgt ttacttgctt caaacggaat gacaagcgag aagtaaaggt tgcccaatta	2400
gctggatcag tggctgaaat gtcttcttat catcatgggtg agatgtcact aatgatgacc	2460
attatcaatt tggctcagaa ttttgtgggt agcaataatc taaacctctt gcagcccatt	2520
ggtcagtttg gtaccaggct acatgggtggc aaggattctg ctagtccacg atacatcttt	2580
acaatgctca gctctttggc tcgattgtta tttccaccaa aagatgatca cacgttgaag	2640
tttttatatg atgacaacca gcgtgttgag cctgaatgggt acattcctat tattcccatg	2700
gtgctgataa atgggtgctga aggaatcgggt actgggtgggt cctgcaaaat cccaacttt	2760
gatgtgcgtg aaattgtaaa taacatcagg cgtttgatgg atggagaaga acctttgcca	2820
atgcttccaa gttacaagaa cttcaagggt actattgaag aactggctcc aaatcaatat	2880
gtgattagtg gtgaagtagc tattcttaat tctacaacca ttgaaatctc agagcttccc	2940
gtcagaacat ggaccagac atacaaagaa caagttctag aacctatgtt gaatggcacc	3000
gagaagacac ctctctcat aacagactat agggaaatacc atacagatac cactgtgaaa	3060
ttgtttgtga agatgactga agaaaaactg gcagaggcag agagagttgg actacacaaa	3120
gtcttcaaac tccaaactag tctcacatgc aactctatgg tgctttttga ccacgtaggc	3180
tgtttaaaga aatatgacac ggtgttggtg attctaagag acttttttga actcagactt	3240
aaatattatg gattaagaaa agaattggctc ctaggaatgc ttgggtgctga atctgctaaa	3300
ctgaataatc aggctcgctt tatcttagag aaaatagatg gcaaaataat cattgaaaat	3360
aagcctaaga aagaattaat taaagttctg attcagaggg gatatgattc ggatcctgtg	3420

aaggcctgga aagaagccca gcaaaaggtt ccagatgaag aagaaaatga agagagtgac	3480
aacgaaaagg aaactgaaaa gagtgactcc gtaacagatt ctggaccaac cttcaactat	3540
cttcttgata tgcccctttg gtatttaacc aaggaaaaga aagatgaact ctgcaggcta	3600
agaaatgaaa aagaacaaga gctggacaca ttaaaaagaa agagtccatc agatttgttg	3660
aaagaagact tggctacatt tattgaagaa ttggaggctg ttgaagccaa ggaaaaacaa	3720
gatgaacaag tcggacttcc tgggaaaggg gggaaggcca aggggaaaaa aacacaaatg	3780
gctgaagttt tgccttctcc gcgtgggtcaa agagtcattc cacgaataac catagaaatg	3840
aaagcagagg cagaaaagaa aaataaaaag aaaattaaga atgaaaatac tgaaggaagc	3900
cctcaagaag atggtgtgga actagaaggc ctaaaacaaa gattagaaaa gaaacagaaa	3960
agagaaccag gtacaaagac aaagaaacaa actacattgg catttaagcc aatcaaaaaa	4020
ggaaagaaga gaaatccctg gtctgattca gaatcagata ggagcagtga cgaaagtaat	4080
tttgatgtcc ctccacgaga aacagagcca cggagagcag caacaaaaac aaaattcaca	4140
atggatttgg attcagatga agatttctca gattttgatg aaaaaactga tgatgaagat	4200
tttgtcccat cagatgctag tccacctaag accaaaactt ccccaaaact tagtaacaaa	4260
gaactgaaac cacagaaaag tgtcgtgtca gaccttgaag ctgatgatgt taagggcagt	4320
gtaccactgt cttcaagccc tcctgctaca catttcccag atgaaactga aattacaaac	4380
ccagttccta aaaagaatgt gacagtgaag aagacagcag caaaaagtca gtcttccacc	4440
tccactaccg gtgccaaaaa aagggtgcc caaaaggaa ctaaaaggga tccagctttg	4500
aattctggtg tctctcaaaa gcctgatcct gccaaaacca agaatcgccg caaaaggaag	4560
ccatccactt ctgatgattc tgactctaatt tttgagaaaa ttgtttcgaa agcagtcaca	4620
agcaagaaat ccaaggggga gagtgatgac ttccatatgg actttgactc agctgtggct	4680
cctcgggcaa aatctgtacg ggcaaagaaa cctataaagt acctggaaga gtcagatgaa	4740
gatgatctgt tttaaaatgt gaggcgatta ttttaagtaa ttatcttacc aagccaaga	4800
ctggttttta agttacctga agctcttaac ttctcccct ctgaatttag tttggggaag	4860
gtgttttttag tacaagacat caaagtgaag taaagcccaa gtgttcttta gctttttata	4920
atactgtcta aatagtgacc atctcatggg cattgttttc ttctctgctt tgtctgtgtt	4980
ttgagtctgc tttcttttgt ctttaaaacc tgatttttaa gttcttctga actgtagaaa	5040

tagctatctg atcacttcag cgtaaagcag tgtgttttatt aaccatccac taagctaaaa	5100
ctagagcagt ttgatttaaa agtgtcactc ttcctccttt tctactttca gtagatatga	5160
gatagagcat aattatctgt tttatcttag ttttatacat aatttaccat cagatagaac	5220
tttatgggtc tagtacagat actctactac actcagcctc ttatgtgcca agtttttctt	5280
taagcaatga gaaattgctc atgttcttca tcttctcaaa tcatcagagg ccgaagaaaa	5340
acactttggc tgtgtctata acttgacaca gtcaatagaa tgaagaaaat tagagtagtt	5400
atgtgattat ttcagctctt gacctgtccc ctctggctgc ctctgagtct gaatctccca	5460
aagagagaaa ccaatttcta agaggactgg attgcagaag actcggggac aacatttgat	5520
ccaagatctt aaatgttata ttgataacca tgctcagcaa tgagctatta gattcatttt	5580
gggaaatctc cataatttca atttgtaaac tttgttaaga cctgtctaca ttgttatatg	5640
tgtgtgactt gagtaatgtt atcaacgttt ttgtaaatat ttactatgtt tttctattag	5700
ctaaattcca acaattttgt actttaataa aatgttctaa acattgcaac ccatagaa	5758

<210> 96

<211> 5389

<212> DNA

<213> Homo sapiens

<220>

<223> >TOP2B|ENSG00000077097|ENST00000435706|5389

<400> 96

tttgagggca gccggcggcg cggcctcctc agcgggctcg gctggacgtc cgctccggat	60
cttcgcgatg gggcgcgggg gtcggcgcgg ctaggagtgc ggcgagtgga gcggtgggtg	120
cggagcggcg gggcccagcg gcccgcaggg aggcgggagc ggcggctgcg gcctcagggc	180
ctgtgagctg gaggcactcg ccatggccaa gtcgggtggc tgcggcgcgg gagccggcgt	240
gggcggcggc aacggggcac tgacctgggt gaacaatgct gcaaaaaaag aagagtcaga	300
aactgccaac aaaaatgatt cttcaaagaa gttgtctgtt gagagagtgt atcagaagaa	360
gacacaactt gaacacattc ttcttcgtcc tgatacatat attgggtcag tggagccatt	420
gacgcagttc atgtgggtgt atgatgaaga tgtaggaatg aattgcaggg aggttacctt	480
tgtgccaggt ttatacaaga tctttgatga aattttggtt aatgctgctg acaataaaca	540
gagggataag aacatgactt gtattaaagt ttctattgat cctgaatcta acattataag	600

catttggaat aatgggaaag gcattccagt agtagaacac aaggtagaga aagtttatgt	660
tcctgcttta atttttggac agcttttaac atccagtaac tatgatgatg atgagaaaaa	720
agttacaggt ggtcgtaatg gttatgggtgc aaaactttgt aatattttca gtacaaagtt	780
tacagtagaa acagcttgca aagaatacaa acacagtttt aagcagacat ggatgaataa	840
tatgatgaag acttctgaag ccaaaattaa acattttgat ggtgaagatt acacatgcat	900
aacattccaa ccagatctgt ccaaatttaa gatggaaaaa cttgacaagg atattgtggc	960
cctcatgact agaagggcat atgatttggc tggttcgtgt agaggggtca aggtcatgtt	1020
taatggaaag aaattgcctg taaatggatt tcgcagttat gtagatcttt atgtgaaaga	1080
caaattggat gaaactgggg tggccctgaa agttattcat gagcttgcaa atgaaagatg	1140
ggatgtttgt ctacattga gtgaaaaagg attccagcaa atcagctttg taaatagtat	1200
tgcaactaca aaaggtggac ggcacgtgga ttatgtggta gatcaagttg ttggtaaact	1260
gattgaagta gttaagaaaa agaacaaagc tgggtgtatca gtgaaaccat ttcaagtaaa	1320
aaaccatata tgggttttta ttaattgcct tattgaaaat ccaacttttg attctcagac	1380
taaggaaaac atgactctgc agcccaaaag ttttgggtct aaatgccagc tgtcagaaaa	1440
attttttaaa gcagcctcta attgtggcat tgtagaaagt atcctgaact gggtgaaatt	1500
taaggctcag actcagctga ataagaagtg ttcacagta aaatacagta aaatcaaagg	1560
tattcccaaa ctggatgatg ctaatgatgc tgggtggtaaa cattccctgg agtgtacact	1620
gatattaaca gagggagact ctgccaaatc actggctgtg tctggattag gtgtgattgg	1680
acgagacaga tacggagttt ttccactcag gggcaaaatt cttaatgtac gggaagcttc	1740
tcataaacag atcatggaaa atgctgaaat aaataatatt attaaaatag ttggtctaca	1800
atataagaaa agttacgatg atgcagaatc tctgaaaacc ttacgctatg gaaagattat	1860
gattatgacc gatcaggatc aagatggttc tcacataaaa ggcctgctta ttaatttcat	1920
ccatcacaat tggccatcac ttttgaagca tggttttctt gaagagttca ttactcctat	1980
tgtaaaggca agcaaaaata agcaggaact ttccttctac agtattcctg aatttgacga	2040
atggaaaaaa catatagaaa accagaaagc ctggaaaata aagtactata aaggattggg	2100
tactagtaca gctaaagaag caaaggaata ttttgctgat atggaaaggc atcgcatctt	2160
gtttagatat gctggtcctg aagatgatgc tgccattacc ttggcattta gtaagaagaa	2220

gattgatgac agaaaagaat ggttaacaaa ttttatggaa gaccggagac agcgtaggct	2280
acatggccta ccagagcaat ttttatatgg tactgcaaca aagcatttga cttataatga	2340
tttcatcaac aaggaattga ttctcttctc aaactcagac aatgaaagat ctataccatc	2400
tcttggtgat ggctttaaac ctggccagcg gaaagtttta tttacctggt tcaagaggaa	2460
tgataaacgt gaagtaaaag ttgccagtt ggctggctct gttgctgaga tgtcggccta	2520
tcatcatgga gaacaagcat tgatgatgac tattgtgaat ttggctcaga actttgtggg	2580
aagtaacaac attaacttgc ttcagcctat tggtcagttt ggaactcggc ttcatggtgg	2640
caaagatgct gcaagccctc gttatatat tacaatgtta agcactttag caaggctact	2700
ttttcctgct gtggatgaca acctccttaa gttcctttat gatgataatc aacgtgtaga	2760
gcctgagtgg tatattccta taattcccat ggttttaata aatggtgctg agggcattgg	2820
tactggatgg gcttgtaaac taccacaacta tgatgctagg gaaattgtga acaatgtcag	2880
acgaatgcta gatggcctgg atcctcatcc catgcttcca aactacaaaa actttaaagg	2940
cacgattcaa gaacttggtc aaaaccagta tgcagtcagt ggtgaaatat ttgtagtgga	3000
cagaaacaca gtagaaatta cagagcttcc agttagaact tggacacagg tatataaaga	3060
acaggtttta gaacctatgc taaatggaac agataaaaaca ccagcattaa tttctgatta	3120
taaagaatat cactactgaca caactgtgaa atttgtggtg aaaatgactg aagagaaaact	3180
agcacaagca gaagctgctg gactgcataa agtttttaaa cttcaaacta ctcttacttg	3240
taattccatg gtactttttg atcatatggg atgtctgaag aaatatgaaa ctgtgcaaga	3300
cattctgaaa gaattctttg atttacgatt aagttattac ggtttacgta aggagtggct	3360
tgtgggaatg ttgggagcag aatctacaaa gcttaacaat caagcccgtt tcattttaga	3420
gaagatacaa gggaaaatta ctatagagaa taggtcaaag aaagatttga ttcaaagtgt	3480
agtccagaga ggttatgaat ctgaccagct gaaagcctgg aaagaagcac aagaaaaggc	3540
agcagaagag gatgaaacac aaaaccagca tgatgatagt tcctccgatt caggaaactcc	3600
ttcaggccca gattttaatt atattttaaa tatgtctctg tggctcttta ctaaagaaaa	3660
agttgaagaa ctgattaaac agagagatgc aaaagggcga gaggtcaatg atcttaaaag	3720
aaaatctcct tcagatcttt ggaaagagga tttagcggca tttgttgaag aactggataa	3780
agtggaatct caagaacgag aagatgttct ggctggaatg tctggaaaag caattaaagg	3840

taaagttggc aaacctaagg tgaagaaact ccagttggaa gagacaatgc cctcacctta	3900
tggcagaaga ataattcctg aaattacagc tatgaaggca gatgccagca aaaagttgct	3960
gaagaagaag aagggtgatc ttgatactgc agcagtaaaa gtggaatttg atgaagaatt	4020
cagtggagca ccagtagaag gtgcaggaga agaggcattg actccatcag ttcctataaa	4080
taaaggtccc aaacctaaga gggagaagaa ggagcctggt accagagtga gaaaaacacc	4140
tacatcatct ggtaaacctg gtgcaaagaa agtgaagaaa cggaatcctt ggtcagatga	4200
tgaatccaag tcagaaagtg atttggaaga aacagaacct gtggttattc caagagattc	4260
tttgcttagg agagcagcag ccgaaagacc taaatacaca tttgatttct cagaagaaga	4320
ggatgatgat gctgatgatg atgatgatga caataatgat ttagaggaat tgaaagttaa	4380
agcatctccc ataacaaatg atggggaaga tgaatttggt ccttcagatg ggtagataa	4440
agatgaatat acattttcac caggcaaadc aaaagccact ccagaaaaat ctttgcatga	4500
caaaaaaagt caggattttg gaaatctctt ctcatttcct tcatattctc agaagtcaga	4560
agatgattca gctaaatttg acagtaatga agaagattct gcttctgttt tttcaccatc	4620
atgtggtctg aaacagacag ataaagtcc aagtaaacg gtagctgcta aaaagggaaa	4680
accgtcttca gatacagtcc ctaagcccaa gagagcccaa aaacagaaga aagtagtaga	4740
ggctgtaaac tctgactcgg attcagaatt tggcattcca aagaagacta caacacccaa	4800
aggtaaaggc cgaggggcaa agaaaaggaa agcatctggc tctgaaaatg aaggcgatta	4860
taaccctggc aggaaaacat ccaaaacaac aagcaagaaa ccgaagaaga catcttttga	4920
tcaggattca gatgtggaca tcttcccctc agacttcctt actgagccac cttctctgcc	4980
acgaaccggt cgggctagga aagaagtaaa atattttgca gagtctgatg aagaagaaga	5040
tgatgttgat tttgcaatgt ttaattaagt gcccaaagag cacaacatt tttcaacaaa	5100
tatcttgtgt tgtccttttg tcttctctgt ctcagacttt tgtacatctg gcttatttta	5160
atgtgatgat gtaattgacg gttttttatt attgtggtag gccttttaac attttgttct	5220
tacacataca gttttatgct cttttttact cattgaaatg tcacgtactg tctgattggc	5280
ttgtagaatt gttatagact gccgtgcatt agcacagatt ttaattgtca tggttacaaa	5340
ctacagacct gctttttgaa atgaaattta aacattaaaa atggaactg	5389

<210> 97
 <211> 5378
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >TOPBP1|ENSG00000163781|ENST00000260810|5378

<400> 97
 ggggtagggg cggcgccgag tcgggggagg gggctgtgcg cgggctggc gcccgacccc 60
 agccaccgcc ctgcggccag cgcgcccc gactcgccgc ccggagaccc cgaggctcca 120
 acgagttcag aaatgtccag aaatgacaaa gaaccgtttt ttgtgaagtt tttaaagtct 180
 tcagacaatt ccaaattgtt ttttaaagct ctcgagtcca taaaagaatt ccaatcagaa 240
 gaatatcttc agattattac agaagaagag gcattgaaga taaaggagaa tgatagatca 300
 ctttatactt gtgacccttt tagtggcggt gtctttgatc acctcaaaaa gcttggctgc 360
 agaattgttg gtcctcaagt agtcatatct tgtatgcacc accagcgatg tgtcccaaga 420
 gccgaacatc cagtttataa tatggttatg tctgatgtaa ccatatcttg tacaagtctg 480
 gaaaaagaaa aaaggggaaga agttcataaa tatgtacaaa tgatgggcgg acgagtatac 540
 agagacctta atgtatcagt aactcacctt attgcaggag aagttggtag caaaaaatat 600
 ttagttgctg caaacctgaa gaaacctatt ttgcttcctt cttggataaa aacacttttg 660
 gagaagtcac aagagaaaaa aataactaga tatactgata taaacatgga agatttcaag 720
 tgtcctatct ttcttggttg cataatctgt gtgactggct tatgtggctt agacaggaaa 780
 gaagttcagc aactcacagt taagcatgga ggtcaatata tgggacaatt gaaaatgaat 840
 gaatgtacac acctcattgt gcaagaacca aaaggtcaga agtatgagtg tgccaagaga 900
 tggaatgtac actgtgtgac cacacagtgg ttttttgaca gtattgagaa aggtttttgt 960
 caggatgaat ccatatacaa gacagaacct agaccagaag caaagactat gcccaattct 1020
 tcaactccta ccagccagat caacacaatt gatagtcgta ctctttcaga tgtcagcaat 1080
 atttccaaca taaatgcaag ttgcgtaagt gaatcaatat gtaattcact taacagcaaa 1140
 ctggagccta cacttgaaaa tctagaaaat ctggatgtca gtgcatttca agcacctgaa 1200
 gatttattag atggttgtcg gatatatctt tgcggtttta gtggcagaaa gctagataaa 1260
 ctgagaagac ttattaacag tggaggtgga gttcgtttta accagctaaa tgaagatgta 1320

actcatgtta ttgtgggaga ttatgatgat gaattgaagc agttttggaa taaatcagcc	1380
cacaggcctc atgtagtggg agcaaagtgg ttgctagagt gtttcagtaa aggttatatg	1440
ctttctgaag aaccatatat ccatgctaata taccagccag tggaaattcc agtttcacat	1500
aagcctgaaa gtaaagcagc tcttttaaaa aagaagaaca gcagcttctc taagaaagac	1560
tttgctccta gtgaaaagca tgagcaagct gatgaagatc tgctctctca atatgaaaat	1620
ggtagctcca cagtagttga ggctaagacg tctgaagcca ggccctttaa tgattctact	1680
catgctgagc ctttgaatga ttctactcac atttctttgc aagaagaaaa ccagtcttct	1740
gtcagtcatt gtgtccctga tgtttctaca attactgaag aaggcttatt tagccaaaag	1800
agtttccttg ttttgggttt tagtaatgaa aatgaatcta acatcgcaaa catcataaaa	1860
gaaaatgctg ggaaaatcat gtcccttctg agcagaactg ttgcggatta tgctgtggtt	1920
cctctgctgg ggtgtgaagt ggaagccact gtgggagaag ttgttacaaa tacatggctg	1980
gttacttgca tagactatca gactttgttt gatccaaagt cgaatcctct cttcacacca	2040
gttccagtaa tgacaggaat gactccttta gaggattgtg ttatttcatt tagccagtgt	2100
gctggagcag aaaaagagtc tttaacattc ctagcaaacc tccttgagc aagtgttcaa	2160
gaatactttg ttcgcaaadc caatgcaaag aaaggcatgt ttgccagtac tcatcttata	2220
ctgaaagaac gtggtggctc taaatatgaa gctgcaaaga agtggaattt acctgccgtt	2280
actatagctt ggctgttgga gactgctaga acgggaaaga gagcagacga aagccatttt	2340
ctgattgaaa attcaactaa agaagaacga agtttggaac cagaaataac aaatggaatc	2400
aatctaaatt cagatactgc agagcatcct ggcacacgcc tgcaaactca cagaaaaacc	2460
gtcgttacac ctttagatat gaaccgcttt cagagtaaag ctttccgtgc tgtggtctca	2520
caacatgcca gacaggctgc agcctcccca gcagtaggac aaccacttca gaaggagccc	2580
tcgttacacc tggatacacc atcaaaattc ctgtccaagg acaaactctt caagccttcc	2640
tttgatgtga aggatgcact tgcagccttg gaaactccag gacgtcccag ccaacagaaa	2700
aggaaaccga gtacgccact ctcagaagtt attgtcaaaa acttgcaact tgctttggca	2760
aatagctctc gaaatgctgt cgctctttct gccagccctc aactgaaaga ggcccagtca	2820
gagaaggaag aagcccaaaa gccacttcac aaagtagtgg tatgtgttag taaaaaactc	2880
agtaagaagc agagtgaact aaatgggatc gcagcctctc taggagcaga ttacaggtgg	2940

agttttgatg aaacagtgac tcatttcatc tatcaagggc ggccaaatga cactaatcgg	3000
gagtataaat ctgtaaaaga aagaggagta cacattgttt ccgagcactg gcttttagat	3060
tgtgccaag agtgtaaaca tcttcctgaa tctctttatc cacatactta taatcccaaa	3120
atgagcttgg atatcagcgc agtgcaagat ggccggctct gtaatagtcg actacttca	3180
gctgtgtctt caacaaagga tgatgagcca gatcctttga ttttagaaga aaatgatgta	3240
gacaatatgg ccaccaataa taaagagtca gcaccatcaa atggaagtgg aaagaatgac	3300
tctaaaggag ttctgacaca gaccttagag atgagagaga actttcagaa gcagttacag	3360
gagataatgt ctgcaacatc aatagtgaac cccaagggc agaggacttc cttttcaaga	3420
agtggttgta acagcgcac ttcaaccct gacagcactc gctctgctcg cagtggacga	3480
agtagagtcc tagaggcact gaggcagtct cgtcagacag tacctgatgt caacacagag	3540
ccttcccaaa atgaacagat catttgggat gaccctacag caaggaggga gagagcaagg	3600
cttgccagca atttgcatg gcctagttgt cccacacaat actctgagct tcaggttgac	3660
attcaaaact tggaggattc tccttttcaa aagcctttac atgattcaga aattgctaaa	3720
caggctgtct gtgatcctgg aaacatacgt gtgactgaag ctcccaaaca cccaatctct	3780
gaagaactgg aaactcccat aaaagacagc cacctgatcc ctacgcctca agccccagt	3840
attgcctttc cactcgcaa cccccctgtg gctccgcacc ctagagaaaa gattataacg	3900
atagaggaga ctcatgaaga attaaaaaaa cagtacatat ttcagttatc atctctgaat	3960
cctcaagaac gtattgacta ttgtcatctg attgagaaac taggtggatt ggtgatagaa	4020
aagcagtgtt ttgatccac ctgtacacac attgttgtgg gacatccact tcgaaacgag	4080
aagtatttag cctcagtggc agctgggaag tgggtgcttc atcgctccta ctttgaagcc	4140
tgcaggactg ctggacactt cgtgcaggaa gaagactatg aatggggaag tagttccata	4200
cttgatgttc tgactggaat caatgtacag caacgaagac tagcacttgc agcaatgaga	4260
tggagaaaaa aaatccagca aagacaagaa tctggcattg ttgagggagc atttagtggg	4320
tggaaggtta ttttacatgt ggatcagtct cgagaagcag gcttcaaacg ctttcttcag	4380
tcaggaggag caaagggtgt acctgggtcat tctgtacctt tatttaaaga ggccacacat	4440
cttttttctg acttgaataa actgaaacca gatgactcag gagttaatat agcagaagct	4500
gctgcccaga acgtgtactg cttgagaaca gaatacattg ctgattatct catgcaggaa	4560

tcacctcctc atgtagaaaa ttactgtcta ccagaagcta tttcatttat tcagaataat	4620
aaggaacttg ggactggatt atcacaaaag aggaaagctc ctacagaaaa aaataaaaatc	4680
aaacgacctg gagtacacta atcgcatcta ccctttagtt accaaacatt aaatgttttt	4740
aaaaattgaa agcctgaatg tgactgtgat agatttgggt agtaatttaa agatgagtac	4800
ctgaagaatt ctgcttcaga gtataatgat gacccttctt gagttttgaa cacctgaaat	4860
tgtaatcact gaaatattaa ctgtttctta ataaaaagtt acctgaaata acaacaaaat	4920
acaactcctc agctagcttg ctgttaaacc acattgaagt ctgttaaaag atatttattt	4980
ttcttgtaaa tatctgaagc tgtagcttag tggaaatttt agcaaggtaa tggattttgc	5040
tttaaaatgt ctgccttaca aattcataac aacaagattt gtcagtcagc atttattcat	5100
gttttccctg atttttatct tctcaccatt ttacctcttt taacaggagc ctgagcacia	5160
ggtttaatga ggaagctggg gctataaata tgtgtgtata tatgtatatg tatgtttgta	5220
caaactctca tgatgtttgc caagtttgaa tgcgcaaac ttggaaaatg tgacaataaa	5280
gaataaaagt agtaactcaa attagtatta agatgtgttt acatagataa attttttaaa	5340
agagcaccct ggcatctgtt cttctgttga agcaacaa	5378

<210> 98

<211> 6231

<212> DNA

<213> Homo sapiens

<220>

<223> >TP53BP1|ENSG00000067369|ENST00000382044|6231

<400> 98

gtgacgggaa agggggagtt cgcgccggt ggcggcggtg gcgacagcgg cgacctaggg	60
atcgatctgg agggacttgg ggagcgtgca gagacctcta gctcgagcgc gagggacctc	120
ccgccgggat gcctggggag cagatggacc ctactggaag tcagttggat tcagatttct	180
ctcagcaaga tactccttgc ctgataattg aagattctca gcctgaaagc caggttctag	240
aggatgattc tggttctcac ttcagtatgc tatctcgaca ccttcctaatt ctccagacgc	300
acaaagaaaa tcctgtgttg gatgttgtgt ccaatcctga acaaacagct ggagaagaac	360
gaggagacgg taatagtggg ttcaatgaac atttgaaaga aaacaagggt gcagaccctg	420
tggattcttc taacttggac acatgtggtt ccatcagtca ggtcattgag cagttacctc	480

agccaaacag gacaagcagt gttctgggaa tgtcagtgga atctgctcct gctgtggagg	540
aagagaaggg agaagagttg gaacagaagg agaaagagaa ggaagaagat acttcaggca	600
atactacaca ttcccttgggt gctgaagata ctgcctcatc acagttgggt tttggggttc	660
tggaactctc ccagagccag gatgttgagg aaaatactgt gccatatgaa gtggacaaag	720
agcagctaca atcagtaacc accaactctg gttataccag gctgtctgat gtggatgcta	780
atactgcaat taagcatgaa gaacagtcca acgaagatat ccccatagca gaacagtcca	840
gcaaggacat ccctgtgaca gcacagccca gtaaggatgt acatgttgta aaagagcaaa	900
atccaccacc tgcaaggta gaggacatgc cttttagccc caaagcatct gttgctgcta	960
tggaagcaaa agaacagttg tctgcacaag aacttatgga aagtggactg cagattcaga	1020
agtcaccaga gcctgagggt ttgtcaactc aggaagactt gtttgaccag agcaataaaa	1080
cagtatcttc tgatggttgc tctactcctt caagggagga aggtgggtgt tctttggctt	1140
ccactcctgc caccactctg catctcctgc agctctctgg tcagagggtcc cttgttcagg	1200
acagtctttc cacgaattct tcagatcttg ttgctccttc tcctgatgct ttccgatcta	1260
ctccttttat cgttcctagc agtcccacag agcaagaagg gagacaagat aagccaatgg	1320
acacgtcagt gttatctgaa gaaggaggag agccttttca gaagaaactt caaagtgggtg	1380
aaccagtgga gttagaaaac cccctctctc tgcctgagtc cactgtatca ccacaagcct	1440
caacaccaat atctcagagc acaccagtct tccctcctgg gtcacttcct atcccatccc	1500
agcctcagtt ttctcatgac atttttattc cttccccaag tctggaagaa caatcaaag	1560
atgggaagaa agatggagat atgcatagtt catctttgac agttgagtgt tctaaaactt	1620
cagagattga accaaagaat tcccctgagg atcttgggct atctttgaca ggggattctt	1680
gcaagttgat gctttctaca agtgaatata gtcagtcccc aaagatggag agcttgagtt	1740
ctcacagaat tgatgaagat ggagaaaaca cacagattga ggatacggaa cccatgtctc	1800
cagttctcaa ttctaaattt gttcctgctg aaaatgatag tatcctgatg aatccagcac	1860
aggatgggtga agtacaactg agtcagaatg atgacaaaac aaaggagat gatacagaca	1920
ccagggatga cattagtatt ttagccactg gttgcaaggg cagagaagaa acggtagcag	1980
aagatgtttg tattgatctc acttgatgatt cggggagtca ggcagttccg tcaccagcta	2040
ctcgatctga ggcactttct agtgtgttag atcaggagga agctatggaa attaaagaac	2100

accatccaga ggaggggtct tcagggctctg aggtggaaga aatccctgag acaccttgtg	2160
aaagtcaagg agaggaactc aaagaagaaa atatggagag tgttccgttg cacctttctc	2220
tgactgaaac tcagtcccaa gggttgtgtc ttcaaaagga aatgccaaaa aaagaatgct	2280
cagaagctat ggaagttgaa accagtgtga ttagtattga ttcccctcaa aagttggcaa	2340
tacttgacca agaattggaa cataaggaac aggaagcttg ggaagaagct acttcagagg	2400
actccagtgt tgtcattgta gatgtgaaag agccatctcc cagagttgat gtttcttgtg	2460
aacctttgga gggagtggag aagtgctcag attcccagtc atgggaggat attgctccag	2520
aaatagaacc atgtgctgag aatagattag acaccaagga agaaaagagt gtagaatatg	2580
aaggagatct gaaatcaggg actgcagaaa cagaacctgt agagcaagat tcttcacagc	2640
cttccttacc tttagtgaga gcagatgatc ctttaagact tgaccaggag ttgcagcagc	2700
cccaaactca ggagaaaaca agtaattcat taacagaaga ctcaaaaatg gctaattgcaa	2760
agcagctaag ctcatatgca gaggcccaga agctggggaa gccctctgcc catgcctcac	2820
aaagcttctg tgaaagttct agtgaaacct catttcattt cactttgcct aaagaaggtg	2880
atatcatccc accattgact ggtgcaacct cacctcttat tgggcaccta aaattggagc	2940
ccaagagaca cagtactcct attggtatta gcaactatcc agaaagcacc atagcaacca	3000
gtgatgtcat gtctgaaagc atggtggaga cccatgatcc catacttggg agtggaagaa	3060
gggatttctg ggctgcccc gacgtggatg ataaattatg tctaagaatg aaactggtta	3120
gtcctgagac tgaggcgagt gaagagtctt tgcagttcaa cctggaaaag cctgcaactg	3180
gtgaaagaaa aaatggatct actgctgttg ctgagtctgt tgccagtccc cagaagacca	3240
tgtctgtgtt gagctgtatc tgtgaagcca ggcaagagaa tgaggctcga agtgaggatc	3300
ccccaccac acccatcagg gggaacttgc tccactttcc aagtttctcaa ggagaagagg	3360
agaaagaaaa attggagggt gaccatacaa tcaggcagag tcaacagcct atgaagccca	3420
ttagtcctgt caaggaccct gtttctcctg cttcccagaa gatggtcata caagggccat	3480
ccagtcctca aggagaggca atggtgacag atgtgctaga agaccagaaa gaaggacgga	3540
gtactaataa ggaaaatcct agtaaggcct tgattgaaag gccagccaa aataacatag	3600
gaatccaaac catggagtgt tccttgaggg tcccagaaac tgtttcagca gcaaccaga	3660
ctataaagaa tgtgtgtgag caggggacca gtacagtgga ccagaacttt ggaaagcaag	3720

atgccacagt tcagactgag aggggggagtg gtgagaaacc agtcagtgct cctgggggatg	3780
atacagagtc gctccatagc cagggagaag aagagtttga tatgcctcag cctccacatg	3840
gccatgtctt acatcgtcac atgagaacaa tccgggaagt acgcacactt gtcactcgtg	3900
tcattacaga tgtgtattat gtggatggaa cagaagtaga aagaaaagta actgaggaga	3960
ctgaagagcc aattgtagag tgtcaggagt gtgaaactga agtttcccct tcacagactg	4020
ggggctcctc aggtgacctg ggggatatca gtccttctc ctccaaggca tccagcttac	4080
accgcacatc aagtgggaca agtctctcag ctatgcacag cagtgggaagc tcagggaag	4140
gagccggacc actcagaggg aaaaccagcg ggacagaacc cgcagatttt gccttaccca	4200
gctcccaggagg agggccagga aaactgagtc ctagaaaagg ggtcagtcag acagggacgc	4260
cagtgtgtga ggaggatggt gatgcaggcc ttggcatcag acagggaggg aaggctccag	4320
tcacgcctcg tgggcgtggg cgaaggggcc gccaccttc tcggaccact ggaaccagag	4380
aaacagctgt gcctggcccc ttgggcatag aggacatttc acctaacttg tcaccagatg	4440
ataaatcctt cagccgtgtc gtgccccgag tgccagactc caccagacga acagatgtgg	4500
gtgctgggtgc tttgcgtcgt agtgactctc cagaaattcc tttccaggct gctgctggcc	4560
cttctgatgg cttagatgcc tcctctccag gaaatagctt tgtagggctc cgtgtttag	4620
ccaagtggtc atccaatggc tacttttact ctgggaaaat cacacgagat gtcggagctg	4680
ggaagtataa attgctcttt gatgatgggt acgaatgtga tgtgttgggc aaagacattc	4740
tgttatgtga ccccatcccg ctggacactg aagtgacggc cctctcggag gatgagtatt	4800
tcagtgcagg agtggtgaaa ggacatagga aggagtctgg ggaactgtac tacagcattg	4860
aaaaagaagg ccaaagaaag tgggtataagc gaatggctgt catcctgtcc ttggagcaag	4920
gaaacagact gagagagcag tatgggcttg gccctatga agcagtaaca cctcttataa	4980
aggcagcaga tatcagctta gacaatttgg tggaaggga gcggaacgg cgcagtaacg	5040
tcagctcccc agccaccct actgcctcca gtagcagcag cacaaccct acccgaaaga	5100
tcacagaaag tcctcgtgcc tccatgggag ttctctcagg caaaagaaaa cttatcattt	5160
ctgaagagga acgggtcccct gccaagcgag gtcgcaagtc tgccacagta aaacctggtg	5220
cagtaggggc aggagagttt gtgagcccct gtgagagtgg agacaacacc ggtgaaccct	5280
ctgccctgga agagcagaga gggcctttgc ctctcaacaa gaccttgttt ctgggctacg	5340

cattttctcct taccatggcc acaaccagtg acaagttggc cagccgctcc aaactgccag	5400
atggtcctac aggaagcagt gaagaagagg aggaatTTTT ggaaattcct cttttcaaca	5460
agcagtatac agaatcccag cttcgagcag gagctggcta tatlcttgaa gatttcaatg	5520
aagcccagtg taacacagct taccagtgtc ttctaattgc ggatcagcat tgtcgaaccc	5580
ggaagtactt cctgtgcctt gccagtggga ttcttTgtgt gtctcatgtc tgggtccatg	5640
atagttgcca tgccaaccag ctccagaact accgtaatta tctgttgcca gctgggtaca	5700
gccttgagga gcaaagaatt ctggactggc aaccccgtga aaatcctttc cagaatctga	5760
aggtactctt ggtatcagac caacagcaga acttcctgga gctctggtct gagatcctca	5820
tgactgggtg tgagcctct gtgaagcagc accattcaag tgcccataac aaagatatTg	5880
ctttaggggt atttgatgtg gtggTgacgg acccctcatg cccagcctcg gtgctgaagt	5940
gtgctgaagc attgcagctg cctgtggTgt cacaagagtg ggtgatccag tgcctcattg	6000
ttggggagag aattggattc aagcagcatc caaaatataa acacgattat gtttctcact	6060
aaagatactt ggtcttactg gttttattcc ctgctatcgt ggagattgtg ttttaaccag	6120
gttttaaatg tgtctTgtgt gtaactggat tccttgcatg gatcttgtat atagttttat	6180
ttgctgaact tttatgataa aataaatgtt gaatctcttt ggTttagta a	6231

<210> 99

<211> 2579

<212> DNA

<213> Homo sapiens

<220>

<223> >TP53|ENSG00000141510|ENST00000269305|2579

<400> 99

gttttcccct cccatgtgct caagactggc gctaaaagtt ttgagcttct caaaagtcta	60
gagccaccgt ccaggagca ggtagctgct gggctccggg gacactttgc gttcgggctg	120
ggagcgtgct ttccacgacg gtgacacgct tccctggatt ggagccaga ctgccttccg	180
ggtcactgcc atggaggagc cgcagtcaga tcctagcgtc gagccccctc tgagtcagga	240
aacattttca gacctatgga aactacttcc tgaaaacaac gttctgtccc ccttgccgtc	300
ccaagcaatg gatgatttga tgctgtcccc ggacgatatt gaacaatggt tcaactgaaga	360

cccaggtcca gatgaagctc ccagaatgcc agaggctgct ccccccgtgg cccctgcacc	420
agcagctcct acaccggcgg cccctgcacc agccccctcc tggcccctgt catcttctgt	480
cccttcccag aaaacctacc agggcagcta cggtttccgt ctgggcttct tgcattctgg	540
gacagccaag tctgtgactt gcacgtactc ccctgccctc aacaagatgt tttgccaact	600
ggccaagacc tgccctgtgc agctgtgggt tgattccaca ccccgcccg gcacccgcgt	660
ccgcgccatg gccatctaca agcagtcaca gcacatgacg gaggttgtga ggcgctgccc	720
ccaccatgag cgctgctcag atagcgatgg tctggcccct cctcagcatc ttatccgagt	780
ggaaggaaat ttgcgtgtgg agtatttggga tgacagaaac acttttcgac atagtgtggt	840
ggtgccctat gagccgcctg aggttggctc tgactgtacc accatccact acaactacat	900
gtgtaacagt tcctgcatgg gcggcatgaa ccggaggccc atcctcacca tcatcacact	960
ggaagactcc agtggtaatc tactgggacg gaacagcttt gaggtgcgtg tttgtgcctg	1020
tcctgggaga gaccggcgca cagaggaaga gaatctccgc aagaaagggg agcctcacca	1080
cgagctgccc ccaggagca ctaagcgagc actgcccac aacaccagct cctctcccca	1140
gccaaagaag aaaccactgg atggagaata tttcacctt cagatccgtg ggcgtgagcg	1200
cttcgagatg ttccgagagc tgaatgaggc cttggaactc aaggatgccc aggctgggaa	1260
ggagccaggg gggagcaggg ctactccag ccacctgaag tccaaaaagg gtcagtctac	1320
ctcccgccat aaaaaactca tgttcaagac agaagggcct gactcagact gacattctcc	1380
acttcttgtt cccactgac agcctccac ccccatctct ccctcccctg ccattttggg	1440
ttttgggtct ttgaaccctt gcttgcaata ggtgtgcgtc agaagcacc aggacttcca	1500
tttgctttgt cccggggctc cactgaacaa gttggcctgc actggtgttt tgttgtgggg	1560
aggaggatgg ggagtaggac ataccagctt agattttaag gtttttactg tgagggatgt	1620
ttgggagatg taagaaatgt tcttgcagtt aagggttagt ttacaatcag ccacattcta	1680
ggtaggggcc cacttcaccg tactaaccag ggaagctgtc cctcactgtt gaattttctc	1740
taacttcaag gcccatatct gtgaaatgct ggcatttgca cctacctcac agagtgcatt	1800
gtgagggtta atgaaataat gtacatctgg ccttgaaacc accttttatt acatggggctc	1860
tagaacttga ccccttagg ggtgcttgtt ccctctccct gttggtcggt gggttggtag	1920
tttctacagt tgggcagctg gttaggtaga gggagttgtc aagtctctgc tggcccagcc	1980

aaaccctgtc tgacaacctc ttggtgaacc ttagtaccta aaaggaaatc tcaccccatc	2040
ccacaccctg gaggatttca tctcttgtat atgatgatct ggatccacca agacttgttt	2100
tatgctcagg gtcaatttct tttttctttt tttttttttt ttttcttttt ctttgagact	2160
gggtctcgct ttgttgccca ggctggagtg gagtggcgtg atcttggctt actgcagcct	2220
ttgcctcccc ggctcgagca gtcctgcctc agcctccgga gtagctggga ccacaggttc	2280
atgccaccat ggccagccaa cttttgcatg tttttagag atggggcttc acagtgttgc	2340
ccaggctggg ctcaaactcc tgggctcagg cgatccacct gtctcagcct cccagagtgc	2400
tgggattaca attgtgagcc accacgtcca gctggaagg tcaacatctt ttacattctg	2460
caagcacatc tgcattttca cccaccctt cccctccttc tcccttttta tatccattt	2520
ttatatcgat ctcttatttt acaataaaac tttgctgcca cctgtgtgtc tgaggggtg	2579

<210> 100
 <211> 12677
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >TRRAP|ENSG00000196367|ENST00000359863|12677

<400> 100	
cgcgccgggg cctggtgctc ggtcggcggg tgctgccgt ttaagcgggg gcgggactgc	60
gcgcggccga gcggttgca cgagggctcg gctgggggtc gccggggctc cgggccgggc	120
ctgcaggagc cgggccgccg aggtcggggc tggttgaact catggacctg atacttttct	180
cttgagaagc aaaccagccc aaaagaaaaa tggcgtttgt tgcaacacag ggggccacgg	240
tggttgacca gaccatttg atgaaaaagt accttcagtt tgtggcagct ctcacagatg	300
tgaatacacc tgatgaaaca aagttgaaaa tgatgcaaga agttagtga aattttgaga	360
atgtcacgtc atctcctcag tattctacat tcctagaaca tatcatcct cgattcctta	420
catttctcca agatggagaa gttcagtttc ttcaggagaa accagcacag caactgcgga	480
agctcgtact tgaaataatt catagaatac caaccaacga acatcttcgt cctcacacaa	540
aaaatgtttt gtctgtgatg tttcgctttt tagagacgga aatgaagaa aatgttctta	600
tttgtctaag aataattatt gagctacaca aacagttcag gccaccgatc acacaagaaa	660
ttcatcattt tctggatttt gtgaaacaga tttaacagga gcttccaaaa gtagtgaacc	720

gctactttga gaaccctcaa gtgatccccg agaacacagt gcctcccca gaaatggttg	780
gtatgataac aacgattgct gtgaaagtca acccggagcg tgaggacagt gagactcgaa	840
cacattccat cattccgagg ggatcacttt ctctgaaagt gttggcagaa ttgccatta	900
ttgttgtttt aatgtatcag ctctacaaac tgaacatcca caatgttggt gctgagtttg	960
tgcccttgat catgaacacc attgccattc aggtgtctgc acaagcgagg caacataagc	1020
tttacaacaa ggagttgtat gctgacttca ttgctgctca gattaaaaca ttgtcatttt	1080
tagcttacat tatcaggatt taccaggagt tggtgactaa gtattctcag cagatggtga	1140
aaggaatgct ccagttactt tcaaattgtc cagcagagac tgcacacctc agaaaggagc	1200
ttctgattgc tgccaaacac atcctcacca cagagctgag aaaccagttc attccttgca	1260
tggacaagct gtttgatgaa tccatactaa ttggctcagg atatactgcc agagagactc	1320
taaggcccct cgcctacagc acgctggccg acctcgtgca ccatgtccgc cagcacctgc	1380
ccctcagcga cctctccctc gccgtccagc tcttcgcaa gaacatcgac gatgagtccc	1440
tgcccagcag catccagacc atgtcctgca agctcctgct gaacctggtg gactgcatcc	1500
gttccaagag cgagcaggag agtggcaatg ggagagacgt cctgatgcgg atgctggagg	1560
ttttcgttct caaattccac acaattgtc ggtaccagct ctctgccatt tttagaagt	1620
gtaagcctca gtcagaactt ggagccgtgg aagcagctct gcctggggtg cccactgccc	1680
ctgcagctcc tggccctgct ccctccccag cccctgtccc tgccccacct ccacccccgc	1740
ccccacccc acctgccacc cctgtgacct cggccccgt gcctcccttc gagaagcaag	1800
gagaaaagga caaggaagac aagcagacat tccaagtcac agactgtcga agtttggtca	1860
aaaccttggt gtgtggtgtc aagacaatca cgtggggcat aacatcatgc aaagcacctg	1920
gtgaagctca gttcattccc aacaagcagt tacaaccaa agagacacag atttacatca	1980
aacttgtaga atatgcaatg caagctttag atatttatca ggtccagata gcaggaaatg	2040
gacagacata catccgtgtg gccaaactgcc agactgtgag aatgaaagag gagaaggagg	2100
tattggagca tttcgctggt gtgttcacaa tgatgaacct cttaacgttc aaagaaatct	2160
tccaaactac ggtcccttat atggtggaga gaatctcaa aaattatgct cttcagattg	2220
ttgccaattc cttcttgga aatcctacta cctctgctct gtttgctacg attctggtgg	2280
aatatctcct tgatcgctg ccagaaatgg gctccaacgt ggagctctcc aacctgtacc	2340

tcaagctgtt caagctggtc tttggctctg tctccctctt tgcagctgaa aatgaacaaa	2400
tgctgaagcc tcacttgcac aagattgtga acagctctat ggagctcgcg cagactgcca	2460
aggaacccta caactacttc ttgctgctac gggcgctgtt tcgctctatt ggtggaggta	2520
gccacgatct cttgtatcag gagttcttgc ctctccttcc aaacctcctg caagggctga	2580
acatgcttca gagtggcctg cacaagcagc acatgaagga cctctttgtg gagctgtgtc	2640
tcaccgtccc tgtgctggctg agctcgcttt tgccgtacct gcccatgctt atggatccct	2700
tggtgtctgc actcaatggg tctcagacat tggtcagcca aggcctcagg acgctggagc	2760
tgtgtgtgga caacctgcag cccgacttcc tctacgacca catccagccg gtgcgcgcag	2820
agctcatgca ggctctgtgg cgcaccttac gcaaccctgc tgacagcatc tcccacgtgg	2880
cctaccgtgt gctcggtgtaag tttggcggca gtaacaggaa gatgctgaag gagtgcgaga	2940
agctgcacta cgttgtgacc gaggttcagg gccccagcat cactgtggag ttttccgact	3000
gcaaagcttc tctccagctc cccatggaga aggccattga aactgctctg gactgcctga	3060
aaagcgccaa cactgagccc tactaccgga ggcaggcgtg ggaagtgatc aaatgcttcc	3120
tggtggccat gatgagcctg gaggacaaca agcacgcact ctaccagctc ctggcacacc	3180
ccaactttac agaaaagacc atccccaatg ttatcatctc acatcgctac aaagcccagg	3240
acactccagc ccggaagact tttgagcagg ccctgacagg cgccttcatg tctgctgtca	3300
ttaaggacct gcggcccagc gccctgccct ttgtcgccag cttgatccgc cactatacga	3360
tggtggcagt cgcccagcag tgtggccctt tcttgctgcc ttgctaccag gtgggcagcc	3420
agcccagcac agccatgttt cacagtgaag aaaatggctc gaaaggaatg gatcctttgg	3480
ttctcattga tgcaattgct atttgtatgg catatgaaga aaaggagctt tgcaaaatcg	3540
gggaggtggc cctagctgtg atatttgatg ttgcaagtat catcctgggc tccaaggaga	3600
gggcctgcca gctgcccctg ttttcttaca tcgtggagcg cctgtgtgca tgttgttatg	3660
aacaggcgtg gtatgcaaag ctgggggggtg tggtgtctat taagtttctc atggagcggc	3720
tgctctcac ttgggttctc cagaaccagc agacattcct gaaagcactt ctctttgtca	3780
tgatggactt aactggagag gtttccaatg gggcagtcgc tatggcaaag accacgtgg	3840
agcagcttct gatgcggtgc gcaacgcctt taaaagacga ggagagagcc gaagagatcg	3900
tggccgcccga ggaaaagtct ttccaccatg tgacacacga cttggttcga gaagtcacct	3960

ctccaaactc cactgtgagg aagcaggcca tgcattcgct gcaggtgttg gcccaggtca	4020
ctgggaagag tgtcacggtg atcatggaac cccacaaaga ggtcctgcag gatatggtcc	4080
cccctaagaa gcacctgctc cgacaccagc ctgccaacgc acagattggc ctgatggagg	4140
ggaacacgtt ctgtaccacg ttgcagccca ggctcttcac aatggacctt aacgtggtgg	4200
agcataaggt gttctacaca gagctgttga atttgtgtga ggctgaagat tcagctttaa	4260
caaagctgcc ctgttataaa agccttccgt cactcgtacc tttacgaatt gcggcattaa	4320
atgcacttgc tgcctgcaat taccttcctc agtccaggga gaaaatcatc gctgcactct	4380
tcaaagccct gaattccacc aatagtgagc tccaagaggc cggagaagcc tgtatgagaa	4440
agtttttaga aggtgctacc atagaagtcg atcaaatcca cacacatatg cgacctttgc	4500
tgatgatgct gggagattac cggagcttga cgctgaatgt tgtgaatcgc ctgacttcgg	4560
tcacgaggct cttcccaaatt tccttcaatg ataaattttg tgatcagatg atgcaacatc	4620
tgcgcaagtg gatggaagtg gtggtgatca cccacaaagg gggccagagg agcgacggaa	4680
acgaaagcat ttccgagtgc gggagatgtc ctttgtctcc attctgtcag tttgaggaaa	4740
tgaagatttg ctcagcaatt ataaaccttt ttcattctgat cccggctgct cctcagacac	4800
tggtgaagcc tttgctagag gttgtcatga aaacggagcg ggcgatgctg atcgaggcgg	4860
ggagtccatt ccgagagccc ctgatcaagt tcctgactcg acatccctcg cagacagtgg	4920
agctgttcat gatggaagcc aactgaacg atccccagtg gagcagaatg tttatgagtt	4980
ttttaaaaca caaagacgcc agacctctgc gggatgtgct ggctgccaac cccaacaggt	5040
tcattaccct gctgctgccg ggggggtgcc agacggctgt gcgccccggt tcgcccagca	5100
ccagcaccat gcgcctggac ctccagttcc aggccatcaa gatcataagc attatagtga	5160
aaaacgatga ctcttggtg gccagccagc actctctggt gagccagttg cgacgtgtgt	5220
gggtgagtga gaacttcaa gagaggcacc gcaaggagaa catggcagcc accaactgga	5280
aggagcccaa gctgctggcc tactgcctgc tgaactactg caaaaggaat tacggagata	5340
tagaattgct gttccagctg ctccgagcct ttactggtcg ttttctctgc aacatgacat	5400
tcttaaaaga gtatatggag gaagagattc caaaaatta cagcatcgct cagaaacgtg	5460
ccctgttctt tcgctttgta gacttcaacg accccaactt cggagatgaa ttaaaagcta	5520
aagttctgca gcatacttg aatcctgctt tcttgtacag ctttgagaag ggggaaggag	5580

agcagctctt	gggacctccc	aatccagaag	gagataaccc	agaaagcatc	accagtgtgt	5640
ttattaccaa	gttcctggac	cccgagaagc	aggcggacat	gctggactcg	ctgcggatct	5700
acctgctgca	gtacgccacg	ctgctgggtgg	agcacgcccc	ccaccacatc	catgacaaca	5760
acaagaaccg	caacagcaag	ctgcgccgcc	tcatgacctt	cgcttgcccc	tgctgtctct	5820
ccaaggcctg	cgtggacca	gcctgcaagt	acagcggaca	cttgctcctg	gcgcacatta	5880
tcgccaaatt	cgccatacac	aagaagatcg	tcctgcaggt	ttttcatagt	ctcctcaagg	5940
ctcacgcaat	ggaagctcga	gcgatcgtca	gacaggcgat	ggccattctg	accccggcgg	6000
tgccggccag	gatggaggac	gggcaccaga	tgctgacca	ctggaccg	aagatcattg	6060
tggaggaggg	gcacaccgtc	ccgcagctgg	tccacattct	gcacctgata	gtgcaacact	6120
tcaaggtgta	ctacccggta	cggcaccact	tggtgcagca	catggtgagc	gccatgcaga	6180
ggctgggctt	cacgcccagt	gtcaccatcg	agcagaggcg	gctggccgtg	gacctgtctg	6240
aagtcgtcat	caagtgggag	ctgcagagga	tcaaggacca	gcagccgat	tcagatatgg	6300
acccaaattc	cagtggagaa	ggagtcaatt	ctgtctcatc	ctccattaag	agaggcctgt	6360
ccgtggattc	tgcccaggaa	gtgaaacgct	ttaggacggc	caccggagcc	atcagtgcag	6420
tctttgggag	gagccagtcg	ctacctggag	cagactctct	cctcgccaag	cccattgaca	6480
agcagcacac	agacactgtg	gtgaacttcc	ttatccgcgt	ggcctgtcag	gttaatgaca	6540
acaccaacac	agcgggggtcc	cctggggagg	tgctctctcg	ccggtgtgtg	aaccttctga	6600
agactgcgtt	gcggccagac	atgtggccca	agtccgaact	caagctgcag	tggttcgaca	6660
agctgctgat	gactgtggag	cagccaaacc	aagtgaacta	tgggaatatc	tgcacgggcc	6720
tagaagtgct	gagcttcctg	ctaactgtcc	tccagtcccc	agccatcctc	agtagcttca	6780
aacctctgca	gcgtggaatt	gccgcctgca	tgacatgtgg	aaacaccaag	gtgttgcgag	6840
ccgtccacag	ccttctctcg	cgctgatga	gcattttccc	aacagagccg	agtacttcca	6900
gtgtggcctc	caaatatgaa	gagctggagt	gcctctacgc	agccgtcgga	aaggtcatct	6960
atgaagggct	caccaactac	gagaaggcca	ccaatgcaa	tccctcccag	ctcttcggga	7020
cccttatgat	cctcaagtct	gcctgcagca	acaaccccag	ctacatagac	aggctgatct	7080
ccgtctttat	gcgctcccctg	cagaagatgg	tccgggagca	tttaaaccct	caggcagcgt	7140
caggaagcac	cgaagccacc	tcaggtacaa	gcgagctggt	gatgctgagt	ctggagctgg	7200

tgaagacgcg cctggcagtg atgagcatgg agatgcggaa gaacttcac	7260
tgacatccct catcgaaaa tcaccagatg ccaaaatcct cggggtgtg	7320
tggaagaatg ggtcaagaat aactcccaa tggcagccaa tcagacac	7380
agaagtccat tttgcttgtg aagatgatga cttacataga aaaacg	7440
ttgaattaaa tgcccagttt ttagatcttg ttaactatgt ctacagg	7500
ctggcagcga gctgacggcg aaacttgagc ctgcctttct ctctgg	7560
agccactcat cagggcaaag tttttcgagg tttttgacaa ctccatg	7620
acgagcgctt gctctatgtg acctgttcgc agaactggga agccatg	7680
ggatcaagca gtgcattgag ctgcttctgg ccgtgtgtga gaagagc	7740
ccagctgcca aggagccatg ctcccgtcca tcaccaacgt catcaac	7800
acgaccgtgc cgccttcgcc atggtcacac atgtcaagca ggagccc	7860
acagcgagtc caaagaggag gatgtagaga tagacatcga actagct	7920
ccagcacgcc caaaaccaa gaactttcag aaaaggacat tggaaacc	7980
taaccaacag gcacgacaag tttctggaca ctctccgaga ggtgaag	8040
tcagcgcttt cgttcagctg tgccacattt ccacgacgt ggcagaga	8100
agcttttccc cagattgtgg aagatcctct ctgacagaca gcagcat	8160
agataagtcc atttctgtgc agcggcagtc accagggtga gcgggac	8220
cgctgaactg ctttgtggaa gccatgtccc agtgcgtgcc gccaatc	8280
gcgtcctgaa gtacctgggg aagacacaca acctctggtt ccggtcc	8340
agcaccaggc ttttgaaaag ggtctgagtc ttcagattaa gccgaag	8400
tttatgagca ggagagcatc accccgccgc agcaggagat actggatt	8460
tttactccct gttacaagag gaagatatgt gggctggtct gtggcaga	8520
actcggagac agcgactgcg attgcttacg agcagcacgg gttctttg	8580
aatcctatga aaaggcaatg gataaagcca aaaaagaaca tgagaggag	8640
ctgctatttt ccctgaatac cagctctggg aagaccactg gattcgat	8700
tgaaccagtg ggaagccctg acggagtacg gtcagtccaa aggccac	8760
tcgtcctgga gtgcgcctgg cgggtgtcca actggactgc catgaagg	8820

aggtggaagt gagctgtccg aaggagatgg cctggaaggt gaacatgtac cgcggataacc	8880
tggccatctg ccaccccag gagcagcagc tcagcttcat cgagcgcctg gtggagatgg	8940
ccagcagcct ggccatccgc gagtggcggc ggctgcccc a cgtagtgtcc cacgtgcaca	9000
cgcctctcct acaggcagcc cagcaaatca tcgaactcca ggaagctgca caaatcaacg	9060
caggcttaca gccaaccaac ctgggaagga acaacagcct gcacgacatg aagacggtgg	9120
tgaagacctg gaggaaccga ctgcccacg tgtctgacga cttgtccac tggagcagca	9180
tcttcatgtg gaggcagcat cattaccagg gtaaaccgac ctggtccggc atgcattcat	9240
catcgattgt aactgcctat gagaatagct ctacagcatga tcccagttca aataacgcta	9300
tgcttggggg tcatgcatca gcttcagcga tcatccagta tggaaaaatc gcccggaac	9360
aaggactggg caatgtagct ctggatatat taagtcggat tcatactatt ccaactgttc	9420
ctatcgtgga ttgcttcag aagattcgac agcaagttaa atgctacctc cagctggcag	9480
gcgtcatggg caaaaacgag tgcattgcagg gccttgaagt tattgaatct acaaatttaa	9540
aatacttcac aaaagagatg acagccgaat tttatgcact gaagggaatg ttcttggtc	9600
agatcaacaa gtccgaggag gcaaacaag ctttctctgc agctgtgcag atgcacgatg	9660
tgctggtgaa agcctggggc atgtggggcg actacctgga gaacatcttt gtgaaggagc	9720
ggcagctgca cctgggctg tctgccatca cctgctacct gcacgcctgc cggcatcaga	9780
acgagagcaa atcgaggaaa tacttagcca aggtgctgtg gcttttgagt tttgatgatg	9840
acaaaaacac tttggcagat gccgtcgaca agtactgcat tgggtgtgcca cccatccagt	9900
ggctggcctg gatcccacag ctgctcacct gcctgggttg ctcggaggga aagctgctct	9960
tgaacctcat tagccaggtt ggacgcgtgt atccccaagc ggtctacttt cccatccgga	10020
ccctgtacct gaccctgaaa atagaacagc gggaacgcta caagagcgat ccagggccca	10080
taagagcaac agcaccatg tggcgctgca gccgaatcat gcacatgcag cgagagctcc	10140
accccaccct tctgtcttcc ctggaaggca tcgtcgatca gatgggtctgg ttacagagaaa	10200
attggcatga agaggttctc aggcagctcc aacagggcct ggcgaaatgt tactccgtgg	10260
cgtttgagaa aagtggagcg gtgtccgatg ctaaaatcac cccccacact ctcaattttg	10320
tgaagaagtt ggtgagcacg tttggggtgg gcctggagaa tgtgtccaac gtctcgacca	10380
tgttctccag cgcagcctct gagtctctgg cccggcgggc gcaggccact gcacaagacc	10440

ctgtctttca gaagctgaaa ggccagttca cgacggattt tgacttcagc gttccaggat	10500
ccatgaagct tcataatctt atttctaagt tgaaaaagt gatcaaaatc ttggaggcca	10560
agaccaagca actcccaaaa ttcttcctca tagaggaaaa gtgccggttc ttgagcaatt	10620
tctcggcaca gacagctgaa gtggaaattc ctggggagtt tctgatgcca aagccaacgc	10680
attattacat caagattgca cggttcatgc cccgggtaga gattgtgcag aagcacaaca	10740
ccgcagcccg gcggctgtac atccggggac acaatggcaa gatctacca tacctcgtca	10800
tgaacgacgc ctgcctcaca gagtcacggc gagaggagcg tgtgttgag ctgctgcgtc	10860
tgctgaaccc ctgtttggag aagagaaagg agaccaccaa gaggcacttg tttttcacag	10920
tgccccgggt tgtggcagtt tccccacaga tgcgcctcgt ggaggacaac ccctcttcac	10980
tttcccttgt ggagatctac aagcagcgt gcgccaagaa gggcatcgag catgacaacc	11040
ccatctcccg ttactatgac cggctggcta cgggtgcaggc gcggggaacc caagccagcc	11100
accaggtcct ccgcgacatc ctcaaggagg ttcagagtaa catggtgccg cgcagcatgc	11160
tcaaggagtg ggcgctgcac accttcccc aatgccacgga ctactggacg ttccggaaga	11220
tgttcaccat ccagctggct ctgataggct tcgcggaatt cgtcctgcat ttaaatagac	11280
tcaaccccga gatgttacag atcgctcagg acactggcaa actgaatgtt gcctactttc	11340
gatttgacat aaacgacgc actggagacc tggatgcaa ccgtcctgtc ccatttcgac	11400
tcacgcccaa catttctgag tttctgacca ccatcgggggt ctccggcccg ttgacagcgt	11460
ccatgattgc ggtcgcccgg tgcttcgccc agccaaactt taaggtggat ggcattctga	11520
aaacggttct ccgggacgag atcattgctt ggcacaaaaa aacacaagag gacacgtcct	11580
ctcctctctc ggccgccggg cagccagaga acatggacag ccagcaactg gtgtccctgg	11640
ttcagaaagc cgtcaccgcc atcatgaccc gcctgcacaa cctcgcccag ttcgaaggcg	11700
gggaaagcaa ggtgaacacc ctggtggccg cggcaaacag cctggacaat ctgtgccgca	11760
tggaccccg cctggacccc tggctgtgac tgtggccgcc acggccaccc ggaatgtgaa	11820
gggcgctccg ggctctgagc ccgcagcttt tacgacttct ccctgcctcg ttccttatat	11880
tcacagaagc cccatagttt cactgggttg cggttatttt cctggtagtt tgcgtgtaag	11940
aaaggagaa tatagtttta gaggaagctg aactatgacg atgctgggcg aagcggttgg	12000
aatggcaga gctgaaactt attccaagct ttcaaaataa tcttttaaga agccaggatt	12060

ctccggtctg gaatttctga gtgagtcctt tttttatggg gtcctccctc tgtgaatgta	12120
caggcggaac tgtacgaaca gctcccttcc atccattttt aactctttcg gaaataacac	12180
ctcacagcag cttcgtgctt ttgtacagac ctttgtaaca agtgtacaga aaactcattt	12240
tgtttgagaa acaggagttg atgaacccat catgctgggt tttctctgag cacaaagttt	12300
taggctgtac acagccagcc ttgggaatct cgttgagcgt tcggcgtgga tccacggggc	12360
caggccaccc tgcgggagcg ccacacgcat ccacttcgga ttcagtgggt gaagacagaa	12420
ctctgagagt ctgcaggcgg ctcctgtgct ttttatttct ggctcttcgg atgtcttcta	12480
gacatttact atcactgcac ctgaagaaaa aatcactttt accttcctaa tttaaaaaga	12540
caaaacagaa atgtacgttc cttcgctagc tttagtcttt ctgttcccat ttttataaat	12600
ctgagcattg ataatgttct atctaaattt gtacagtgtg attttttttt ttagaataaa	12660
tattttataa aagggtt	12677

<210> 101
 <211> 3671
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >UBE2N|ENSG00000177889|ENST00000318066|3671

<400> 101	
ggttaagaga gacgcgcgcg cagtcgcgcg cgggtcgtgc cgtaccaccg tcgcgggcag	60
gctcggccac gagcgccaga gccccgcgcc tcccctcgcg gcctgtccca agtccctgcc	120
ccgcaacaga gcgtcacttc cgccatcccc ggcagcgggtt ggggcggggc gcacggggga	180
gggggcccagg tcggagggaa gcccgccgt gcccgagccc gcgcccgagc agggactaca	240
tttcccaggg ggcctcggcg gcggctgcgg cgacgggcgc ggcaacgtcc cccggaagtg	300
gagcccggga cttccactcg tgcgtgaggc gagaggagcc ggagacgaga ccagaggccg	360
aactcggggt ctgacaagat ggccgggctg ccccgagga tcatcaagga aaccagcgt	420
ttgctggcag aaccagttcc tggcatcaaa gccgaaccag atgagagcaa cgcccgttat	480
tttcatgtgg tcattgctgg ccctcaggat tcccccttg agggaggac ttttaaactt	540
gaactattcc ttccagaaga atacccaatg gcagccccta agtacgttt catgaccaa	600

atttatcatc ctaatgtaga caagttggga agaatatgtt tagatatattt gaaagataag	660
tggtccccag cactgcagat ccgcacagtt ctgctatcga tccaggcctt gttaagtgtc	720
cccaatccag atgatccatt agcaaagat gtagcggagc agtggaagac caacgaagcc	780
caagccatag aaacagctag agcatggact aggctatatg ccatgaataa tatttaaatt	840
gatacgatca tcaagtgtgc atcacttctc ctgttctgcc aagacttcct cctctttgtt	900
tgcatttaat ggacacagtc ttagaaacat tacagaataa aaaagcccag acatcttcag	960
tcctttggtg attaaatgca cattagcaaa tctatgtctt gtcctgattc actgtcataa	1020
agcatgagca gaggctagaa gtatcatctg gattgttgtg aaacgtttaa aagcagtggc	1080
ccctccctgc ttttattcat ttccccatc ctggtttaag tataaagcac tgtgaatgaa	1140
ggtagttgtc aggttagctg caggggtgtg ggtgttttta ttttatttta ttttatttta	1200
tttttgaggg gggaggtagt ttaattttat gggctccttt cccccttttt tggtgatcta	1260
attgcattgg ttaaaagcag ctaaccaggt ctttagaata tgctctagcc aagtctaact	1320
ttatttagac gctgtagatg gacaagcttg attgttggaa ccaaaatggg aacattaaac	1380
aaacatcaca gccctcacta ataacattgc tgtcaagtgt agattccccc cttcaaaaaa	1440
agcttgtgac cattttgtat ggcttgtctg gaaacttctg taaatcttat gttttagtaa	1500
aatatttttt gttattctac tttgcctttg tacagtttat tttactgtgt ttatttcatt	1560
ttccaattt gacaatcgta ttttaaaatt gaaactgatg gaacattctt tcttggtctt	1620
caccatctga caaattgaat ggcaagaggt ggattttgcc agtttctttt cactgatgca	1680
gatttgtgtt aagatagtag tgaatggagt atttataaac tggccctgag catgcataaa	1740
gcatcagtat ctgacctttt tttaaccttc taggaatttg aaataaatgt gtttgtgttg	1800
tctgattaga tgatcattgg tgtcttgcca caatgtttaa aaattactgt acaggaaagt	1860
cacagcaaag atagcagttg tgactgacat gtaggacttt cacagttgtg ccacattttt	1920
gcctaaaatt tgggttatga catttttctt ggttcttatc tgaaaatttc atctgtaacc	1980
tttcatgtgt gttaagaaac actgatctga tcatttggga tttgctgagg catttgtgag	2040
tcttccttat aaacctgatg agcagatctc aactatctag cttgtgtgtc atcagaaagg	2100
tttatccctt tgagagtatc aagtcctcag ttaatgattc ttgctttcat ccctccagta	2160
tttgctgtgg gagctcgttt tattctttta tttggaattc agtaattttt cttctttatt	2220

gacgaattcc tcccctcaca aaactgttct ttcccacctc tctccatata taattcctga	2280
ttcttggttat ttttaagtca taaatgtagc cagtcataaa tacataaatg ttaaccttcg	2340
ggttgcaacc ttgtctcttg cagtttaagg taatggatat tgtagcccat ttgaattttc	2400
ttcactctta ttctcgtaat tctggagttt cttcagattg tgggtgtattt tattgtgctc	2460
ctatgtaaga tgaagaatta actattaaaa ttacattttc aacatacaaa agcttttgat	2520
gactggtaac tggatcctt ccaaataaat gcattgcttg gtaacaaatg ttagcttaat	2580
tcagaatgaa atagttaaca accctaatat tgtaaaattt gaaaactgat agcagctgtc	2640
atttataaca attcttgggc tgtttaaagg gtacatttga ctatagctaa ggaaattgaa	2700
tctcacttta aaaatgaacc acacaaaact tcaagaaacg aaatagaaat caacatttac	2760
tgcccgtaat attgctcatg gttatagcta gtatcacata gtgatttaca aaccatgaaa	2820
ttccagattt aaatctgttg cactacaaaa atgtctaaat gatgtctgct aaatttgtga	2880
gttatagaag tacagaaata agtcattttc atttcaaagc cagtgaattt ctaatttcac	2940
ggtatcatat ttttttcctt tgcagttttg gtaagactcc cccattactt tgctaagact	3000
tgctttctca accagtggga tgaaacgagt gccagtgtgt tgactgagcc ttcttaggac	3060
ttaaggggct ctgggagagg ccaaggcagc agttgaaaac tgcagcttca cctggccttt	3120
gggcctgttt tccctggccc acacagtgtg gctgggttaa acattctgaa tttgatgcca	3180
acatttttaa aaattgggag atttgcatga aaatccagaa ttcgggtaac ttttgaatga	3240
tctgaagatc tggcaacagt agacctgacc cacctttgtg cacgggactt aactgagagg	3300
atctgagtag cagacaccct aagagcatgt cattgtcttc tcaccagtta ctctccagtt	3360
ggctagagtc ccctccgagc acacctgcta gtccttcccc cttgtaagcc tttgactttg	3420
taatccctgg gctaaaggaa acacattatt caacaaaagt gtttcatgtt tgcccactag	3480
ttgttgggga tgattaggtg gtgagtgaat ggcattggggc tttcctttgg ggaccatata	3540
ctctaattggg gaagggggac agacaaatag ttgtaactct gtcattgaaa agagaagtaa	3600
cttgacaga cttttggatt caggccagat gaacatctcc aggcagaggg aacagcagca	3660
tctgcaaagt c	3671

<210> 102

<211> 2570

<212> DNA

<213> Homo sapiens

<220>

<223> >UIMC1|ENSG00000087206|ENST00000377227|2570

<400> 102

gggctcggga tgcggggctg ggaccctccc gattccgggg cggattccgg acgccgggac	60
cggccattac tgggtgccggg ttgggcttct ccagatgccg gggctgggtc cttcccaagg	120
ttgagacaaa aggatgccac ggagaaagaa aaaagttaaa gaagtctccg aatctcggaa	180
cctggagaag aaggatgtgg aaactaccag ttctgtcagt gtgaagagga agcgtagact	240
tgaggatgca ttcattgtga tatccgatag tgatggagag gaaccaaagg aggaaaatgg	300
gttcagaaaa acgaagacaa aacagtcgaa tagagcaaag tgtttgcca aaagaaaaat	360
cgcacagatg acagaagaag aacagtttgc tctggctctc aaaatgagtg agcaggaagc	420
tagggaggtg aacagccagg aggaggaaga agaggagctc ttgaggaaag ccattgctga	480
aagcctgaat agttgccggc cttctgatgc ttccgctacc agatctcgac ctctggccac	540
tggaccgtct tcccagtccc atcaagagaa aaccacagac tctgggctca ctgaaggcat	600
atggcagctg gtacctccat cactgtttaa aggctcacat atcagtcagg gaaacgaggc	660
tgaggaaaga gaggagcctt gggaccacac tgaaaaaact gaagaggagc cggctctctgg	720
cagctcagga agctgggacc agtcaagcca gccagtgttt gagaatgtga acgttaaatac	780
ttttgacaga tgtactggcc actcggctga gcacacacag tgtgggaagc cacaggaaag	840
tactgggagg ggttctgctt ttctcaaagc tgtccagggt agcggggaca catctaggca	900
ctgtctacct accctagcag atgccaaagg tctccaggac actgggggca ctgtgaacta	960
tttctgggggt attccattct gccctgatgg agtagaccct aaccagtata ccaaggtcac	1020
tctctgccag ttggagggtt atcaaaagag cctgaaaatg gctcagaggc agctccttaa	1080
taaaaaaggt tttggggaac cagtgttacc tagacctctt tctctgatcc agaataatg	1140
tggccaagga gagcaggcta gtgagaaaaa tgaatgcata tcagaagata tgggagatga	1200
agacaaagag gagaggcagg agtctagggc atctgactgg cactcaaaaa ccaaggattt	1260
ccaggaaagc tcaattaaaa gcttgaaaga gaaacttttg ttggaggaag aaccaacaac	1320
cagtcatggg cagtcttccc aagggttgt tgaagaaact tctgaagagg gaaactctgt	1380
acctgcttca caaagtgttg ctgctttgac cagtaagaga agcttagtcc ttatgccaga	1440

gagttctgca gaagaaatca ctgtttgtcc tgagaccag ctaagttcct ctgaaacttt	1500
tgaccttgaa agagaagtct ctccaggtag cagagatatt ttggatggag tcagaataat	1560
aatggcagat aaggagggtg gtaacaagga agatgctgag aaggaagtag ctatttctac	1620
cttctcatcc agtaaccagg tctcctgccc gctatgtgac caatgctttc caccacaaa	1680
gattgaacga catgccatgt actgcaatgg tctgatggag gaagatacag tattgactcg	1740
gagacaaaaa gaggccaaga ccaagagtga cagtgggaca gctgccaga cttctctaga	1800
cattgacaag aatgagaagt gttacctctg taaatccctg gtccattta gagagtatca	1860
gtgtcatgtg gactcctgtc tccagcttgc aaaggctgac caaggagatg gacctgaagg	1920
gagtggaga gcatgttcaa ctgtggaggg gaagtggcag cagaggctga agaaccaaa	1980
ggaaaaaggc cacagtgaag gccgactcct tagtttcttg gaacagtctg agcacaagac	2040
ttcagatgca gacatcaagt cttcagaaac aggagccttc aggggtgcctt caccagggat	2100
ggaagaggca ggctgcagca gagagatgca gagttctttc acacgtcgtg acttaaatga	2160
atctcccgctc aagtcttttg tttccatttc agaagccaca gattgcttag tggactttaa	2220
aaagcaagtt actgtccagc caggtagtcg gacacggacc aaagctggca gaggaagaag	2280
gagaaaattc tgaatttcta gggtcacaaa gttgacaaaa ccattagtag gaggggtggg	2340
ccatgttcat taagccatag tgggtccctag ttcattgttg agcaagtttt agccctgcag	2400
ttttcaccac cagcacctac ccagcattct ggtttttatg ttttttatga tctatgcaga	2460
caactgtgta ttctgtttta taacagtttg tttgaattta cttacagtta aaaaatttaa	2520
atatatttat gtttgtacga aatcttattt caatagatgg aaattttaat	2570

<210> 103

<211> 3996

<212> DNA

<213> Homo sapiens

<220>

<223> >USP1|ENSG00000162607|ENST00000339950|3996

<400> 103

gtggcccatc ccgtctgctc tctgtcccgt ctcccagccc gtacgctttt ccctcaactc	60
---	----

gcgcgcacga atgggttaaaa aaagggtgtg aggctcgagc ccgccagcca ggccgctagc	120
--	-----

acctcgcgcg cgccctcagc gaggacccgc acgagctgcc cggctctgtg cctgcgttgt	180
ttgaaactgt ggaacccgtt aggcttttgg ggaactacag tcccagctgg ctccggagcc	240
cgtgcgtgcc aggggcgagt ggccgtcgcg agacggctcc tgcctttcgt gtctctgcag	300
cgtggagact ggaaccggca atttcaaagg acgccacgtt caatcgcagc gctggcgcgg	360
gcggaggcta aaacacgggg gtcctgagac tgaggaaaac gcgccaagtt cccctcgggtg	420
gcggagtgtt aaagacccta gcggttcagg cgttcggcga gcggggccgc tgcttgttgc	480
gtcctgggtt ctcccggggc gggcgcagat gggcgccgct cccgggatgt agttggtgtt	540
ggtgcaagac gggagcgcgc ggcggtcggg gttcccgctc ttgggagcgg atggtcactc	600
ccccgcgggg agggcgagcc gaccagattt tcctggggcc ggggacccgg cgggctcggg	660
gcagggactc acctgtcgca cccacactca ttcgggttgg acttgccggc gtcaccgccg	720
cggacttcgc tttgggcat gaccagatat aattggtgat tacaactttc ctctataaat	780
taactcttga cactccttgg gatttgaaga aaaaaatgcc tgggtgcata cctagtgaag	840
gtaatggact ttcaagaggt agcccttcaa agaaaaacag actttcctta aagttttttc	900
agaaaaagga aactaagaga gctttggatt tcacagattc tcaagaaaat gaagaaaaag	960
cttctgaata tagagcatct gaaattgatc aagttgttcc tgcagcacag tcttcaccta	1020
taaactgtga gaagagagaa aacttggttac catttgtggg actgaataat ctcggcaata	1080
cttgctatct taatagtata cttcaggtat tatatttttg tcccggtttt aaatctggag	1140
taaagcactt atttaatat atttcaagga agaaagaagc tctaaaggat gaagccaatc	1200
aaaaagacaa gggaaattgc aaagaagatt ctttggcaag ttatgaattg atatgcagtt	1260
tacagtcctt aatcatttcg gttgaacagc tccaggctag ttttctctta aatccagaga	1320
aatatactga tgaacttgcc actcagccaa ggcgactgct taacacactg agggaactca	1380
accctatgta tgaaggatat ctacagcatg atgcacagga agtattacaa tgtatttttg	1440
gaaacattca agaaacatgc caactcctaa aaaaagaaga agtaaaaaat gtggcagaat	1500
tacctactaa ggtagaagaa atacctcatc cgaaagagga aatgaatggt attaacagca	1560
tagagatgga cagtatgagg cattctgaag acttttaaaga gaaactccca aaaggaaatg	1620
ggaaaagaaa aagtgcact gaatttggtt acatgaagaa aaaagttaaa ttatccaagg	1680
aacaccagtc attggaagag aaccagagac aaactagatc aaaaagaaaa gctacaagtg	1740

atacattaga	gagtcctcct	aaaataattc	ccaagtatat	ttctgaaaat	gagagtccaa	1800
gaccctcaca	aaagaaatca	agagttaaaa	taaattgggt	aaagtctgca	actaagcaac	1860
ccagcattct	ttctaaatth	tgtagtctgg	gaaaaataac	aacaaaccaa	ggagtcaaag	1920
gacaatctaa	agaaaaatga	tgtgatcctg	aagaggactt	ggggaagtgt	gaaagtgata	1980
acacaactaa	tggttggtga	cttgaatctc	caggaaatac	tgttacacct	gtaaatgtta	2040
atgaagttaa	acccataaac	aaaggtgaag	aacaaattgg	ttttgagcta	gtggagaaat	2100
tatttcaagg	tcagctggta	ttaaggacgc	gttgcttgga	atgtgaaagt	ttaacagaaa	2160
gaagagaaga	ttttcaagac	atcagtgtgc	cagtacaaga	agatgagctt	tccaaagtag	2220
aggagagttc	tgaaatthct	ccagagccaa	aaacagaaat	gaagaccctg	agatgggcaa	2280
tttcacaatt	tgcttcagta	gaaaggattg	taggagaaga	taaatatttc	tgtgaaaact	2340
gccatcatta	tactgaagct	gaacgaagtc	ttttgtttga	caaatgcct	gaagttataa	2400
ctattcattt	gaagtgctth	gctgctagt	gtttggagtt	tgattgttat	ggtggtggac	2460
ttccaagat	caacactcct	ttattgacac	ctcttaaatt	gtcactagaa	gaatggagca	2520
caaagccaac	taacgacagc	tatggattat	ttgcggttgt	gatgcatagt	ggcattacaa	2580
ttagtagtgg	gcattacact	gcttctgtta	aagtcactga	ccttaacagt	ttagaactag	2640
ataaaggaaa	ttttgtgggt	gaccaaattg	gtgaaatagg	taagccagaa	ccattgaatg	2700
aggaggaagc	aaggggtgtg	gttgagaatt	ataatgatga	agaagtgtca	attagagttg	2760
gtggaaatac	acagccaagt	aaagtthtga	acaaaaaaaa	tgtagaagct	attggacttc	2820
ttggaggaca	aaagagcaaa	gcagattatg	agctatacaa	caaagcctct	aatcctgata	2880
aggttgctag	tacagcgtth	gctgaaaata	gaaattctga	gactagtgat	actactggga	2940
cccatgaatc	tgatagaaac	aaggaatcca	gtgaccaaac	aggcattaat	attagtggat	3000
ttgagaacaa	aatttcatac	gtagtgcaaa	gcttaaagga	gtatgagggg	aagtggttgc	3060
tttttgatga	ttctgaagtc	aaagttactg	aagagaagga	ctttctgaat	tctctttccc	3120
cttctacatc	tcctacttct	actccttact	tgctattthta	taagaaatta	tagagtgagt	3180
gtattthtct	tgtgtatata	ttaaacacac	ccatacaaac	attggtaaag	ttgattacat	3240
caaagaatct	ttagcttatc	ttttgaagct	actggatatt	attggtctct	ctaggtthtt	3300
atataaatag	tgaaatthga	attactgaaa	accatgttaa	tttttagaac	tcattthtct	3360

cagtagagac tagtgatgca ttagcttctg ggaacaaact tgtatcggtt ctttaattaaa	3420
ttatccaaaa cggaggcatt taaacacttg gatttacacc agtcttttgt gtttgctttt	3480
taaaataaag tgctcgtatt tgtattctcc atattttgga gtaattatct acatgatgtt	3540
tatagttcct gtggtttttc acccaagaag cagaatctca ttcagtacat ttagttttat	3600
aagagtcatg aagctaaatc cttgggctat gtcagaggca caaagtctag aatgtgtgta	3660
ttcacaatgg tgtatgtaca ttttgtgcct tgattcactt agaagtgtct cagaaaacct	3720
ggacagttcg cttctacaca agaattttat atgtatttat gaagatgatt ctgtacccta	3780
gtatatcttt ttgggcatgg actaatttgt atctgtttta ctcatatctt gcacgatctg	3840
tatatagtac atcaaactta gaggtgtgac cttaaattta acttttttta aaaactggga	3900
ggtcaataaa atttaaaactg cttaaactatg tatatgaata tttgaatttt ttacttgtat	3960
atttttataa atacagctga gttttcttaa agcgaa	3996

<210> 104
 <211> 3707
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >WDR48|ENSG00000114742|ENST00000302313|3707

<400> 104	
aagtgacgtc ggagtgtcaa catgcaagat ggcggcccat caccggcaga acacagcagg	60
gcggaggaaa gtgcaggttt cctatgttat tcgagatgaa gtggagaagt acaaccgaaa	120
tggagtcaat gctctgcagc tggatccagc actaaataga cttttcacag ccggtcgaga	180
ctctatcata agaatatgga gtgtcaatca gcacaagcaa gatccatata tagcatctat	240
ggaacaccat actgattggg taaacgacat tgtactctgt tgtaatggga aaacattaat	300
atctgcttct tctgacacga cagtaaaagt atggaatgca cacaagggat tttgcatgtc	360
aacattaagg acacataagg attacgtaaa ggccttagca tatgccaagg ataaagaact	420
agtagcatca gctgggttgg acagacaaat attcctttgg gatgtgaata ctctaacagc	480
attgactgcc tcaaataaca ctgtcacaac ttcttcttta agtggaaaca aagattccat	540
ttatagcctg gccatgaatc aactgggaac aatcattgta tcagggtcca ctgaaaaggt	600
gttacgggta tgggatccaa gaacatgtgc aaaactaatg aagcttaaag ggcacacgga	660

taatgtgaag gcattgctat taaacagaga tggcacgcaa tgcctgtcag gcagttctga	720
tgggacaatt cgccttttggc cccttggcca gcagagatgt atagcaacat accgagtcca	780
tgatgaaggt gtttgggcgc tgcaagtcaa tgatgccttc acacatgtgt attctggtgg	840
aaggacagg aagatttatt gtacagacct aagaaacct gacattcggg tgctaatttg	900
tgaagaaaa gcaccagttc tcaagatgga gcttgataga tcagctgac ctcctcctgc	960
aatttgggtt gcaacaacta agtctacagt aaataaatgg actttgaaag gaattcataa	1020
ttttagagcc tctggagatt atgacaatga ctgtacaaat cctataacac ctctttgtac	1080
acaacctgac caggttatta aagggggtgc tagtattatt cagtgccaca ttcttaatga	1140
taagagacat atattaacca aagataccaa taataatgtg gcatattggg atgtattgaa	1200
ggcatgtaaa gttgaagatc tgggcaaagt ggattttgaa gatgaaatta agaaaagatt	1260
taaaatggtg tatgtgcaa attggttctc agtagactta aaaacaggga tgттаactat	1320
tactttggat gaaagtgatt gttttgctgc ctgggtttct gcaaaagatg ctggtttcag	1380
cagccctgat gggtcagatc caaaattgaa tttaggagga cttttactcc aagcactcct	1440
ggaatatggc cctagaacac atgtgaatcc aatggatgaa gaggaaaatg aagtaaacca	1500
tgtaaatggg gagcaggaga accgagtgc gaagggaat ggatattttc aagtgcctcc	1560
acatacacc gtgatctttg gtgaagctgg aggtcgcaca ctgttcaggc tgctctgccg	1620
agattccggg ggtgagactg agtctatgct tcttaatgaa acagtgccac aatgggtaat	1680
tgacatcact gtggataaaa atatgcccaa attcaacaaa attcctttct acctccaacc	1740
tcatgcatct tcaggagcaa aaaccttaaa aaaagataga ctctctgcta gtgacatgct	1800
ccaagtccga aaagttatgg aacatgttta tgaaaaaatt atcaacttgg ataatgagtc	1860
tcaaaccact agctcttcta ataataaaaa accaggagaa caggaaaaag aagaagatat	1920
tgctgtgttg gcagaggaga aaattgaact tttgtgccag gaccaggttt tggatccaaa	1980
tatggacctt cgaacagtga aacacttcat atggaagagc ggtggagacc tcacctcca	2040
ttaccgtcag aagtccacgt gaaggctggg ctaatgctcc tggatattca tttacgacct	2100
tcctctatgg cccaagagt agtcctagga agcccactga tcccaacgg gagcaagact	2160
tctaacggct gattggtatg gaccgagatt atctttcaat tgaagtgact aatcgagatg	2220
taatatagaa accagtctcc atgtgtagat taagctgtcc ccagggaagc agagtgaag	2280

agcagaagag ccccaagcag actatgtctt tcaagatgta cagaagactg acaacaggcc	2340
agtgcagact ttgcttcctc cttttgtttc aaaggacctt atctacccat taacacttgt	2400
tagagacacc tcagtcgggc cacaactgtc tctgtttcaa ctttaagccc acacactctg	2460
tagcttttct aaaacagtac attccattca tgcagtagat atgaaggctt tttccaagtt	2520
tgtcatataa agtaatcaaa ttgttttcac cgtttaagac agttatTTTT aatgggaatt	2580
tgctgttgca caatttgaga accagcccat ttcataataa aagattaatg ttgggcttgt	2640
tgtaatatatt gcatctcaga tgtgttcttg taggaagatc tgttgaaatt ccaggtgctg	2700
gatcatagca cttggactgt tgggccggag tttgtgcagc ccatgcctag cgatgctgca	2760
cctcccatga tgcctggct ttgtggcaac ccagggagg tgcagcagtt tcactccttc	2820
ctccactgtg tccagtttta ttacacaaa gcgagctgag accacagggtt cttgctctgg	2880
atgagcagca tgacaccatc aaaatttaag accatgatga aatttcagtt tcattcaaat	2940
gttacctaaa atttatgggtg tcagaataaa gggagatcat agtgagttaa tttgatattc	3000
atatcctgga agtatacata tttgtttttc caaagtttta tatgcagggtt tttgttgtag	3060
ctgtatccag atcttctttt cactgttcta acaatctaac actttcatag aatcatttgg	3120
atcttgtata gagtgttaact tattggggat aaacacttca acttttgga gaaaatactt	3180
tggattctcc ccaaggatcat ttgtattcag agtaaatcag tgggtccacc ccataaaaact	3240
gtaagacaag cctcctgtat gaagccagaa gcacaagtgc cttcagaggg cagttgattc	3300
cagtccagtt gcctccctag cttgtgtgtg gcctgcctcc accatagcag ggactaggga	3360
gggaggcagg gaacgttttt tcttttctac atcttcacag gttcggctgg gggcagctga	3420
taggcctaag gccatacctt actatttaag atactctgat cgcacaactg cagagacagg	3480
gttccctcac tgctagtcag cttcttttgg aaactggaca ggccattgcc acctgcactt	3540
tgaagaagtg agcaaagacg tggccacatt tctaataagt tgaaatggtc tttctccctt	3600
cctcagagag caaccagctt cttttttttt aaaagtcctt tctatctgtt ggatacaaca	3660
gtgtgctttt ggtgcacatt ttttagcaat agttgtgaat taatggc	3707

<210> 105

<211> 5215

<212> DNA

<213> Homo sapiens

<220>

<223> >WRN|ENSG00000165392|ENST00000298139|5215

<400> 105

ggagctgatg tgtacttgtt gcgccgggga ggcgccggct tgtactcggc agcgcgggaa	60
taaagtttgc tgatttggtg tctagcctgg atgcctgggt tgcaggccct gcttgtggtg	120
gcgctccaca gtcacccggc tgaagaagac ctgttggact ggatcttctc gggttttctt	180
tcagatattg ttttgtattt acccatgaag acattgtttt ttggactctg caaataggac	240
atttcaaaga tgagtgaata aaaattggaa acaactgcac agcagcggaa atgtcctgaa	300
tggatgaatg tgcagaataa aagatgtgct gtagaagaaa gaaaggcatg tgttcggaag	360
agtgtttttg aagatgacct ccccttctta gaattcactg gatccattgt gtatagttac	420
gatgctagtg attgctcttt cctgtcagaa gatattagca tgagtctatc agatggggat	480
gtggtaggat ttgacatgga gtggccacca ttatacaata gagggaaact tggcaaagtt	540
gcactaattc agtttgttgt ttctgagagc aaatgttact tgttccacgt ttcttccatg	600
tcagtttttc cccagggaatt aaaaatgttg cttgaaaata aagcagttaa aaaggcaggt	660
gtaggaattg aaggagatca gtggaaactt ctacgtgact ttgatataa attgaagaat	720
tttgtggagt tgacagatgt tgccaataaa aagctgaaat gcacagagac ctggagcctt	780
aacagtctgg ttaaacacct cttaggtaaa cagctcctga aagacaagtc tatccgctgt	840
agcaattgga gtaaatctc tctcactgag gaccagaaac tgtatgcagc cactgatgct	900
tatgctggtt ttattattta ccgaaattta gagatttttg atgatactgt gcaaaggttt	960
gctataaata aagaggaaga aatcctactt agcgacatga acaaacagtt gacttcaatc	1020
tctgaggaag tgatggatct ggctaagcat cttcctcatg ctttcagtaa attggaaaac	1080
ccacggaggg tttctatctt actaaaggat atttcagaaa atctatattc actgaggagg	1140
atgataattg ggtctactaa cattgagact gaactgaggc ccagcaataa tttaaactta	1200
ttatcctttg aagattcaac tactggggga gtacaacaga aacaaattag agaacatgaa	1260
gttttaattc acgttgaaga tgaaacatgg gaccaaacac ttgatcattt agctaaacat	1320
gatggagaag atgtacttgg aaataaagtg gaacgaaaag aagatggatt tgaagatgga	1380
gtagaagaca acaaattgaa agagaatatg gaaagagctt gtttgatgtc gttagatatt	1440

acagaacatg aactccaaat tttggaacag cagtctcagg aagaatatct tagtgatatt	1500
gcttataaat ctactgagca tttatctccc aatgataatg aaaacgatac gtcctatgta	1560
attgagagtg atgaagattt agaaatggag atgcttaagc atttatctcc caatgataat	1620
gaaaacgata cgtcctatgt aattgagagt gatgaagatt tagaaatgga gatgcttaag	1680
tctttagaaa acctcaatag tggcacggta gaaccaactc attctaaatg cttaaaaatg	1740
gaaagaaatc tgggtcttcc tactaaagaa gaagaagaag atgatgaaaa tgaagctaata	1800
gaaggggaag aagatgatga taaggacttt ttgtggccag cacccaatga agagcaagtt	1860
acttgcctca agatgtactt tggccattcc agttttaaac cagttcagtg gaaagtgatt	1920
cattcagtat tagaagaaag aagagataat gttgctgtca tggcaactgg atatggaaag	1980
agtttgtgct tccagtatcc acctgtttat gtaggcaaga ttggccttgt tatctctccc	2040
cttatttctc tgatggaaga ccaagtgtca cagcttaaaa tgtccaacat cccagcttgc	2100
ttccttgat cagcacagtc agaaaatgtt ctaacagata ttaaattagg taaataccgg	2160
attgtatacg taactccaga atactgttca ggtaacatgg gcctgctcca gcaacttgag	2220
gctgatattg gtatcacgct cattgctgtg gatgaggctc actgtatttc tgagtggggg	2280
catgatttta gggattcatt caggaagttg ggctccctaa agacagcact gccaatggtt	2340
ccaatcgttg cacttactgc tactgcaagt tcttcaatcc gggaagacat tgtacgttgc	2400
ttaaactctga gaaatcctca gatcacctgt actggttttg atcgaccaa cctgtattta	2460
gaagttaggc gaaaaacagg gaatatcctt caggatctgc agccatttct tgtcaaaaca	2520
agttcccact gggaatttga aggtccaaca atcatctact gtccttctag aaaaatgaca	2580
caacaagtta caggtgaact taggaaactg aatctatcct gtggaacata ccatgcgggc	2640
atgagtttta gcacaaggaa agacattcat cataggtttg taagagatga aattcagtg	2700
gtcatagcta ccatagcttt tggaatgggc attaataaag ctgacattcg ccaagtcatt	2760
cattacggtg ctcctaagga catggaatca tattatcagg agattggtag agctggctcg	2820
gatggacttc aaagttcttg tcacgtcctc tgggctcctg cagacattaa cttaaatagg	2880
caccttctta ctgagatacg taatgagaag tttcgattat acaaattaaa gatgatggca	2940
aagatggaag aatatcttca ttctagcaga tgtaggagac aatcatctt gtctcatttt	3000
gaggacaaac aagtacaaaa agcctccttg ggaattatgg gaactgaaaa atgctgtgat	3060

aattgcaggt ccagattgga tcattgctat tccatggatg actcagagga tacatcctgg	3120
gactttgggtc cacaagcatt taagcttttg tctgctgtgg acatcttagg cgaaaaatth	3180
ggaattgggc ttccaattht atthtctccga ggatctaatt ctacagcgtct tgccgatcaa	3240
tatcgcaggc acagthttatt tggcactggc aaggatcaaa cagagagttg gtggaaggct	3300
thttcccgctc agctgatcac tgagggattc ttggtagaag thttctcggt aatacaatth	3360
atgaagatth gcgcccttac gaaaaagggt agaaattggc thcataaagc taatacagaa	3420
tctcagagcc tcatccttca agctaataaa gaattgtgtc caaagaagtt gcttctgcct	3480
agttcgaaaa ctgtatcttc gggcaccaaa gagcattgtt ataataagc accagttgaa	3540
ttaagtacag agaagaagtc taacttggag aagttatatt cttataaacc atgtgataag	3600
atthtctctg ggagtaacat thctaaaaaa agtatcatgg tacagtcacc agaaaaagct	3660
tacagttcct cacagcctgt tathctggca caagagcagg agactcagat tgtgttatat	3720
ggcaaatgg tagaagctag gcagaaacat gccataaaaa tggatgttcc cccagctatt	3780
ctggcaacaa acaagatact ggtggatatg gccaaaatga gaccaactac ggttgaaaac	3840
gtaaaaagga ttgatgggtg thctgaaggc aaagctgcc agttggtccc tctgttgga	3900
gtcatcaaac atthctgcc aacaaatagt gttcagacag acctcttht aagtacaaaa	3960
cctcaagaag aacagaagac gagtctggt acaaaaaata aaatatgcac actthcacag	4020
tctatggcca tcacatactc thtattccaa gaaaagaaga tgcctttgaa gagcatagct	4080
gagagcagga thctgcctct catgacaatt ggcattgcact tatcccaagc ggtgaaagct	4140
ggctgcccc ttgatttgga gcgagcaggc ctgactccag aggttcagaa gattattgct	4200
gatgttatcc gaaaccctc cgtcaactca gatatgagta aaattagcct aatcagaatg	4260
ttagttcctg aaaacattga cacgtacct atccacatgg caattgagat ctttaaacat	4320
ggctctgaca gcggacttca accttcatgt gatgtcaaca aaaggagatg thttcccggt	4380
tctgaagaga tctgttcaag thctaagaga agcaaggaag aagtaggcat caatactgag	4440
acttcatctg cagagagaaa gagacgatta cctgtgtggt ttgcaaaagg aagtataacc	4500
agcaagaaat taatggacaa aacgaaaagg ggaggtctth ttagttaagc tggcaattac	4560
cagaacaatt atgtthcttg ctgtattata agaggatagc tataththt thctgaagag	4620
taaggagtag taththtgct taaaaatcat tctaattaca aagttcactg thtattgaag	4680

aactggcatc ttaaatcagc cttccgcaat tcatgtagtt tctgggtctt ctgggagcct	4740
acgtgagtac atcacctaac agaataattaa attagacttc ctgtaagatt gctttaagaa	4800
actgttactg tcctgttttc taatctcttt attaaaacag tgtatttgga aaatgttatg	4860
tgctctgatt tgatatagat aacagattag tagttacatg gtaattatgt gatataaaat	4920
attcatatat tatcaaaatt ctgttttgta aatgtaagaa agcatagtta ttttaciaat	4980
tgtttttact gtcttttgaa gaagttctta aatacgttgt taaatggtat tagttgacca	5040
gggcagtga aatgaaaccg cattttgggt gccattaaat agggaaaaaa catgtaaaaa	5100
atgtaaaatg gagaccaatt gcactaggca agtgtatatt ttgtatttta tatacaattt	5160
ctattatttt tcaagtaata aaacaatgtt tttcatactg aatattatat atata	5215

<210> 106
 <211> 1417
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >XPA|ENSG00000136936|ENST00000375128|1417

<400> 106	
ggctgcctc ggcgtgcagt gcgcgtgcgt ggagctggga gctaggtcct cggagtgggc	60
cagagatggc ggcggccgac ggggctttgc cggaggcggc ggcttttagag caaccgcgg	120
agctgcctgc ctcggtgcgg gcgagtatcg agcggaaagc gcagcgggca ctgatgctgc	180
gccaggcccc gctggctgcc cggccctact cggcgacggc ggctgcggct actggaggca	240
tggctaattgt aaaagcagcc ccaaagataa ttgacacagg aggaggcttc attttagaag	300
aggaagaaga agaagaacag aaaattggaa aagttgttca tcaaccagga cctgttatgg	360
aatttgatta tgtaatatgc gaagaatgtg ggaaagaatt tatggattct tatcttatga	420
accactttga ttgccaact tgtgataact gcagagatgc tgatgataaa cacaagctta	480
taaccaaaaac agaggcaaaa caagaatatc ttctgaaaga ctgtgattta gaaaaaagag	540
agccacctct taaatttatt gtgaagaaga atccacatca ttcacaatgg ggtgatatga	600
aactctactt aaagttacag attgtgaaga ggtctcttga agtttgggggt agtcaagaag	660
cattagaaga agcaaaggaa gtccgacagg aaaaccgaga aaaaatgaaa cagaagaat	720
ttgataaaaa agtaaaagaa ttgcggcgag cagtaagaag cagcgtgtgg aaaagggaga	780

cgattgttca tcaacatgag tatggaccag aagaaaacct agaagatgac atgtaccgta	840
agacttgtag tatgtgtggc catgaactga catatgaaaa aatgtgattt tttagttcag	900
tgacctgttt tatagaattt tatattttaaa taaaggaaat ttagattggg ctttttcaaa	960
attcaaaaaa aaaagcaaca tcttcataga tgaatgaaac cttgtataa gtaatacttc	1020
agtaataatt atgtatgtta tggcttaaaa gcaagtttca gtgaagggtca cctggcctgg	1080
ttgtgtgcac aatgtcatgt ctgtgattgc cttcttacia cagagatggg agctgagtgc	1140
tagagtaggt gcagaagtgg taggtcagct acaaatttga ggacaagata ccaaggcaaa	1200
ccctagattg gggtagaggg aaaagggttc aacaaaggct gaactggatt cttaaccaag	1260
aaacaaataa tagcaatggt ggtgcaccac tgtacccag gttctagtca tgtgtttttt	1320
aggacgattt ctgtctccac gatgggtgga acagtgggga actactgctg gaaaaagccc	1380
taatagcaga aataaacatt gagttgtacg agtctga	1417

<210> 107
 <211> 2802
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >XRCC1|ENSG00000073050|ENST00000262887|2802

<400> 107	
tcccttggcc ccaggagaca ggggttgtag aaagccgaga tcgtgccact gcactccatc	60
ctgggtgaga gagcaagacc ctgtctcaac aaaaaatttt taaaaaataa aataaataat	120
aatacagcaa aaagatttgc tttctcggct tcagtgtggg cggttaactcc atcgtgcaat	180
gagaaaggcg aatttcttcc agacaccaat cccggagggtc gcttctgttg ctaggctccc	240
agaaagcagg gttcggacgt cattgggagg cgaggctaga gcgggggttg gtgtggcgga	300
gggaggcggg gctggaggaa acgctcgttg ctaaggaacg cagcgtcttt cccgctctgg	360
agaggcgcga ctgggcttgc gcagtgtcga cgccggcgcc ggcgcgccgg ggtttgaaag	420
gcccagacct cgcgcgcttg cgcactttag ccagcgcagg gcgcaccccg cccctccca	480
ctctccctgc ccctcggacc ccatactcta cctcatcctt ctggccaggc gaagcccacg	540
acgttgacat gccggagatc cgcctccgcc atgtcgtgtc ctgcagcagc caggactcga	600

ctcactgtgc agaaaatctt ctcaaggcag acacttaccg aaaatggcgg gcagccaagg	660
caggcgagaa gaccatctct gtggtcctac agttggagaa ggaggagcag atacacagt	720
tggacattgg gaatgatggc tcagctttcg tggagggtgct ggtgggcagt tcagctggag	780
gcgctgggga gcaagactat gaggtccttc tggtcacctc atctttcatg tcccccttcg	840
agagccgcag tggctcaaac cccaaccgcg ttcgcatgtt tgggcctgac aagctggtcc	900
gggcagccgc cgagaagcgc tgggaccggg tcaaaattgt ttgcagccag ccctacagca	960
aggactcccc ctttggcttg agttttgtac ggtttcatag cccccagac aaagatgagg	1020
cagaggcccc gtcccagaag gtgacagtga ccaagcttgg ccagttccgt gtgaaggagg	1080
aggatgagag cgccaactct ctgaggccgg gggctctctt cttcagccgg atcaacaaga	1140
catccccagt cacagccagc gaccagcag gacctagcta tgcagctgct accctccagg	1200
cttctagtgc tgcctcctca gcctctccag tctccagggc cataggcagc acctccaagc	1260
cccaggagtc tcccaaaggg aagaggaagt tggatttgaa ccaagaagaa aagaagaccc	1320
ccagcaaacc accagcccag ctgtcgccat ctgttcccaa gagacctaaa ttgccagctc	1380
caactcgta cccagccaca gccccagtcc ctgcccagagc acagggggca gtgacaggca	1440
aaccccgagg agaaggcacc gagcccagc gaccccgagc tggcccagag gagctgggga	1500
agatccttca ggggtgtgga gtggtgctga gtggcttcca gaaccccttc cgctccgagc	1560
tgcgagataa ggccctagag cttggggcca agtatcggcc agactggacc cgggacagca	1620
cgcacctcat ctgtgccttt gccaacaccc ccaagtacag ccaggtccta ggcctgggag	1680
gccgcatcgt gcgtaaggag tgggtgctgg actgtcaccg catgcgtcgg cggctgcct	1740
cccagaggta cctcatggca gggccagggt ccagcagtga ggaggatgag gcctctcaca	1800
gcggtggcag cggagatgaa gcccccaagc ttcctcagaa gcaaccccag accaaaacca	1860
agcccactca ggcagctgga cccagctcac cccagaagcc cccaaccct gaagagacca	1920
aagcagcctc accagtgtc caggaagata tagacattga gggggtacag tcagaaggac	1980
aggacaatgg ggcggaagat tctggggaca cagaggatga gctgaggagg gtggcagagc	2040
agaaggaaca cagactgccc cctggccagg aggagaatgg ggaagacccg tatgcaggct	2100
ccacggatga gaacacggac agtgaggaac accaggagcc tcctgatctg ccagtccttg	2160
agctcccaga tttcttccag ggcaagcact tctttcttta cggggagttc cctggggacg	2220

agcggcggaa actcatccga tacgtcacag ctttcaatgg ggagctcgag gacaatatga	2280
gtgaccgggt tcagtttgtg atcacagcac aggaatggga tcccagcttt gaggaggccc	2340
tgatggacaa cccctccctg gcattcgttc gtccccgatg gatctacagt tgcaatgaga	2400
agcagaagtt acttcctcac cagctctatg ggggtgggtgcc gcaagcctga agtatgtgct	2460
atacacacac acacacacac acacacacac acacacacac gatgcattta ataaagatga	2520
gttggtttctc atccaagagt ctcccaaaac tctaagaggc tccctgggac ctggggaaga	2580
atgctgggca cctccgtcag agatctggta gagaaggaa tctttgtctc ttctgcttgg	2640
ccccttatcc ctgtgttggc aagaggcagg gaactgggaa tctgaccctc agcactgccc	2700
ctcaactttt tctggccctc tgagccacac ctgtatcttg gctgtccctt tgtggctgga	2760
ggcctgggta cccatgaggc ttgtctctct cctgaagcct ca	2802

<210> 108
 <211> 3067
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >XRCC2|ENSG00000196584|ENST00000359321|3067

<400> 108	
tttgactggc cgtagagtct gcgcagttgg tgaatggcgt tggtaggggg aaagttgagt	60
ctctcctgcg ccgagccttc ggggcatgt gtagtgcctt ccatagggct gagtctggga	120
ccgagctcct tgcccgaact gaaggtaga gttccttgaa agaaataga ccaaatctgt	180
ttgctgatga agattcacct gtgcatgggt atattcttga atttcatggc ccagaaggaa	240
caggaaaaac agaaatgctt tatcacctaa cagcacgatg tatacttccc aaatcagaag	300
gtggcctgga agtagaagtc ttatttattg atacagatta ccactttgat atgctccggc	360
tagttacaat tcttgagcac agactatccc aaagctctga agaaataatc aaatactgcc	420
tggaagatt ttttttgggt tactgcagta gtagcaccca cttacttctt acactttact	480
cactagaaag tatgttttgt agtcacccat ctctctgcct tttgattttg gatagcctgt	540
cagcttttta ctggatagac cgcgtcaatg gaggagaaag tgtgaactta caggagtcta	600
ctctgaggaa atgttctcag tgcttagaga agcttgtaaa tgactatcgc ctggttcttt	660
ttgcaacgac acaaactata atgcagaaag cctcgagctc atcagaagaa ctttctcatg	720

cctctcgacg actgtgtgat gtggacatag actacagacc ttatctctgt aaggcatggc	780
agcaactggg gaagcacagg atgtttttct ccaaacaaga tgattctcaa agcagcaacc	840
aattttcatt agtttcacgt tgtttaaaaa gtaacagttt aaaaaaacat ttttttatta	900
ttggagaaaag tgggggttgaa ttttgttgat atacatcata aaatagtctt ttgcagggta	960
ctacgcaagc cttaaaatth ttcttaagac agagtcttgc tctgtctccc aggctggagt	1020
gcagtggcac aatcatggct cactgcagcc ttgaactcct ggcctcaagg gatcctccta	1080
tgtgtgcctc ctagagtga gggattacag gcgtgagcca ctgctcgtgg ccaaaagttt	1140
tctttttttt tttttttctt tttgaaacag tcttactctg tctcccaggc tgctggagtg	1200
cagtggcaca atctcggccc gctgcagcct ctgcctcttg ggttcaagtg attctttcac	1260
ctcagcctcc caggtagctg ggattacagg caccaccac cacgcctggc taatttttgt	1320
atttttaata gagacggggt ttcacatgt tggccaggct ggtctcgaac tcctgacctc	1380
aagtgatcca cccacctcgg cctcccaaag tgctaggatt acaggcccgt gccagccct	1440
aaagttttaa actctagggg aattaacagt atttctttac agaattggatt tgttaaacta	1500
gcacagtaaa agtaaagact attctgtttc taggctgttg aatcaaagtg attttagcaa	1560
ttaaactttg tattaattta ccaccaatat ttcttcacaa aggaactttt aaaagattat	1620
ctcagaaaagt aaatctgaga ggtaagaagt aataatgagt aaatggtaag tactttagta	1680
aatctaaaga aatattgata gtaaggcaat cctaagcaaa aagaacaaag ctggaggcat	1740
cacgctaccc agcttcaaac tatactacaa ggctacagta accaaaacag catagtactg	1800
gcacaaaaac acacgtagac tgatggaaca gaatagagaa tttagaaatg agaccacaca	1860
cctataattt ttttgatctt cgatgaacct gacaaaaaca agcaatgggc aatggattct	1920
ctattcaata aatcgtgctg ggataactgg ccagccatat ggaaaagatt gaaaatggac	1980
gccttcctta tgccatatac aaaaattaac tcaagatgga ttaaagactt aatgtaaaac	2040
ccaaaacagt aaaaatcctg gaagacaacc caggcagtag cattcaggac ataggcacag	2100
gcaaagattt catgacgaag acgccaaaaa caattgcaac agaagcaaaa attcacaaat	2160
gggatctaata taaactaaag agctgcacag caaaagaaac tatcaagaga gtaaacagac	2220
agcttacaga atgggagaaa attgttgcaa actatgcatc tgagaaaggt ctgaaatcca	2280
gcatctatac gtaatttaaa caaattdaga agaaaaaacc accccattaa aaagtgggca	2340

aaggacatga acagacactt ttcaaaagaa gacatctgtg gccaacaatc ctatggaaaa	2400
aagcccagca tctactgatca ttagagaaat gcaaatcgaa acaacaacga gataccatct	2460
cacaccagtc caaatggcta ttataaaaat gtcagaaaat aacagatgct ggtgaggttg	2520
tggagaaaaa gatatgctta tacactgttg gtggaaatgt aaattaaatt agttcagcca	2580
ttgtggaaga cagtgtgggg ataaagacag agataccatt caaccagca atctcattac	2640
tgggtatata cccaaaggaa tagaaatcat tggtataaag acacatgcac gcgtatgttc	2700
gttgacgac tgcccatcag tgacagactg gattaaaaaa atgtggtaca tacacaccag	2760
ggaatactat acagccataa aaaggaacaa gactgactgg gcgtggtggc tcatgcctgt	2820
gacccatgca ctttgcgagg ccgagggtggg tggattgccc gcgctcagga ggtcaagacc	2880
agcctgggca acacggtgaa accccatctc tattaaaata caaaaaatta gctgggcatg	2940
gtggtgcgtg cctgtagtgc cagctactca ggaggccgag gcaggagaat tgctggaacc	3000
caggaggtgg aggttgcagt gagctgagat cgcgccattg cactcccgcc tgggcgactc	3060
catctct	3067

<210> 109
 <211> 3018
 <212> DNA
 <213> Homo sapiens

<220>
 <223> >XRCC3|ENSG00000126215|ENST00000553264|3018

<400> 109	
ctccctgagc aggctgaaca tgggcatacg agtgaccagc gtgatcgtgt ccaggccagt	60
gccgcattct gtgtggtcca agggaaaacc ttctccaagt ccacggcccg ccctttcctc	120
tgctgggtta cggccaaggg cgccaggcct cactggagga catggatcct ggacccggcc	180
ctcgaggctg ctgccttcca gtttccccat gtcctggaca ctcttggtct caccatttgg	240
gaggccgaca ggggatgcag agtggcctgt tggaaaggcc ttatgtgacc ccaggctgag	300
tccctgagca gcctgctgag gccggagggg agaagagctg agcttctccc ttgctgggcc	360
aggctcgggc tgtggaatgg ctgccagctc ctttctggg agaggctggg tggggtccac	420
agcaccttcc tcccgtcca caggaagctg tcttctgctg ttggctgctg ggggcgactg	480

atgtttggat ccctgagcac cctgccaatc ccagactcac cttcccactc tttccagagt	540
ccccagggag acacttaagg gaaattaaac tgcagagtgc aagagatgcc tcagtcaagt	600
cagccaaaaa cacgcgggtc atccccaagc cccagagagt gacagagccc cgatgacacg	660
gacacctcgg ctgctgtcac ttccctgggt cgggcctccc acaggctttg aattgaaggc	720
gagtgcctca gaatttgcac ccattgttct gtctttcctg ggaagttatt catcctgggtg	780
gccagcccac cgacaaaatg gatttggatc tactggacct gaatcccaga attattgctg	840
caattaagaa agccaaactg aaatcggtaa aggaggtttt acacttttct ggaccagact	900
tgaagagact gaccaacctc tccagccccg aggtctggca cttgctgaga acggcctcct	960
tacacttgcg gggaagcagc atccttacag cactgcagct gcaccagcag aaggagcggc	1020
tccccacgca gcaccagcgc ctgagcctgg gctgcccggt gctggacgcg ctgctccgcg	1080
gtggcctgcc cctggacggc atcactgagc tggccggacg cagctcggca gggaagaccc	1140
agctggcgct gcagctctgc ctggctgtgc agttcccgcg gcagcacgga ggcctggagg	1200
ctggagccgt ctacatctgc acggaagacg ctttcccgcg caagcgcctg cagcagctca	1260
tggcccagca gccgcggctg cgcactgacg ttccaggaga gctgcttcag aagctccgat	1320
ttggcagcca gatcttcacg gagcacgtgg ccgatgtgga caccttgttg gagtgtgtga	1380
ataagaaggt ccccgactg ctgtctcggg gcatggctcg cctggtggtc atcgactcgg	1440
tggcagcccc attccgctgt gaatttgaca gccaggcctc cgccccagg gccaggcatc	1500
tgagtcctt gggggccacg ctgcgtgagc tgagcagtgc cttccagagc cctgtgtgtg	1560
gcatcaacca ggtgacagag gccatggagg agcagggcgc agcacacggg ccgctggggc	1620
tctgggacga acgtgtttcc ccagcccttg gcataacctg ggctaaccag ctctgtgtga	1680
gactgctggc tgaccggctc cgcgaggaag aggctgccct cggctgcca gcccggaacc	1740
tgcggtgct ctctgcccc cacctgcccc ctcctcctg ttctacacg atcagtgccg	1800
aaggggtgcg agggacacct gggaccagt cccactgaca cgggtggcgc tgcacaacag	1860
ccctgcctga gaagccccga cacacggggc tcgggccttt aaaacgcgtc tgcctgggcc	1920
gtggcacagc tgggagcctg gttcagacac agctcttcca gggcagcggc tccactttct	1980
catccgaaga tgggtggccac agactgacct ccatctgagc tggggggatg ttctgcctct	2040
ccctgggtct ggggacaggc ccgcttgctg ggtacctggt cccactgct gagctggccc	2100

ttggggagag gtgattctca gggctggagc ctgggggtgtc ctacagtac tccctgggag	2160
ccgcctgctt cttctctcca catggaagcc caactgggggt tgcgtctgag gcctgcccc	2220
tgggctgggg cctcagaccc cctcagcctt gggaccgtgc ccacgagggt ctcccctcct	2280
gcacacaggg cagtccttac tccccacca ctcaggccac agtggggctg caggcaggcg	2340
gctcctcctc acccacctct gggtccttgg ctcccggggg cccacctcg gcacacactg	2400
tgccccacaa aacttcagt tgggtacaagg tggagaaagc atatcccacc aacctccagt	2460
gtcagggtcc aggagagcct ggggggtgggg ggactgcctt gtctctagta gtgtggcctg	2520
tgccagcacc acagccggtc agaggagcgc aggcagcgca gggctggcac gtgacaggct	2580
cgtcagccac ctgggaacac agttctgggc aaagaggatc cgaggttgag aggaaggagg	2640
gtcccgggtg atcctggccc tgggggtctg ggcgtccagc tcagccctgg cctggctggg	2700
tgggtattctg gtagggatat ggcaggactc ctggcagggc cacctgcagg accctgtcct	2760
gcagtccac actgtgcaga cccagtcca cactgtggcc aggccttaca tctggctgga	2820
aagcagagcc tcctgggaac acatctggct gcacaggctg aaatatccac ccagcaggca	2880
gagtggcgtg gcctccccat gggcacagtg gtgacccct tgattccac cgtacaaccc	2940
cctccacccc ccactcagt cctccacatg ctgcctggca cagaccaggc ctttgacaaa	3000
taaatgttca atggatgc	3018

<210> 110

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<223> >XRCC4|ENSG00000152422|ENST00000511817|1636

<400> 110

accggaagta gagtcacgga gaggtaggat ccggaagtgg ggctgcctct ttaaataaca	60
aaaatctgag gtattaagaa atggagagaa aaataagcag aatccacctt gtttctgaac	120
ccagtataac tcattttcta caagtatctt gggagaaaac actggaatct ggttttgtta	180
ttacacttac tgatggatcat tcagcatgga ctgggacagt ttctgaatca gagatttccc	240
aagaagctga tgacatggca atggaaaaag ggaaatatgt tggatgaactg agaaaagcat	300
tgttgtcagg agcaggacca gctgatgtat acacgtttaa tttttctaaa gagtcttgtt	360

atctcttctt tgagaaaaac ctgaaagatg tctcattcag acttggttcc ttcaacctag	420
agaaagttga aaaccagct gaagtcatta gagaacttat ttgttattgc ttggacacca	480
ttgcagaaaa tcaagccaaa aatgagcacc tgcagaaaga aaatgaaagg cttctgagag	540
attggaatga tgttcaagga cgatttgaaa aatgtgtgag tgctaaggaa gctttggaga	600
ctgatcttta taagcggttt attctggtgt tgaatgagaa gaaaacaaaa atcagaagtt	660
tgcataataa attattaaat gcagctcaag aacgagaaaa ggacatcaaa caagaagggg	720
aaactgcaat ctgttctgaa atgactgctg accgagatcc agtctatgat gagagtactg	780
atgaggaaag tgaaaaccaa actgatctct ctgggttggc ttcagctgct gtaagtaaag	840
atgattccat tatttcaagt cttgatgtca ctgatattgc accaagtaga aaaaggagac	900
agcgaatgca aagaaatctt gggacagaac ctaaaatggc tcctcaggag aatcagcttc	960
aagaaaagga aaattctagg cctgattctt cactacctga gacgtctaaa aaggagcaca	1020
tctcagctga aaacatgtct ttagaaactc tgagaaacag cagcccagaa gacctcttg	1080
atgagattta acagtctcaa aaaatacttt gatgttcact agactatggt ttctattcat	1140
ttctttaaaa tgaaaaagga gaatttcaag tcagcagccg ctattaccgt atcttacaat	1200
ttaattacat acacagtga ttgaaacat tgtgcaaaat ggattacaca tgtatacaaa	1260
gatacgattt gatgatgaca ctggcacatt attctaaact attcattcag catgcctata	1320
attacataaa ttgtatgaga ctttttgttg caaaggacac atttatcata ttcatcaca	1380
catattatat gtgatagctg tccaacatcc tgtctgggaa gattttgaaa acaggacaaa	1440
gaaaacatca ttttaaaatg tcttcagctt tttttgaata gacgtattca aacatattct	1500
gaacattgat gtttgaacat ttttaattgt gtgatgatgt agaaaatata attttagttt	1560
gtacataaac attgtgaaaa tctgataata aaatTTTTga tacattgaag attttTgtgtt	1620
tttaataaaa tgtgtt	1636

<210> 111

<211> 2709

<212> DNA

<213> Homo sapiens

<220>

<223> >XRCC6|ENSG00000196419|ENST00000359308|2709

<400> 111

gcgggccgtt atccatttgt gttgttcgcc agctaggcct ggcctcgtcc cgcttcgctc	60
ggtcgggtctc gcgcgcccc atagccttgc tagagggtta gcgttagcct taagtgtgcg	120
aatccgagga gcagcgacag actcgagacc acgctccttc ctcggaagg aggcggcacc	180
tcgcgtttga ggcccgctg cgtttgaggc ccgcctgcgc ttgcggcccg cctgcgcttg	240
aggcctgtct gcgtttgaga tctcattggg cgtgattgag gaatttgggg aggtttttgg	300
gcgggtattga ggacgagggg gtccgttagt cagcatagaa tcctggagcg ggaatccctc	360
accgtctaaa tggcgtcggg ggcgggacct ccgggatctg gcttccgcgg gccgccgccc	420
gccctgaaac gtgagggata gctgagatga ggcagctact gggatggccc ccatgcgcat	480
ttacatgcag tccgactgcc gagctttcga ggcagcagga tttaccgtcc acattcctca	540
ctactaacca agcttttaga acagatctca caagaaccta gaggtcggta ttttttcgat	600
ttaaatttgc ctgttactga cgttaacgtc tttcgcctag tgagcagtag ccaacatgtc	660
agggtgggag tcatattaca aaaccgaggg cgatgaagaa gcagaggaag aacaagaaga	720
gaaccttgaa gcaagtggag actataaata ttcaggaaga gatagtttga tttttttggt	780
tgatgcctcc aaggctatgt ttgaatctca gagtgaagat gagttgacac cttttgacat	840
gagcatccag tgtatccaaa gtgtgtacat cagtaagatc ataagcagtg atcgagatct	900
cttggctgtg gtgttctatg gtaccgagaa agacaaaaat tcagtgaatt ttaaaaatat	960
ttacgtctta caggagctgg ataatccagg tgcaaacga attctagagc ttgaccagtt	1020
taaggggcag cagggacaaa aacgtttcca agacatgatg ggccacggat ctgactactc	1080
actcagtga gtgctgtggg tctgtgccaa cctctttagt gatgtccaat tcaagatgag	1140
tcataagagg atcatgctgt tcaccaatga agacaacccc catggcaatg acagtgccaa	1200
agccagccgg gccaggacca aagccggtga tctccgagat acaggcatct tccttgactt	1260
gatgcacctg aagaaacctg ggggctttga catatccttg ttctacagag atatcatcag	1320
catagcagag gatgaggacc tcagggttca ctttgaggaa tccagcaagc tagaagacct	1380
gttgcggaag gttcgcgcca aggagaccag gaagcgagca ctcagcaggt taaagctgaa	1440
gctcaacaaa gatatagtga tctctgtggg catttataat ctggtccaga aggctctcaa	1500
gcctcctcca ataaagctct atcgggaaac aaatgaacca gtgaaaacca agacccggac	1560

ctttaataca agtacaggcg gtttgcttct gcctagcgat accaagaggt ctcagatcta	1620
tgggagtcgt cagattatac tggagaaaga ggaaacagaa gagctaaaac ggtttgatga	1680
tccaggtttg atgctcatgg gtttcaagcc gttgggtactg ctgaagaaac accattacct	1740
gaggccctcc ctgttcgtgt acccagagga gtcgctgggtg attgggagct caaccctgtt	1800
cagtgccttg ctcatcaagt gtctggagaa ggaggttgca gcattgtgca gatacacacc	1860
ccgcaggaac atccctcctt attttgtggc tttggtgcca caggaagaag agttggatga	1920
ccagaaaatt caggtgactc ctccaggctt ccagctgggtc tttttaccct ttgctgatga	1980
taaaaggaag atgcccttta ctgaaaaaat catggcaact ccagagcagg tgggcaagat	2040
gaaggctatc gttgagaagc ttcgcttcac atacagaagt gacagctttg agaaccccg	2100
gctgcagcag cacttcagga acctggaggc cttggccttg gatttgatgg agccggaaca	2160
agcagtggac ctgacattgc ccaaggttga agcaatgaat aaaagactgg gctccttggt	2220
ggatgagttt aaggagcttg tttaccacc agattacaat cctgaaggga aagttaccaa	2280
gagaaaacac gataatgaag gttctggaag caaaaggccc aaggtggagt attcagaaga	2340
ggagctgaag acccacatca gcaagggtac gctgggcaag ttcactgtgc ccatgctgaa	2400
agaggcctgc cgggcttacg ggctgaagag tgggctgaag aagcaggagc tgctggaagc	2460
cctcaccaag cacttccagg actgaccaga ggccgcgcgt ccagctgccc ttccgcagtg	2520
tggccaggct gcctggcctt gtcctcagcc agttaaaatg tgtttctcct gagctaggaa	2580
gagtctaccc gacataagtc gagggacttt atgtttttga ggctttctgt tgccatggtg	2640
atggtgtagc cctcccactt tgctgttcct tactttactg cctgaataaa gagccctaag	2700
tttgtacta	2709

<210> 112
 <211> 2843
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >APC|ENSG00000134982|ENST00000457016|8532

<400> 112
 Met Ala Ala Ala Ser Tyr Asp Gln Leu Leu Lys Gln Val Glu Ala Leu
 1 5 10 15
 Lys Met Glu Asn Ser Asn Leu Arg Gln Glu Leu Glu Asp Asn Ser Asn

450		455		460															
Ala	Met	Asn	Glu	Leu	Gly	Gly	Leu	Gln	Ala	Ile	Ala	Glu	Leu	Leu	Gln				
465					470					475					480				
Val	Asp	Cys	Glu	Met	Tyr	Gly	Leu	Thr	Asn	Asp	His	Tyr	Ser	Ile	Thr				
				485					490						495				
Leu	Arg	Arg	Tyr	Ala	Gly	Met	Ala	Leu	Thr	Asn	Leu	Thr	Phe	Gly	Asp				
			500					505						510					
Val	Ala	Asn	Lys	Ala	Thr	Leu	Cys	Ser	Met	Lys	Gly	Cys	Met	Arg	Ala				
		515					520					525							
Leu	Val	Ala	Gln	Leu	Lys	Ser	Glu	Ser	Glu	Asp	Leu	Gln	Gln	Val	Ile				
	530					535						540							
Ala	Ser	Val	Leu	Arg	Asn	Leu	Ser	Trp	Arg	Ala	Asp	Val	Asn	Ser	Lys				
545					550					555					560				
Lys	Thr	Leu	Arg	Glu	Val	Gly	Ser	Val	Lys	Ala	Leu	Met	Glu	Cys	Ala				
				565					570					575					
Leu	Glu	Val	Lys	Lys	Glu	Ser	Thr	Leu	Lys	Ser	Val	Leu	Ser	Ala	Leu				
			580					585					590						
Trp	Asn	Leu	Ser	Ala	His	Cys	Thr	Glu	Asn	Lys	Ala	Asp	Ile	Cys	Ala				
		595					600						605						
Val	Asp	Gly	Ala	Leu	Ala	Phe	Leu	Val	Gly	Thr	Leu	Thr	Tyr	Arg	Ser				
	610					615						620							
Gln	Thr	Asn	Thr	Leu	Ala	Ile	Ile	Glu	Ser	Gly	Gly	Gly	Ile	Leu	Arg				
625					630					635					640				
Asn	Val	Ser	Ser	Leu	Ile	Ala	Thr	Asn	Glu	Asp	His	Arg	Gln	Ile	Leu				
				645					650					655					
Arg	Glu	Asn	Asn	Cys	Leu	Gln	Thr	Leu	Leu	Gln	His	Leu	Lys	Ser	His				
			660					665					670						
Ser	Leu	Thr	Ile	Val	Ser	Asn	Ala	Cys	Gly	Thr	Leu	Trp	Asn	Leu	Ser				
		675					680					685							
Ala	Arg	Asn	Pro	Lys	Asp	Gln	Glu	Ala	Leu	Trp	Asp	Met	Gly	Ala	Val				
	690					695					700								
Ser	Met	Leu	Lys	Asn	Leu	Ile	His	Ser	Lys	His	Lys	Met	Ile	Ala	Met				
705					710					715					720				
Gly	Ser	Ala	Ala	Ala	Leu	Arg	Asn	Leu	Met	Ala	Asn	Arg	Pro	Ala	Lys				
				725					730					735					
Tyr	Lys	Asp	Ala	Asn	Ile	Met	Ser	Pro	Gly	Ser	Ser	Leu	Pro	Ser	Leu				
		740						745					750						
His	Val	Arg	Lys	Gln	Lys	Ala	Leu	Glu	Ala	Glu	Leu	Asp	Ala	Gln	His				
		755				760						765							
Leu	Ser	Glu	Thr	Phe	Asp	Asn	Ile	Asp	Asn	Leu	Ser	Pro	Lys	Ala	Ser				
	770					775					780								
His	Arg	Ser	Lys	Gln	Arg	His	Lys	Gln	Ser	Leu	Tyr	Gly	Asp	Tyr	Val				
785					790					795					800				
Phe	Asp	Thr	Asn	Arg	His	Asp	Asp	Asn	Arg	Ser	Asp	Asn	Phe	Asn	Thr				
			805						810					815					
Gly	Asn	Met	Thr	Val	Leu	Ser	Pro	Tyr	Leu	Asn	Thr	Thr	Val	Leu	Pro				
			820					825					830						
Ser	Ser	Ser	Ser	Ser	Arg	Gly	Ser	Leu	Asp	Ser	Ser	Arg	Ser	Glu	Lys				
		835					840					845							
Asp	Arg	Ser	Leu	Glu	Arg	Glu	Arg	Gly	Ile	Gly	Leu	Gly	Asn	Tyr	His				
	850					855					860								
Pro	Ala	Thr	Glu	Asn	Pro	Gly	Thr	Ser	Ser	Lys	Arg	Gly	Leu	Gln	Ile				
865					870					875					880				
Ser	Thr	Thr	Ala	Ala	Gln	Ile	Ala	Lys	Val	Met	Glu	Glu	Val	Ser	Ala				

Ile	His	Thr	Ser	885	Gln	Glu	Asp	Arg	Ser	890	Ser	Gly	Ser	Thr	Thr	895	Glu	Leu
			900						905							910		
His	Cys	Val	Thr	Asp	Glu	Arg	Asn	Ala	Leu	Arg	Arg	Ser	Ser	Ala	Ala			
		915					920							925				
His	Thr	His	Ser	Asn	Thr	Tyr	Asn	Phe	Thr	Lys	Ser	Glu	Asn	Ser	Asn			
		930				935						940						
Arg	Thr	Cys	Ser	Met	Pro	Tyr	Ala	Lys	Leu	Glu	Tyr	Lys	Arg	Ser	Ser			
945				950						955						960		
Asn	Asp	Ser	Leu	Asn	Ser	Val	Ser	Ser	Ser	Asp	Gly	Tyr	Gly	Lys	Arg			
			965						970					975				
Gly	Gln	Met	Lys	Pro	Ser	Ile	Glu	Ser	Tyr	Ser	Glu	Asp	Asp	Glu	Ser			
		980						985					990					
Lys	Phe	Cys	Ser	Tyr	Gly	Gln	Tyr	Pro	Ala	Asp	Leu	Ala	His	Lys	Ile			
		995					1000					1005						
His	Ser	Ala	Asn	His	Met	Asp	Asp	Asn	Asp	Gly	Glu	Leu	Asp	Thr	Pro			
1010					1015						1020							
Ile	Asn	Tyr	Ser	Leu	Lys	Tyr	Ser	Asp	Glu	Gln	Leu	Asn	Ser	Gly	Arg			
1025				1030						1035					1040			
Gln	Ser	Pro	Ser	Gln	Asn	Glu	Arg	Trp	Ala	Arg	Pro	Lys	His	Ile	Ile			
			1045						1050					1055				
Glu	Asp	Glu	Ile	Lys	Gln	Ser	Glu	Gln	Arg	Gln	Ser	Arg	Asn	Gln	Ser			
		1060					1065						1070					
Thr	Thr	Tyr	Pro	Val	Tyr	Thr	Glu	Ser	Thr	Asp	Asp	Lys	His	Leu	Lys			
		1075					1080					1085						
Phe	Gln	Pro	His	Phe	Gly	Gln	Gln	Glu	Cys	Val	Ser	Pro	Tyr	Arg	Ser			
1090					1095						1100							
Arg	Gly	Ala	Asn	Gly	Ser	Glu	Thr	Asn	Arg	Val	Gly	Ser	Asn	His	Gly			
1105				1110					1115					1120				
Ile	Asn	Gln	Asn	Val	Ser	Gln	Ser	Leu	Cys	Gln	Glu	Asp	Asp	Tyr	Glu			
			1125					1130						1135				
Asp	Asp	Lys	Pro	Thr	Asn	Tyr	Ser	Glu	Arg	Tyr	Ser	Glu	Glu	Glu	Gln			
		1140					1145						1150					
His	Glu	Glu	Glu	Glu	Arg	Pro	Thr	Asn	Tyr	Ser	Ile	Lys	Tyr	Asn	Glu			
		1155					1160					1165						
Glu	Lys	Arg	His	Val	Asp	Gln	Pro	Ile	Asp	Tyr	Ser	Leu	Lys	Tyr	Ala			
1170					1175						1180							
Thr	Asp	Ile	Pro	Ser	Ser	Gln	Lys	Gln	Ser	Phe	Ser	Phe	Ser	Lys	Ser			
1185				1190						1195					1200			
Ser	Ser	Gly	Gln	Ser	Ser	Lys	Thr	Glu	His	Met	Ser	Ser	Ser	Ser	Glu			
			1205					1210						1215				
Asn	Thr	Ser	Thr	Pro	Ser	Ser	Asn	Ala	Lys	Arg	Gln	Asn	Gln	Leu	His			
		1220					1225						1230					
Pro	Ser	Ser	Ala	Gln	Ser	Arg	Ser	Gly	Gln	Pro	Gln	Lys	Ala	Ala	Thr			
		1235					1240					1245						
Cys	Lys	Val	Ser	Ser	Ile	Asn	Gln	Glu	Thr	Ile	Gln	Thr	Tyr	Cys	Val			
1250					1255						1260							
Glu	Asp	Thr	Pro	Ile	Cys	Phe	Ser	Arg	Cys	Ser	Ser	Leu	Ser	Ser	Leu			
1265				1270					1275						1280			
Ser	Ser	Ala	Glu	Asp	Glu	Ile	Gly	Cys	Asn	Gln	Thr	Thr	Gln	Glu	Ala			
			1285					1290						1295				
Asp	Ser	Ala	Asn	Thr	Leu	Gln	Ile	Ala	Glu	Ile	Lys	Glu	Lys	Ile	Gly			
		1300						1305					1310					
Thr	Arg	Ser	Ala	Glu	Asp	Pro	Val	Ser	Glu	Val	Pro	Ala	Val	Ser	Gln			

Page 349

1745		1750		1755		1760
Asn Lys Asn Gln Leu Asp Gly Lys Lys Lys Lys Pro Thr Ser Pro Val						
	1765		1770		1775	
Lys Pro Ile Pro Gln Asn Thr Glu Tyr Arg Thr Arg Val Arg Lys Asn						
	1780		1785		1790	
Ala Asp Ser Lys Asn Asn Leu Asn Ala Glu Arg Val Phe Ser Asp Asn						
	1795		1800		1805	
Lys Asp Ser Lys Lys Gln Asn Leu Lys Asn Asn Ser Lys Val Phe Asn						
	1810		1815		1820	
Asp Lys Leu Pro Asn Asn Glu Asp Arg Val Arg Gly Ser Phe Ala Phe						
	1825		1830		1835	1840
Asp Ser Pro His His Tyr Thr Pro Ile Glu Gly Thr Pro Tyr Cys Phe						
	1845		1850		1855	
Ser Arg Asn Asp Ser Leu Ser Ser Leu Asp Phe Asp Asp Asp Asp Val						
	1860		1865		1870	
Asp Leu Ser Arg Glu Lys Ala Glu Leu Arg Lys Ala Lys Glu Asn Lys						
	1875		1880		1885	
Glu Ser Glu Ala Lys Val Thr Ser His Thr Glu Leu Thr Ser Asn Gln						
	1890		1895		1900	
Gln Ser Ala Asn Lys Thr Gln Ala Ile Ala Lys Gln Pro Ile Asn Arg						
	1905		1910		1915	1920
Gly Gln Pro Lys Pro Ile Leu Gln Lys Gln Ser Thr Phe Pro Gln Ser						
	1925		1930		1935	
Ser Lys Asp Ile Pro Asp Arg Gly Ala Ala Thr Asp Glu Lys Leu Gln						
	1940		1945		1950	
Asn Phe Ala Ile Glu Asn Thr Pro Val Cys Phe Ser His Asn Ser Ser						
	1955		1960		1965	
Leu Ser Ser Leu Ser Asp Ile Asp Gln Glu Asn Asn Asn Lys Glu Asn						
	1970		1975		1980	
Glu Pro Ile Lys Glu Thr Glu Pro Pro Asp Ser Gln Gly Glu Pro Ser						
	1985		1990		1995	2000
Lys Pro Gln Ala Ser Gly Tyr Ala Pro Lys Ser Phe His Val Glu Asp						
	2005		2010		2015	
Thr Pro Val Cys Phe Ser Arg Asn Ser Ser Leu Ser Ser Leu Ser Ile						
	2020		2025		2030	
Asp Ser Glu Asp Asp Leu Leu Gln Glu Cys Ile Ser Ser Ala Met Pro						
	2035		2040		2045	
Lys Lys Lys Lys Pro Ser Arg Leu Lys Gly Asp Asn Glu Lys His Ser						
	2050		2055		2060	
Pro Arg Asn Met Gly Gly Ile Leu Gly Glu Asp Leu Thr Leu Asp Leu						
	2065		2070		2075	2080
Lys Asp Ile Gln Arg Pro Asp Ser Glu His Gly Leu Ser Pro Asp Ser						
	2085		2090		2095	
Glu Asn Phe Asp Trp Lys Ala Ile Gln Glu Gly Ala Asn Ser Ile Val						
	2100		2105		2110	
Ser Ser Leu His Gln Ala Ala Ala Ala Cys Leu Ser Arg Gln Ala						
	2115		2120		2125	
Ser Ser Asp Ser Asp Ser Ile Leu Ser Leu Lys Ser Gly Ile Ser Leu						
	2130		2135		2140	
Gly Ser Pro Phe His Leu Thr Pro Asp Gln Glu Glu Lys Pro Phe Thr						
	2145		2150		2155	2160
Ser Asn Lys Gly Pro Arg Ile Leu Lys Pro Gly Glu Lys Ser Thr Leu						
	2165		2170		2175	
Glu Thr Lys Lys Ile Glu Ser Glu Ser Lys Gly Ile Lys Gly Gly Lys						

2610	2615	2620
Ser Thr Ser Gln Thr Val Ser Ser Gly Ala Thr Asn Gly Ala Glu Ser		
2625	2630	2635
Lys Thr Leu Ile Tyr Gln Met Ala Pro Ala Val Ser Lys Thr Glu Asp		2640
	2645	2650
Val Trp Val Arg Ile Glu Asp Cys Pro Ile Asn Asn Pro Arg Ser Gly		2655
	2660	2665
Arg Ser Pro Thr Gly Asn Thr Pro Pro Val Ile Asp Ser Val Ser Glu		2670
	2675	2680
Lys Ala Asn Pro Asn Ile Lys Asp Ser Lys Asp Asn Gln Ala Lys Gln		2685
	2690	2695
Asn Val Gly Asn Gly Ser Val Pro Met Arg Thr Val Gly Leu Glu Asn		2700
2705	2710	2715
Arg Leu Asn Ser Phe Ile Gln Val Asp Ala Pro Asp Gln Lys Gly Thr		2720
	2725	2730
Glu Ile Lys Pro Gly Gln Asn Asn Pro Val Pro Val Ser Glu Thr Asn		2735
	2740	2745
Glu Ser Ser Ile Val Glu Arg Thr Pro Phe Ser Ser Ser Ser Ser Ser		2750
	2755	2760
Lys His Ser Ser Pro Ser Gly Thr Val Ala Ala Arg Val Thr Pro Phe		2765
	2770	2775
Asn Tyr Asn Pro Ser Pro Arg Lys Ser Ser Ala Asp Ser Thr Ser Ala		2780
2785	2790	2795
Arg Pro Ser Gln Ile Pro Thr Pro Val Asn Asn Asn Thr Lys Lys Arg		2800
	2805	2810
Asp Ser Lys Thr Asp Ser Thr Glu Ser Ser Gly Thr Gln Ser Pro Lys		2815
	2820	2825
Arg His Ser Gly Ser Tyr Leu Val Thr Ser Val		2830
	2835	2840

<210> 113
 <211> 2285
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >ARID1A|ENSG00000117713|ENST00000324856|6858

<400> 113
Met Ala Ala Gln Val Ala Pro Ala Ala Ser Ser Leu Gly Asn Pro
1 5 10 15
Pro Pro Pro Pro Pro Ser Glu Leu Lys Lys Ala Glu Gln Gln Arg
20 25 30
Glu Glu Ala Gly Gly Glu Ala Ala Ala Ala Ala Ala Glu Arg Gly
35 40 45
Glu Met Lys Ala Ala Ala Gly Gln Glu Ser Glu Gly Pro Ala Val Gly
50 55 60
Pro Pro Gln Pro Leu Gly Lys Glu Leu Gln Asp Gly Ala Glu Ser Asn
65 70 75 80
Gly Gly Gly Gly Gly Gly Gly Ala Gly Ser Gly Gly Gly Pro Gly Ala
85 90 95
Glu Pro Asp Leu Lys Asn Ser Asn Gly Asn Ala Gly Pro Arg Pro Ala
100 105 110

Leu	Asn	Asn	Asn	Leu	Thr	Glu	Pro	Pro	Gly	Gly	Gly	Gly	Gly	Gly	Ser
		115					120					125			
Ser	Asp	Gly	Val	Gly	Ala	Pro	Pro	His	Ser	Ala	Ala	Ala	Ala	Leu	Pro
	130					135					140				
Pro	Pro	Ala	Tyr	Gly	Phe	Gly	Gln	Pro	Tyr	Gly	Arg	Ser	Pro	Ser	Ala
145					150					155					160
Val	Ala	Ala	Ala	Ala	Ala	Ala	Val	Phe	His	Gln	Gln	His	Gly	Gly	Gln
				165					170					175	
Gln	Ser	Pro	Gly	Leu	Ala	Ala	Leu	Gln	Ser	Gly	Gly	Gly	Gly	Gly	Leu
			180					185					190		
Glu	Pro	Tyr	Ala	Gly	Pro	Gln	Gln	Asn	Ser	His	Asp	His	Gly	Phe	Pro
		195					200				205				
Asn	His	Gln	Tyr	Asn	Ser	Tyr	Tyr	Pro	Asn	Arg	Ser	Ala	Tyr	Pro	Pro
	210					215					220				
Pro	Ala	Pro	Ala	Tyr	Ala	Leu	Ser	Ser	Pro	Arg	Gly	Gly	Thr	Pro	Gly
225					230					235					240
Ser	Gly	Ala	Ala	Ala	Ala	Ala	Gly	Ser	Lys	Pro	Pro	Pro	Ser	Ser	Ser
			245						250					255	
Ala	Ser	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Phe	Ala	Gln	Gln	Arg	Phe	Gly
			260					265					270		
Ala	Met	Gly	Gly	Gly	Gly	Pro	Ser	Ala	Ala	Gly	Gly	Gly	Thr	Pro	Gln
		275					280					285			
Pro	Thr	Ala	Thr	Pro	Thr	Leu	Asn	Gln	Leu	Leu	Thr	Ser	Pro	Ser	Ser
	290					295					300				
Ala	Arg	Gly	Tyr	Gln	Gly	Tyr	Pro	Gly	Gly	Asp	Tyr	Ser	Gly	Gly	Pro
305					310					315					320
Gln	Asp	Gly	Gly	Ala	Gly	Lys	Gly	Pro	Ala	Asp	Met	Ala	Ser	Gln	Cys
			325						330					335	
Trp	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Gly
		340					345						350		
Ala	Gln	Gln	Arg	Ser	His	His	Ala	Pro	Met	Ser	Pro	Gly	Ser	Ser	Gly
	355						360					365			
Gly	Gly	Gly	Gln	Pro	Leu	Ala	Arg	Thr	Pro	Gln	Pro	Ser	Ser	Pro	Met
	370					375					380				
Asp	Gln	Met	Gly	Lys	Met	Arg	Pro	Gln	Pro	Tyr	Gly	Gly	Thr	Asn	Pro
385				390						395					400
Tyr	Ser	Gln	Gln	Gln	Gly	Pro	Pro	Ser	Gly	Pro	Gln	Gln	Gly	His	Gly
			405						410					415	
Tyr	Pro	Gly	Gln	Pro	Tyr	Gly	Ser	Gln	Thr	Pro	Gln	Arg	Tyr	Pro	Met
			420					425				430			
Thr	Met	Gln	Gly	Arg	Ala	Gln	Ser	Ala	Met	Gly	Gly	Leu	Ser	Tyr	Thr
	435						440					445			
Gln	Gln	Ile	Pro	Pro	Tyr	Gly	Gln	Gln	Gly	Pro	Ser	Gly	Tyr	Gly	Gln
	450					455					460				
Gln	Gly	Gln	Thr	Pro	Tyr										

Pro	Gln	Ser	Gln	Pro	Pro	Tyr	Ser	Gln	Pro	Gln	Ala	Gln	Ser	Pro	Tyr
545					550					555					560
Gln	Gln	Gln	Gln	Pro	Gln	Gln	Pro	Ala	Pro	Ser	Thr	Leu	Ser	Gln	Gln
				565					570						575
Ala	Ala	Tyr	Pro	Gln	Pro	Gln	Ser	Gln	Gln	Ser	Gln	Gln	Thr	Ala	Tyr
			580					585						590	
Ser	Gln	Gln	Arg	Phe	Pro	Pro	Pro	Gln	Glu	Leu	Ser	Gln	Asp	Ser	Phe
		595					600					605			
Gly	Ser	Gln	Ala	Ser	Ser	Ala	Pro	Ser	Met	Thr	Ser	Ser	Lys	Gly	Gly
610						615					620				
Gln	Glu	Asp	Met	Asn	Leu	Ser	Leu	Gln	Ser	Arg	Pro	Ser	Ser	Leu	Pro
625					630					635					640
Asp	Leu	Ser	Gly	Ser	Ile	Asp	Asp	Leu	Pro	Met	Gly	Thr	Glu	Gly	Ala
				645					650					655	
Leu	Ser	Pro	Gly	Val	Ser	Thr	Ser	Gly	Ile	Ser	Ser	Ser	Gln	Gly	Glu
			660					665					670		
Gln	Ser	Asn	Pro	Ala	Gln	Ser	Pro	Phe	Ser	Pro	His	Thr	Ser	Pro	His
		675					680					685			
Leu	Pro	Gly	Ile	Arg	Gly	Pro	Ser	Pro	Ser	Pro	Val	Gly	Ser	Pro	Ala
690						695					700				
Ser	Val	Ala	Gln	Ser	Arg	Ser	Gly	Pro	Leu	Ser	Pro	Ala	Ala	Val	Pro
705					710					715					720
Gly	Asn	Gln	Met	Pro	Pro	Arg	Pro	Pro	Ser	Gly	Gln	Ser	Asp	Ser	Ile
				725					730					735	
Met	His	Pro	Ser	Met	Asn	Gln	Ser	Ser	Ile	Ala	Gln	Asp	Arg	Gly	Tyr
			740					745					750		
Met	Gln	Arg	Asn	Pro	Gln	Met	Pro	Gln	Tyr	Ser	Ser	Pro	Gln	Pro	Gly
		755					760					765			
Ser	Ala	Leu	Ser	Pro	Arg	Gln	Pro	Ser	Gly	Gly	Gln	Ile	His	Thr	Gly
770						775					780				
Met	Gly	Ser	Tyr	Gln	Gln	Asn	Ser	Met	Gly	Ser	Tyr	Gly	Pro	Gln	Gly
785					790					795					800
Gly	Gln	Tyr	Gly	Pro	Gln	Gly	Gly	Tyr	Pro	Arg	Gln	Pro	Asn	Tyr	Asn
				805					810					815	
Ala	Leu	Pro	Asn	Ala	Asn	Tyr	Pro	Ser	Ala	Gly	Met	Ala	Gly	Gly	Ile
			820					825					830		
Asn	Pro	Met	Gly	Ala	Gly	Gly	Gln	Met	His	Gly	Gln	Pro	Gly	Ile	Pro
		835					840					845			
Pro	Tyr	Gly	Thr	Leu	Pro	Pro	Gly	Arg	Met	Ser	His	Ala	Ser	Met	Gly
					855					860					
Asn	Arg	Pro	Tyr	Gly	Pro	Asn	Met	Ala	Asn	Met	Pro	Pro	Gln	Val	Gly
865					870					875					880
Ser	Gly	Met	Cys	Pro	Pro	Pro	Gly	Gly	Met	Asn	Arg	Lys	Thr	Gln	Glu
				885					890					895	
Thr	Ala	Val	Ala	Met	His	Val	Ala	Ala	Asn	Ser	Ile	Gln	Asn	Arg	Pro
			900					905					910		
Pro	Gly	Tyr	Pro	Asn	Met	Asn	Gln	Gly	Gly	Met	Met	Gly	Thr	Gly	Pro
		915					920					925			
Pro	Tyr	Gly	Gln	Gly	Ile	Asn	Ser	Met	Ala	Gly	Met	Ile	Asn	Pro	Gln
					935						940				
Gly	Pro	Pro	Tyr	Ser	Met	Gly	Gly	Thr	Met	Ala	Asn	Asn	Ser	Ala	Gly
945					950					955					960
Met	Ala	Ala	Ser	Pro	Glu	Met	Met	Gly	Leu	Gly	Asp	Val	Lys	Leu	Thr
				965					970					975	

Pro Ala Thr Lys Met Asn Asn Lys Ala Asp Gly Thr Pro Lys Thr Glu
 980 985 990
 Ser Lys Ser Lys Lys Ser Ser Ser Ser Thr Thr Thr Asn Glu Lys Ile
 995 1000 1005
 Thr Lys Leu Tyr Glu Leu Gly Gly Glu Pro Glu Arg Lys Met Trp Val
 1010 1015 1020
 Asp Arg Tyr Leu Ala Phe Thr Glu Glu Lys Ala Met Gly Met Thr Asn
 1025 1030 1035 1040
 Leu Pro Ala Val Gly Arg Lys Pro Leu Asp Leu Tyr Arg Leu Tyr Val
 1045 1050 1055
 Ser Val Lys Glu Ile Gly Gly Leu Thr Gln Val Asn Lys Asn Lys Lys
 1060 1065 1070
 Trp Arg Glu Leu Ala Thr Asn Leu Asn Val Gly Thr Ser Ser Ser Ala
 1075 1080 1085
 Ala Ser Ser Leu Lys Lys Gln Tyr Ile Gln Cys Leu Tyr Ala Phe Glu
 1090 1095 1100
 Cys Lys Ile Glu Arg Gly Glu Asp Pro Pro Pro Asp Ile Phe Ala Ala
 1105 1110 1115 1120
 Ala Asp Ser Lys Lys Ser Gln Pro Lys Ile Gln Pro Pro Ser Pro Ala
 1125 1130 1135
 Gly Ser Gly Ser Met Gln Gly Pro Gln Thr Pro Gln Ser Thr Ser Ser
 1140 1145 1150
 Ser Met Ala Glu Gly Gly Asp Leu Lys Pro Pro Thr Pro Ala Ser Thr
 1155 1160 1165
 Pro His Ser Gln Ile Pro Pro Leu Pro Gly Met Ser Arg Ser Asn Ser
 1170 1175 1180
 Val Gly Ile Gln Asp Ala Phe Asn Asp Gly Ser Asp Ser Thr Phe Gln
 1185 1190 1195 1200
 Lys Arg Asn Ser Met Thr Pro Asn Pro Gly Tyr Gln Pro Ser Met Asn
 1205 1210 1215
 Thr Ser Asp Met Met Gly Arg Met Ser Tyr Glu Pro Asn Lys Asp Pro
 1220 1225 1230
 Tyr Gly Ser Met Arg Lys Ala Pro Gly Ser Asp Pro Phe Met Ser Ser
 1235 1240 1245
 Gly Gln Gly Pro Asn Gly Gly Met Gly Asp Pro Tyr Ser Arg Ala Ala
 1250 1255 1260
 Gly Pro Gly Leu Gly Asn Val Ala Met Gly Pro Arg Gln His Tyr Pro
 1265 1270 1275 1280
 Tyr Gly Gly Pro Tyr Asp Arg Val Arg Thr Glu Pro Gly Ile Gly Pro
 1285 1290 1295
 Glu Gly Asn Met Ser Thr Gly Ala Pro Gln Pro Asn Leu Met Pro Ser
 1300 1305 1310
 Asn Pro Asp Ser Gly Met Tyr Ser Pro Ser Arg Tyr Pro Pro Gln Gln
 1315 1320 1325
 Gln Gln Gln Gln Gln Gln Arg His Asp Ser Tyr Gly Asn Gln Phe Ser
 1330 1335 1340
 Thr Gln Gly Thr Pro Ser Gly Ser Pro Phe Pro Ser Gln Gln Thr Thr
 1345 1350 1355 1360
 Met Tyr Gln Gln Gln Gln Gln Asn Tyr Lys Arg Pro Met Asp Gly Thr
 1365 1370 1375
 Tyr Gly Pro Pro Ala Lys Arg His Glu Gly Glu Met Tyr Ser Val Pro
 1380 1385 1390
 Tyr Ser Thr Gly Gln Gly Gln Pro Gln Gln Gln Gln Leu Pro Pro Ala
 1395 1400 1405

Gln	Pro	Gln	Pro	Ala	Ser	Gln	Gln	Gln	Ala	Ala	Gln	Pro	Ser	Pro	Gln	
1410						1415					1420					
Gln	Asp	Val	Tyr	Asn	Gln	Tyr	Gly	Asn	Ala	Tyr	Pro	Ala	Thr	Ala	Thr	
1425					1430					1435					1440	
Ala	Ala	Thr	Glu	Arg	Arg	Pro	Ala	Gly	Gly	Pro	Gln	Asn	Gln	Phe	Pro	
				1445					1450					1455		
Phe	Gln	Phe	Gly	Arg	Asp	Arg	Val	Ser	Ala	Pro	Pro	Gly	Thr	Asn	Ala	
			1460					1465					1470			
Gln	Gln	Asn	Met	Pro	Pro	Gln	Met	Met	Gly	Gly	Pro	Ile	Gln	Ala	Ser	
		1475					1480					1485				
Ala	Glu	Val	Ala	Gln	Gln	Gly	Thr	Met	Trp	Gln	Gly	Arg	Asn	Asp	Met	
1490						1495					1500					
Thr	Tyr	Asn	Tyr	Ala	Asn	Arg	Gln	Ser	Thr	Gly	Ser	Ala	Pro	Gln	Gly	
1505					1510					1515					1520	
Pro	Ala	Tyr	His	Gly	Val	Asn	Arg	Thr	Asp	Glu	Met	Leu	His	Thr	Asp	
				1525					1530					1535		
Gln	Arg	Ala	Asn	His	Glu	Gly	Ser	Trp	Pro	Ser	His	Gly	Thr	Arg	Gln	
			1540					1545					1550			
Pro	Pro	Tyr	Gly	Pro	Ser	Ala	Pro	Val	Pro	Pro	Met	Thr	Arg	Pro	Pro	
		1555					1560					1565				
Pro	Ser	Asn	Tyr	Gln	Pro	Pro	Pro	Ser	Met	Gln	Asn	His	Ile	Pro	Gln	
		1570				1575				1580						
Val	Ser	Ser	Pro	Ala	Pro	Leu	Pro	Arg	Pro	Met	Glu	Asn	Arg	Thr	Ser	
1585					1590					1595					1600	
Pro	Ser	Lys	Ser	Pro	Phe	Leu	His	Ser	Gly	Met	Lys	Met	Gln	Lys	Ala	
				1605					1610					1615		
Gly	Pro	Pro	Val	Pro	Ala	Ser	His	Ile	Ala	Pro	Ala	Pro	Val	Gln	Pro	
			1620					1625					1630			
Pro	Met	Ile	Arg	Arg	Asp	Ile	Thr	Phe	Pro	Pro	Gly	Ser	Val	Glu	Ala	
		1635					1640					1645				
Thr	Gln	Pro	Val	Leu	Lys	Gln	Arg	Arg	Arg	Leu	Thr	Met	Lys	Asp	Ile	
1650						1655					1660					
Gly	Thr	Pro	Glu	Ala	Trp	Arg	Val	Met	Met	Ser	Leu	Lys	Ser	Gly	Leu	
1665					1670					1675					1680	
Leu	Ala	Glu	Ser	Thr	Trp	Ala	Leu	Asp	Thr	Ile	Asn	Ile	Leu	Leu	Tyr	
				1685					1690					1695		
Asp	Asp	Asn	Ser	Ile	Met	Thr	Phe	Asn	Leu	Ser	Gln	Leu	Pro	Gly	Leu	
			1700					1705					1710			
Leu	Glu	Leu	Leu	Val	Glu	Tyr	Phe	Arg	Arg	Cys	Leu	Ile	Glu	Ile	Phe	
		1715					1720					1725				
Gly	Ile	Leu	Lys	Glu	Tyr	Glu	Val	Gly	Asp	Pro	Gly	Gln	Arg	Thr	Leu	
		1730				1735					1740					
Leu	Asp	Pro	Gly	Arg	Phe	Ser	Lys	Val	Ser	Ser	Pro	Ala	Pro	Met	Glu	
1745					1750					1755					1760	
Gly	Gly	Glu	Glu	Glu	Glu	Glu	Leu	Leu	Gly	Pro	Lys	Leu	Glu	Glu	Glu	
				1765					1770					1775		
Glu	Glu	Glu	Glu	Val	Val	Glu	Asn	Asp	Glu	Glu	Ile	Ala	Phe	Ser	Gly	
				1780				1785					1790			
Lys	Asp	Lys	Pro	Ala	Ser	Glu	Asn	Ser	Glu	Glu	Lys	Leu	Ile	Ser	Lys	
		1795					1800					1805				
Phe	Asp	Lys	Leu	Pro	Val	Lys	Ile	Val	Gln	Lys	Asn	Asp	Pro	Phe	Val	
		1810				1815					1820					
Val	Asp	Cys	Ser	Asp	Lys	Leu	Gly	Arg	Val	Gln	Glu	Phe	Asp	Ser	Gly	
1825					1830					1835					1840	

Leu Leu His Trp Arg Ile Gly Gly Gly Asp Thr Thr Glu His Ile Gln
 1845 1850 1855
 Thr His Phe Glu Ser Lys Thr Glu Leu Leu Pro Ser Arg Pro His Ala
 1860 1865 1870
 Pro Cys Pro Pro Ala Pro Arg Lys His Val Thr Thr Ala Glu Gly Thr
 1875 1880 1885
 Pro Gly Thr Thr Asp Gln Glu Gly Pro Pro Pro Asp Gly Pro Pro Glu
 1890 1895 1900
 Lys Arg Ile Thr Ala Thr Met Asp Asp Met Leu Ser Thr Arg Ser Ser
 1905 1910 1915 1920
 Thr Leu Thr Glu Asp Gly Ala Lys Ser Ser Glu Ala Ile Lys Glu Ser
 1925 1930 1935
 Ser Lys Phe Pro Phe Gly Ile Ser Pro Ala Gln Ser His Arg Asn Ile
 1940 1945 1950
 Lys Ile Leu Glu Asp Glu Pro His Ser Lys Asp Glu Thr Pro Leu Cys
 1955 1960 1965
 Thr Leu Leu Asp Trp Gln Asp Ser Leu Ala Lys Arg Cys Val Cys Val
 1970 1975 1980
 Ser Asn Thr Ile Arg Ser Leu Ser Phe Val Pro Gly Asn Asp Phe Glu
 1985 1990 1995 2000
 Met Ser Lys His Pro Gly Leu Leu Leu Ile Leu Gly Lys Leu Ile Leu
 2005 2010 2015
 Leu His His Lys His Pro Glu Arg Lys Gln Ala Pro Leu Thr Tyr Glu
 2020 2025 2030
 Lys Glu Glu Glu Gln Asp Gln Gly Val Ser Cys Asn Lys Val Glu Trp
 2035 2040 2045
 Trp Trp Asp Cys Leu Glu Met Leu Arg Glu Asn Thr Leu Val Thr Leu
 2050 2055 2060
 Ala Asn Ile Ser Gly Gln Leu Asp Leu Ser Pro Tyr Pro Glu Ser Ile
 2065 2070 2075 2080
 Cys Leu Pro Val Leu Asp Gly Leu Leu His Trp Ala Val Cys Pro Ser
 2085 2090 2095
 Ala Glu Ala Gln Asp Pro Phe Ser Thr Leu Gly Pro Asn Ala Val Leu
 2100 2105 2110
 Ser Pro Gln Arg Leu Val Leu Glu Thr Leu Ser Lys Leu Ser Ile Gln
 2115 2120 2125
 Asp Asn Asn Val Asp Leu Ile Leu Ala Thr Pro Pro Phe Ser Arg Leu
 2130 2135 2140
 Glu Lys Leu Tyr Ser Thr Met Val Arg Phe Leu Ser Asp Arg Lys Asn
 2145 2150 2155 2160
 Pro Val Cys Arg Glu Met Ala Val Val Leu Leu Ala Asn Leu Ala Gln
 2165 2170 2175
 Gly Asp Ser Leu Ala Ala Arg Ala Ile Ala Val Gln Lys Gly Ser Ile
 2180 2185 2190
 Gly Asn Leu Leu Gly Phe Leu Glu Asp Ser Leu Ala Ala Thr Gln Phe
 2195 2200 2205
 Gln Gln Ser Gln Ala Ser Leu Leu His Met Gln Asn Pro Pro Phe Glu
 2210 2215 2220
 Pro Thr Ser Val Asp Met Met Arg Arg Ala Ala Arg Ala Leu Leu Ala
 2225 2230 2235 2240
 Leu Ala Lys Val Asp Glu Asn His Ser Glu Phe Thr Leu Tyr Glu Ser
 2245 2250 2255
 Arg Leu Leu Asp Ile Ser Val Ser Pro Leu Met Asn Ser Leu Val Ser
 2260 2265 2270

Gln Val Ile Cys Asp Val Leu Phe Leu Ile Gly Gln Ser
 2275 2280 2285

<210> 114
 <211> 275
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >ATG5|ENSG00000057663|ENST00000369076|828

<400> 114
 Met Thr Asp Asp Lys Asp Val Leu Arg Asp Val Trp Phe Gly Arg Ile
 1 5 10 15
 Pro Thr Cys Phe Thr Leu Tyr Gln Asp Glu Ile Thr Glu Arg Glu Ala
 20 25 30
 Glu Pro Tyr Tyr Leu Leu Leu Pro Arg Val Ser Tyr Leu Thr Leu Val
 35 40 45
 Thr Asp Lys Val Lys Lys His Phe Gln Lys Val Met Arg Gln Glu Asp
 50 55 60
 Ile Ser Glu Ile Trp Phe Glu Tyr Glu Gly Thr Pro Leu Lys Trp His
 65 70 75 80
 Tyr Pro Ile Gly Leu Leu Phe Asp Leu Leu Ala Ser Ser Ser Ala Leu
 85 90 95
 Pro Trp Asn Ile Thr Val His Phe Lys Ser Phe Pro Glu Lys Asp Leu
 100 105 110
 Leu His Cys Pro Ser Lys Asp Ala Ile Glu Ala His Phe Met Ser Cys
 115 120 125
 Met Lys Glu Ala Asp Ala Leu Lys His Lys Ser Gln Val Ile Asn Glu
 130 135 140
 Met Gln Lys Lys Asp His Lys Gln Leu Trp Met Gly Leu Gln Asn Asp
 145 150 155 160
 Arg Phe Asp Gln Phe Trp Ala Ile Asn Arg Lys Leu Met Glu Tyr Pro
 165 170 175
 Ala Glu Glu Asn Gly Phe Arg Tyr Ile Pro Phe Arg Ile Tyr Gln Thr
 180 185 190
 Thr Thr Glu Arg Pro Phe Ile Gln Lys Leu Phe Arg Pro Val Ala Ala
 195 200 205
 Asp Gly Gln Leu His Thr Leu Gly Asp Leu Leu Lys Glu Val Cys Pro
 210 215 220
 Ser Ala Ile Asp Pro Glu Asp Gly Glu Lys Lys Asn Gln Val Met Ile
 225 230 235 240
 His Gly Ile Glu Pro Met Leu Glu Thr Pro Leu Gln Trp Leu Ser Glu
 245 250 255
 His Leu Ser Tyr Pro Asp Asn Phe Leu His Ile Ser Ile Ile Pro Gln
 260 265 270
 Pro Thr Asp
 275

<210> 115
 <211> 3056
 <212> PRT
 <213> Homo sapiens

<220>

<223> >ATM|ENSG00000149311|ENST00000278616|9171

<400> 115

Met	Ser	Leu	Val	Leu	Asn	Asp	Leu	Leu	Ile	Cys	Cys	Arg	Gln	Leu	Glu
1				5					10					15	
His	Asp	Arg	Ala	Thr	Glu	Arg	Lys	Lys	Glu	Val	Glu	Lys	Phe	Lys	Arg
			20					25					30		
Leu	Ile	Arg	Asp	Pro	Glu	Thr	Ile	Lys	His	Leu	Asp	Arg	His	Ser	Asp
		35					40					45			
Ser	Lys	Gln	Gly	Lys	Tyr	Leu	Asn	Trp	Asp	Ala	Val	Phe	Arg	Phe	Leu
	50					55					60				
Gln	Lys	Tyr	Ile	Gln	Lys	Glu	Thr	Glu	Cys	Leu	Arg	Ile	Ala	Lys	Pro
65					70					75					80
Asn	Val	Ser	Ala	Ser	Thr	Gln	Ala	Ser	Arg	Gln	Lys	Lys	Met	Gln	Glu
				85					90					95	
Ile	Ser	Ser	Leu	Val	Lys	Tyr	Phe	Ile	Lys	Cys	Ala	Asn	Arg	Arg	Ala
			100					105					110		
Pro	Arg	Leu	Lys	Cys	Gln	Glu	Leu	Leu	Asn	Tyr	Ile	Met	Asp	Thr	Val
		115					120					125			
Lys	Asp	Ser	Ser	Asn	Gly	Ala	Ile	Tyr	Gly	Ala	Asp	Cys	Ser	Asn	Ile
	130				135						140				
Leu	Leu	Lys	Asp	Ile	Leu	Ser	Val	Arg	Lys	Tyr	Trp	Cys	Glu	Ile	Ser
145					150					155					160
Gln	Gln	Gln	Trp	Leu	Glu	Leu	Phe	Ser	Val	Tyr	Phe	Arg	Leu	Tyr	Leu
				165					170					175	
Lys	Pro	Ser	Gln	Asp	Val	His	Arg	Val	Leu	Val	Ala	Arg	Ile	Ile	His
			180				185						190		
Ala	Val	Thr	Lys	Gly	Cys	Cys	Ser	Gln	Thr	Asp	Gly	Leu	Asn	Ser	Lys
		195					200					205			
Phe	Leu	Asp	Phe	Phe	Ser	Lys	Ala	Ile	Gln	Cys	Ala	Arg	Gln	Glu	Lys
	210					215					220				
Ser	Ser	Ser	Gly	Leu	Asn	His	Ile	Leu	Ala	Ala	Leu	Thr	Ile	Phe	Leu
225					230					235					240
Lys	Thr	Leu	Ala	Val	Asn	Phe	Arg	Ile	Arg	Val	Cys	Glu	Leu	Gly	Asp
				245					250					255	
Glu	Ile	Leu	Pro	Thr	Leu	Leu	Tyr	Ile	Trp	Thr	Gln	His	Arg	Leu	Asn
			260				265						270		
Asp	Ser	Leu	Lys	Glu	Val	Ile	Ile	Glu	Leu	Phe	Gln	Leu	Gln	Ile	Tyr
		275					280					285			
Ile	His	His	Pro	Lys	Gly	Ala	Lys	Thr	Gln	Glu	Lys	Gly	Ala	Tyr	Glu
	290					295					300				
Ser	Thr	Lys	Trp	Arg	Ser	Ile	Leu	Tyr	Asn	Leu	Tyr	Asp	Leu	Leu	Val
305					310					315					320
Asn	Glu	Ile	Ser	His	Ile	Gly	Ser	Arg	Gly	Lys	Tyr	Ser	Ser	Gly	Phe
				325					330					335	
Arg	Asn	Ile	Ala	Val	Lys	Glu	Asn	Leu	Ile	Glu	Leu	Met	Ala	Asp	Ile
			340					345					350		
Cys	His	Gln	Val	Phe	Asn	Glu	Asp	Thr	Arg	Ser	Leu	Glu	Ile	Ser	Gln
		355					360					365			
Ser	Tyr	Thr	Thr	Thr	Gln	Arg	Glu	Ser	Ser	Asp	Tyr	Ser	Val	Pro	Cys
	370					375					380				

Lys	Arg	Lys	Lys	Ile	Glu	Leu	Gly	Trp	Glu	Val	Ile	Lys	Asp	His	Leu	385	390	395	400
Gln	Lys	Ser	Gln	Asn	Asp	Phe	Asp	Leu	Val	Pro	Trp	Leu	Gln	Ile	Ala	405	410	415	
Thr	Gln	Leu	Ile	Ser	Lys	Tyr	Pro	Ala	Ser	Leu	Pro	Asn	Cys	Glu	Leu	420	425	430	
Ser	Pro	Leu	Met	Ile	Leu	Ser	Gln	Leu	Leu	Pro	Gln	Gln	Arg	His		435	440	445	
Gly	Glu	Arg	Thr	Pro	Tyr	Val	Leu	Arg	Cys	Leu	Thr	Glu	Val	Ala	Leu	450	455	460	
Cys	Gln	Asp	Lys	Arg	Ser	Asn	Leu	Glu	Ser	Ser	Gln	Lys	Ser	Asp	Leu	465	470	475	480
Leu	Lys	Leu	Trp	Asn	Lys	Ile	Trp	Cys	Ile	Thr	Phe	Arg	Gly	Ile	Ser	485	490	495	
Ser	Glu	Gln	Ile	Gln	Ala	Glu	Asn	Phe	Gly	Leu	Leu	Gly	Ala	Ile	Ile	500	505	510	
Gln	Gly	Ser	Leu	Val	Glu	Val	Asp	Arg	Glu	Phe	Trp	Lys	Leu	Phe	Thr	515	520	525	
Gly	Ser	Ala	Cys	Arg	Pro	Ser	Cys	Pro	Ala	Val	Cys	Cys	Leu	Thr	Leu	530	535	540	
Ala	Leu	Thr	Thr	Ser	Ile	Val	Pro	Gly	Thr	Val	Lys	Met	Gly	Ile	Glu	545	550	555	560
Gln	Asn	Met	Cys	Glu	Val	Asn	Arg	Ser	Phe	Ser	Leu	Lys	Glu	Ser	Ile	565	570	575	
Met	Lys	Trp	Leu	Leu	Phe	Tyr	Gln	Leu	Glu	Gly	Asp	Leu	Glu	Asn	Ser	580	585	590	
Thr	Glu	Val	Pro	Pro	Ile	Leu	His	Ser	Asn	Phe	Pro	His	Leu	Val	Leu	595	600	605	
Glu	Lys	Ile	Leu	Val	Ser	Leu	Thr	Met	Lys	Asn	Cys	Lys	Ala	Ala	Met	610	615	620	
Asn	Phe	Phe	Gln	Ser	Val	Pro	Glu	Cys	Glu	His	His	Gln	Lys	Asp	Lys	625	630	635	640
Glu	Glu	Leu	Ser	Phe	Ser	Glu	Val	Glu	Glu	Leu	Phe	Leu	Gln	Thr	Thr	645	650	655	
Phe	Asp	Lys	Met	Asp	Phe	Leu	Thr	Ile	Val	Arg	Glu	Cys	Gly	Ile	Glu	660	665	670	
Lys	His	Gln	Ser	Ser	Ile	Gly	Phe	Ser	Val	His	Gln	Asn	Leu	Lys	Glu	675	680	685	
Ser	Leu	Asp	Arg	Cys	Leu	Leu	Gly	Leu	Ser	Glu	Gln	Leu	Leu	Asn	Asn	690	695	700	
Tyr	Ser	Ser	Glu	Ile	Thr	Asn	Ser	Glu	Thr	Leu	Val	Arg	Cys	Ser	Arg	705	710	715	720
Leu	Leu	Val	Gly	Val	Leu	Gly	Cys	Tyr	Cys	Tyr	Met	Gly	Val	Ile	Ala	725	730	735	
Glu	Glu	Glu	Ala	Tyr	Lys	Ser	Glu	Leu	Phe	Gln	Lys	Ala	Lys	Ser	Leu	740	745	750	
Met	Gln	Cys	Ala	Gly	Glu	Ser	Ile	Thr	Leu	Phe	Lys	Asn	Lys	Thr	Asn	755	760	765	
Glu	Glu	Phe	Arg	Ile	Gly	Ser	Leu	Arg	Asn	Met	Met	Gln	Leu	Cys	Thr	770	775	780	
Arg	Cys	Leu	Ser	Asn	Cys	Thr	Lys	Lys	Ser	Pro	Asn	Lys	Ile	Ala	Ser	785	790	795	800
Gly	Phe	Phe	Leu	Arg	Leu	Leu	Thr	Ser	Lys	Leu	Met	Asn	Asp	Ile	Ala	805	810	815	

Asp	Ile	Cys	Lys	Ser	Leu	Ala	Ser	Phe	Ile	Lys	Lys	Pro	Phe	Asp	Arg	
			820					825					830			
Gly	Glu	Val	Glu	Ser	Met	Glu	Asp	Asp	Thr	Asn	Gly	Asn	Leu	Met	Glu	
			835				840					845				
Val	Glu	Asp	Gln	Ser	Ser	Met	Asn	Leu	Phe	Asn	Asp	Tyr	Pro	Asp	Ser	
			850			855					860					
Ser	Val	Ser	Asp	Ala	Asn	Glu	Pro	Gly	Glu	Ser	Gln	Ser	Thr	Ile	Gly	
865					870					875					880	
Ala	Ile	Asn	Pro	Leu	Ala	Glu	Glu	Tyr	Leu	Ser	Lys	Gln	Asp	Leu	Leu	
				885				890						895		
Phe	Leu	Asp	Met	Leu	Lys	Phe	Leu	Cys	Leu	Cys	Val	Thr	Thr	Ala	Gln	
			900					905					910			
Thr	Asn	Thr	Val	Ser	Phe	Arg	Ala	Ala	Asp	Ile	Arg	Arg	Lys	Leu	Leu	
			915				920					925				
Met	Leu	Ile	Asp	Ser	Ser	Thr	Leu	Glu	Pro	Thr	Lys	Ser	Leu	His	Leu	
						935					940					
His	Met	Tyr	Leu	Met	Leu	Leu	Lys	Glu	Leu	Pro	Gly	Glu	Glu	Tyr	Pro	
945					950					955					960	
Leu	Pro	Met	Glu	Asp	Val	Leu	Glu	Leu	Leu	Lys	Pro	Leu	Ser	Asn	Val	
				965				970						975		
Cys	Ser	Leu	Tyr	Arg	Arg	Asp	Gln	Asp	Val	Cys	Lys	Thr	Ile	Leu	Asn	
			980					985					990			
His	Val	Leu	His	Val	Val	Lys	Asn	Leu	Gly	Gln	Ser	Asn	Met	Asp	Ser	
			995				1000					1005				
Glu	Asn	Thr	Arg	Asp	Ala	Gln	Gly	Gln	Phe	Leu	Thr	Val	Ile	Gly	Ala	
						1015					1020					
Phe	Trp	His	Leu	Thr	Lys	Glu	Arg	Lys	Tyr	Ile	Phe	Ser	Val	Arg	Met	
1025					1030					1035					1040	
Ala	Leu	Val	Asn	Cys	Leu	Lys	Thr	Leu	Leu	Glu	Ala	Asp	Pro	Tyr	Ser	
				1045					1050					1055		
Lys	Trp	Ala	Ile	Leu	Asn	Val	Met	Gly	Lys	Asp	Phe	Pro	Val	Asn	Glu	
			1060					1065					1070			
Val	Phe	Thr	Gln	Phe	Leu	Ala	Asp	Asn	His	His	Gln	Val	Arg	Met	Leu	
			1075				1080					1085				
Ala	Ala	Glu	Ser	Ile	Asn	Arg	Leu	Phe	Gln	Asp	Thr	Lys	Gly	Asp	Ser	
						1095					1100					
Ser	Arg	Leu	Leu	Lys	Ala	Leu	Pro	Leu	Lys	Leu	Gln	Gln	Thr	Ala	Phe	
1105					1110					1115					1120	
Glu	Asn	Ala	Tyr	Leu	Lys	Ala	Gln	Glu	Gly	Met	Arg	Glu	Met	Ser	His	
				1125					1130					1135		
Ser	Ala	Glu	Asn	Pro	Glu	Thr	Leu	Asp	Glu	Ile	Tyr	Asn	Arg	Lys	Ser	
			1140					1145					1150			
Val	Leu	Leu	Thr	Leu	Ile	Ala	Val	Val	Leu	Ser	Cys	Ser	Pro	Ile	Cys	
			1155				1160					1165				
Glu	Lys	Gln	Ala	Leu	Phe	Ala	Leu	Cys	Lys	Ser	Val	Lys	Glu	Asn	Gly	
						1175					1180					
Leu	Glu	Pro	His	Leu	Val	Lys	Lys	Val	Leu	Glu	Lys	Val	Ser	Glu	Thr	
1185					1190					1195					1200	
Phe	Gly	Tyr	Arg	Arg	Leu	Glu	Asp	Phe	Met	Ala	Ser	His	Leu	Asp	Tyr	
				1205					1210					1215		
Leu	Val	Leu	Glu	Trp	Leu	Asn	Leu	Gln	Asp	Thr	Glu	Tyr	Asn	Leu	Ser	
			1220					1225					1230			
Ser	Phe	Pro	Phe	Ile	Leu	Leu	Asn	Tyr	Thr	Asn	Ile	Glu	Asp	Phe	Tyr	
			1235				1240					1245				

Arg	Ser	Cys	Tyr	Lys	Val	Leu	Ile	Pro	His	Leu	Val	Ile	Arg	Ser	His	1250	1255	1260	
Phe	Asp	Glu	Val	Lys	Ser	Ile	Ala	Asn	Gln	Ile	Gln	Glu	Asp	Trp	Lys	1265	1270	1275	1280
Ser	Leu	Leu	Thr	Asp	Cys	Phe	Pro	Lys	Ile	Leu	Val	Asn	Ile	Leu	Pro	1285	1290	1295	
Tyr	Phe	Ala	Tyr	Glu	Gly	Thr	Arg	Asp	Ser	Gly	Met	Ala	Gln	Gln	Arg	1300	1305	1310	
Glu	Thr	Ala	Thr	Lys	Val	Tyr	Asp	Met	Leu	Lys	Ser	Glu	Asn	Leu	Leu	1315	1320	1325	
Gly	Lys	Gln	Ile	Asp	His	Leu	Phe	Ile	Ser	Asn	Leu	Pro	Glu	Ile	Val	1330	1335	1340	
Val	Glu	Leu	Leu	Met	Thr	Leu	His	Glu	Pro	Ala	Asn	Ser	Ser	Ala	Ser	1345	1350	1355	1360
Gln	Ser	Thr	Asp	Leu	Cys	Asp	Phe	Ser	Gly	Asp	Leu	Asp	Pro	Ala	Pro	1365	1370	1375	
Asn	Pro	Pro	His	Phe	Pro	Ser	His	Val	Ile	Lys	Ala	Thr	Phe	Ala	Tyr	1380	1385	1390	
Ile	Ser	Asn	Cys	His	Lys	Thr	Lys	Leu	Lys	Ser	Ile	Leu	Glu	Ile	Leu	1395	1400	1405	
Ser	Lys	Ser	Pro	Asp	Ser	Tyr	Gln	Lys	Ile	Leu	Leu	Ala	Ile	Cys	Glu	1410	1415	1420	
Gln	Ala	Ala	Glu	Thr	Asn	Asn	Val	Tyr	Lys	Lys	His	Arg	Ile	Leu	Lys	1425	1430	1435	1440
Ile	Tyr	His	Leu	Phe	Val	Ser	Leu	Leu	Leu	Lys	Asp	Ile	Lys	Ser	Gly	1445	1450	1455	
Leu	Gly	Gly	Ala	Trp	Ala	Phe	Val	Leu	Arg	Asp	Val	Ile	Tyr	Thr	Leu	1460	1465	1470	
Ile	His	Tyr	Ile	Asn	Gln	Arg	Pro	Ser	Cys	Ile	Met	Asp	Val	Ser	Leu	1475	1480	1485	
Arg	Ser	Phe	Ser	Leu	Cys	Cys	Asp	Leu	Leu	Ser	Gln	Val	Cys	Gln	Thr	1490	1495	1500	
Ala	Val	Thr	Tyr	Cys	Lys	Asp	Ala	Leu	Glu	Asn	His	Leu	His	Val	Ile	1505	1510	1515	1520
Val	Gly	Thr	Leu	Ile	Pro	Leu	Val	Tyr	Glu	Gln	Val	Glu	Val	Gln	Lys	1525	1530	1535	
Gln	Val	Leu	Asp	Leu	Leu	Lys	Tyr	Leu	Val	Ile	Asp	Asn	Lys	Asp	Asn	1540	1545	1550	
Glu	Asn	Leu	Tyr	Ile	Thr	Ile	Lys	Leu	Leu	Asp	Pro	Phe	Pro	Asp	His	1555	1560	1565	
Val	Val	Phe	Lys	Asp	Leu	Arg	Ile	Thr	Gln	Gln	Lys	Ile	Lys	Tyr	Ser	1570	1575	1580	
Arg	Gly	Pro	Phe	Ser	Leu	Leu	Glu	Glu	Ile	Asn	His	Phe	Leu	Ser	Val	1585	1590	1595	1600
Ser	Val	Tyr	Asp	Ala	Leu	Pro	Leu	Thr	Arg	Leu	Glu	Gly	Leu	Lys	Asp	1605	1610	1615	
Leu	Arg	Arg	Gln	Leu	Glu	Leu	His	Lys	Asp	Gln	Met	Val	Asp	Ile	Met	1620	1625	1630	
Arg	Ala	Ser	Gln	Asp	Asn	Pro	Gln	Asp	Gly	Ile	Met	Val	Lys	Leu	Val	1635	1640	1645	
Val	Asn	Leu	Leu	Gln	Leu	Ser	Lys	Met	Ala	Ile	Asn	His	Thr	Gly	Glu	1650	1655	1660	
Lys	Glu	Val	Leu	Glu	Ala	Val	Gly	Ser	Cys	Leu	Gly	Glu	Val	Gly	Pro	1665	1670	1675	1680

Page 363

Thr	Ser	Val	Ser	Lys	Glu	Val	Glu	Gly	Thr	Ser	Tyr	His	Glu	Ser	Leu	2115	2120	2125
Tyr	Asn	Ala	Leu	Gln	Ser	Leu	Arg	Asp	Arg	Glu	Phe	Ser	Thr	Phe	Tyr	2130	2135	2140
Glu	Ser	Leu	Lys	Tyr	Ala	Arg	Val	Lys	Glu	Val	Glu	Glu	Met	Cys	Lys	2145	2150	2155
Arg	Ser	Leu	Glu	Ser	Val	Tyr	Ser	Leu	Tyr	Pro	Thr	Leu	Ser	Arg	Leu	2165	2170	2175
Gln	Ala	Ile	Gly	Glu	Leu	Glu	Ser	Ile	Gly	Glu	Leu	Phe	Ser	Arg	Ser	2180	2185	2190
Val	Thr	His	Arg	Gln	Leu	Ser	Glu	Val	Tyr	Ile	Lys	Trp	Gln	Lys	His	2195	2200	2205
Ser	Gln	Leu	Leu	Lys	Asp	Ser	Asp	Phe	Ser	Phe	Gln	Glu	Pro	Ile	Met	2210	2215	2220
Ala	Leu	Arg	Thr	Val	Ile	Leu	Glu	Ile	Leu	Met	Glu	Lys	Glu	Met	Asp	2225	2230	2235
Asn	Ser	Gln	Arg	Glu	Cys	Ile	Lys	Asp	Ile	Leu	Thr	Lys	His	Leu	Val	2245	2250	2255
Glu	Leu	Ser	Ile	Leu	Ala	Arg	Thr	Phe	Lys	Asn	Thr	Gln	Leu	Pro	Glu	2260	2265	2270
Arg	Ala	Ile	Phe	Gln	Ile	Lys	Gln	Tyr	Asn	Ser	Val	Ser	Cys	Gly	Val	2275	2280	2285
Ser	Glu	Trp	Gln	Leu	Glu	Glu	Ala	Gln	Val	Phe	Trp	Ala	Lys	Lys	Glu	2290	2295	2300
Gln	Ser	Leu	Ala	Leu	Ser	Ile	Leu	Lys	Gln	Met	Ile	Lys	Lys	Leu	Asp	2305	2310	2315
Ala	Ser	Cys	Ala	Ala	Asn	Asn	Pro	Ser	Leu	Lys	Leu	Thr	Tyr	Thr	Glu	2325	2330	2335
Cys	Leu	Arg	Val	Cys	Gly	Asn	Trp	Leu	Ala	Glu	Thr	Cys	Leu	Glu	Asn	2340	2345	2350
Pro	Ala	Val	Ile	Met	Gln	Thr	Tyr	Leu	Glu	Lys	Ala	Val	Glu	Val	Ala	2355	2360	2365
Gly	Asn	Tyr	Asp	Gly	Glu	Ser	Ser	Asp	Glu	Leu	Arg	Asn	Gly	Lys	Met	2370	2375	2380
Lys	Ala	Phe	Leu	Ser	Leu	Ala	Arg	Phe	Ser	Asp	Thr	Gln	Tyr	Gln	Arg	2385	2390	2395
Ile	Glu	Asn	Tyr	Met	Lys	Ser	Ser	Glu	Phe	Glu	Asn	Lys	Gln	Ala	Leu	2405	2410	2415
Leu	Lys	Arg	Ala	Lys	Glu	Glu	Val	Gly	Leu	Leu	Arg	Glu	His	Lys	Ile	2420	2425	2430
Gln	Thr	Asn	Arg	Tyr	Thr	Val	Lys	Val	Gln	Arg	Glu	Leu	Glu	Leu	Asp	2435	2440	2445
Glu	Leu	Ala	Leu	Arg	Ala	Leu	Lys	Glu	Asp	Arg	Lys	Arg	Phe	Leu	Cys	2450	2455	2460
Lys	Ala	Val	Glu	Asn	Tyr	Ile	Asn	Cys	Leu	Leu	Ser	Gly	Glu	Glu	His	2465	2470	2475
Asp	Met	Trp	Val	Phe	Arg	Leu	Cys	Ser	Leu	Trp	Leu	Glu	Asn	Ser	Gly	2485	2490	2495
Val	Ser	Glu	Val	Asn	Gly	Met	Met	Lys	Arg	Asp	Gly	Met	Lys	Ile	Pro	2500	2505	2510
Thr	Tyr	Lys	Phe	Leu	Pro	Leu	Met	Tyr	Gln	Leu	Ala	Ala	Arg	Met	Gly	2515	2520	2525
Thr	Lys	Met	Met	Gly	Gly	Leu	Gly	Phe	His	Glu	Val	Leu	Asn	Asn	Leu	2530	2535	2540

Ile	Ser	Arg	Ile	Ser	Met	Asp	His	Pro	His	His	Thr	Leu	Phe	Ile	Ile	2545	2550	2555	2560
Leu	Ala	Leu	Ala	Asn	Ala	Asn	Arg	Asp	Glu	Phe	Leu	Thr	Lys	Pro	Glu		2565	2570	2575
Val	Ala	Arg	Arg	Ser	Arg	Ile	Thr	Lys	Asn	Val	Pro	Lys	Gln	Ser	Ser	2580	2585	2590	
Gln	Leu	Asp	Glu	Asp	Arg	Thr	Glu	Ala	Ala	Asn	Arg	Ile	Ile	Cys	Thr	2595	2600	2605	
Ile	Arg	Ser	Arg	Arg	Pro	Gln	Met	Val	Arg	Ser	Val	Glu	Ala	Leu	Cys	2610	2615	2620	
Asp	Ala	Tyr	Ile	Ile	Leu	Ala	Asn	Leu	Asp	Ala	Thr	Gln	Trp	Lys	Thr	2625	2630	2635	2640
Gln	Arg	Lys	Gly	Ile	Asn	Ile	Pro	Ala	Asp	Gln	Pro	Ile	Thr	Lys	Leu	2645	2650	2655	
Lys	Asn	Leu	Glu	Asp	Val	Val	Val	Pro	Thr	Met	Glu	Ile	Lys	Val	Asp	2660	2665	2670	
His	Thr	Gly	Glu	Tyr	Gly	Asn	Leu	Val	Thr	Ile	Gln	Ser	Phe	Lys	Ala	2675	2680	2685	
Glu	Phe	Arg	Leu	Ala	Gly	Gly	Val	Asn	Leu	Pro	Lys	Ile	Ile	Asp	Cys	2690	2695	2700	
Val	Gly	Ser	Asp	Gly	Lys	Glu	Arg	Arg	Gln	Leu	Val	Lys	Gly	Arg	Asp	2705	2710	2715	2720
Asp	Leu	Arg	Gln	Asp	Ala	Val	Met	Gln	Gln	Val	Phe	Gln	Met	Cys	Asn	2725	2730	2735	
Thr	Leu	Leu	Gln	Arg	Asn	Thr	Glu	Thr	Arg	Lys	Arg	Lys	Leu	Thr	Ile	2740	2745	2750	
Cys	Thr	Tyr	Lys	Val	Val	Pro	Leu	Ser	Gln	Arg	Ser	Gly	Val	Leu	Glu	2755	2760	2765	
Trp	Cys	Thr	Gly	Thr	Val	Pro	Ile	Gly	Glu	Phe	Leu	Val	Asn	Asn	Glu	2770	2775	2780	
Asp	Gly	Ala	His	Lys	Arg	Tyr	Arg	Pro	Asn	Asp	Phe	Ser	Ala	Phe	Gln	2785	2790	2795	2800
Cys	Gln	Lys	Lys	Met	Met	Glu	Val	Gln	Lys	Lys	Ser	Phe	Glu	Glu	Lys	2805	2810	2815	
Tyr	Glu	Val	Phe	Met	Asp	Val	Cys	Gln	Asn	Phe	Gln	Pro	Val	Phe	Arg	2820	2825	2830	
Tyr	Phe	Cys	Met	Glu	Lys	Phe	Leu	Asp	Pro	Ala	Ile	Trp	Phe	Glu	Lys	2835	2840	2845	
Arg	Leu	Ala	Tyr	Thr	Arg	Ser	Val	Ala	Thr	Ser	Ser	Ile	Val	Gly	Tyr	2850	2855	2860	
Ile	Leu	Gly	Leu	Gly	Asp	Arg	His	Val	Gln	Asn	Ile	Leu	Ile	Asn	Glu	2865	2870	2875	2880
Gln	Ser	Ala	Glu	Leu	Val	His	Ile	Asp	Leu	Gly	Val	Ala	Phe	Glu	Gln	2885	2890	2895	
Gly	Lys	Ile	Leu	Pro	Thr	Pro	Glu	Thr	Val	Pro	Phe	Arg	Leu	Thr	Arg	2900	2905	2910	
Asp	Ile	Val	Asp	Gly	Met	Gly	Ile	Thr	Gly	Val	Glu	Gly	Val	Phe	Arg	2915	2920	2925	
Arg	Cys	Cys	Glu	Lys	Thr	Met	Glu	Val	Met	Arg	Asn	Ser	Gln	Glu	Thr	2930	2935	2940	
Leu	Leu	Thr	Ile	Val	Glu	Val	Leu	Leu	Tyr	Asp	Pro	Leu	Phe	Asp	Trp	2945	2950	2955	2960
Thr	Met	Asn	Pro	Leu	Lys	Ala	Leu	Tyr	Leu	Gln	Gln	Arg	Pro	Glu	Asp	2965	2970	2975	

Glu	Thr	Glu	Leu	His	Pro	Thr	Leu	Asn	Ala	Asp	Asp	Gln	Glu	Cys	Lys
			2980					2985					2990		
Arg	Asn	Leu	Ser	Asp	Ile	Asp	Gln	Ser	Phe	Asn	Lys	Val	Ala	Glu	Arg
	2995						3000					3005			
Val	Leu	Met	Arg	Leu	Gln	Glu	Lys	Leu	Lys	Gly	Val	Glu	Glu	Gly	Thr
	3010					3015					3020				
Val	Leu	Ser	Val	Gly	Gly	Gln	Val	Asn	Leu	Leu	Ile	Gln	Gln	Ala	Ile
3025				3030					3035						3040
Asp	Pro	Lys	Asn	Leu	Ser	Arg	Leu	Phe	Pro	Gly	Trp	Lys	Ala	Trp	Val
			3045					3050						3055	

<210> 116
 <211> 2644
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >ATR|ENSG00000175054|ENST00000350721|7935

<400> 116

Met	Gly	Glu	His	Gly	Leu	Glu	Leu	Ala	Ser	Met	Ile	Pro	Ala	Leu	Arg
1			5					10						15	
Glu	Leu	Gly	Ser	Ala	Thr	Pro	Glu	Glu	Tyr	Asn	Thr	Val	Val	Gln	Lys
		20					25					30			
Pro	Arg	Gln	Ile	Leu	Cys	Gln	Phe	Ile	Asp	Arg	Ile	Leu	Thr	Asp	Val
	35					40					45				
Asn	Val	Val	Ala	Val	Glu	Leu	Val	Lys	Lys	Thr	Asp	Ser	Gln	Pro	Thr
	50				55					60					
Ser	Val	Met	Leu	Leu	Asp	Phe	Ile	Gln	His	Ile	Met	Lys	Ser	Ser	Pro
65				70					75						80
Leu	Met	Phe	Val	Asn	Val	Ser	Gly	Ser	His	Glu	Ala	Lys	Gly	Ser	Cys
			85					90					95		
Ile	Glu	Phe	Ser	Asn	Trp	Ile	Ile	Thr	Arg	Leu	Leu	Arg	Ile	Ala	Ala
	100						105					110			
Thr	Pro	Ser	Cys	His	Leu	Leu	His	Lys	Lys	Ile	Cys	Glu	Val	Ile	Cys
	115					120					125				
Ser	Leu	Leu	Phe	Leu	Phe	Lys	Ser	Lys	Ser	Pro	Ala	Ile	Phe	Gly	Val
	130				135					140					
Leu	Thr	Lys	Glu	Leu	Leu	Gln	Leu	Phe	Glu	Asp	Leu	Val	Tyr	Leu	His
145				150					155						160
Arg	Arg	Asn	Val	Met	Gly	His	Ala	Val	Glu	Trp	Pro	Val	Val	Met	Ser
			165					170						175	
Arg	Phe	Leu	Ser	Gln	Leu	Asp	Glu	His	Met	Gly	Tyr	Leu	Gln	Ser	Ala
	180					185							190		
Pro	Leu	Gln	Leu	Met	Ser	Met	Gln	Asn	Leu	Glu	Phe	Ile	Glu	Val	Thr
	195					200						205			
Leu	Leu	Met	Val	Leu	Thr	Arg	Ile	Ile	Ala	Ile	Val	Phe	Phe	Arg	Arg
	210					215					220				
Gln	Glu	Leu	Leu	Leu	Trp	Gln	Ile	Gly	Cys	Val	Leu	Leu	Glu	Tyr	Gly
225				230					235						240
Ser	Pro	Lys	Ile	Lys	Ser	Leu	Ala	Ile	Ser	Phe	Leu	Thr	Glu	Leu	Phe

				245					250					255			
Gln	Leu	Gly	Gly	Leu	Pro	Ala	Gln	Pro	Ala	Ser	Thr	Phe	Phe	Ser	Ser		
			260					265					270				
Phe	Leu	Glu	Leu	Leu	Lys	His	Leu	Val	Glu	Met	Asp	Thr	Asp	Gln	Leu		
		275					280					285					
Lys	Leu	Tyr	Glu	Glu	Pro	Leu	Ser	Lys	Leu	Ile	Lys	Thr	Leu	Phe	Pro		
		290				295					300						
Phe	Glu	Ala	Glu	Ala	Tyr	Arg	Asn	Ile	Glu	Pro	Val	Tyr	Leu	Asn	Met		
305					310				315						320		
Leu	Leu	Glu	Lys	Leu	Cys	Val	Met	Phe	Glu	Asp	Gly	Val	Leu	Met	Arg		
			325						330					335			
Leu	Lys	Ser	Asp	Leu	Leu	Lys	Ala	Ala	Leu	Cys	His	Leu	Leu	Gln	Tyr		
			340					345					350				
Phe	Leu	Lys	Phe	Val	Pro	Ala	Gly	Tyr	Glu	Ser	Ala	Leu	Gln	Val	Arg		
		355					360					365					
Lys	Val	Tyr	Val	Arg	Asn	Ile	Cys	Lys	Ala	Leu	Leu	Asp	Val	Leu	Gly		
		370				375					380						
Ile	Glu	Val	Asp	Ala	Glu	Tyr	Leu	Leu	Gly	Pro	Leu	Tyr	Ala	Ala	Leu		
385					390				395						400		
Lys	Met	Glu	Ser	Met	Glu	Ile	Ile	Glu	Glu	Ile	Gln	Cys	Gln	Thr	Gln		
				405					410					415			
Gln	Glu	Asn	Leu	Ser	Ser	Asn	Ser	Asp	Gly	Ile	Ser	Pro	Lys	Arg	Arg		
			420					425					430				
Arg	Leu	Ser	Ser	Ser	Leu	Asn	Pro	Ser	Lys	Arg	Ala	Pro	Lys	Gln	Thr		
		435					440					445					
Glu	Glu	Ile	Lys	His	Val	Asp	Met	Asn	Gln	Lys	Ser	Ile	Leu	Trp	Ser		
		450				455					460						
Ala	Leu	Lys	Gln	Lys	Ala	Glu	Ser	Leu	Gln	Ile	Ser	Leu	Glu	Tyr	Ser		
465					470					475					480		
Gly	Leu	Lys	Asn	Pro	Val	Ile	Glu	Met	Leu	Glu	Gly	Ile	Ala	Val	Val		
				485					490					495			
Leu	Gln	Leu	Thr	Ala	Leu	Cys	Thr	Val	His	Cys	Ser	His	Gln	Asn	Met		
			500					505					510				
Asn	Cys	Arg	Thr	Phe	Lys	Asp	Cys	Gln	His	Lys	Ser	Lys	Lys	Lys	Pro		
		515					520					525					
Ser	Val	Val	Ile	Thr	Trp	Met	Ser	Leu	Asp	Phe	Tyr	Thr	Lys	Val	Leu		
		530				535					540						
Lys	Ser	Cys	Arg	Ser	Leu	Glu	Ser	Val	Gln	Lys	Leu	Asp	Leu	Glu			
545					550				555					560			
Ala	Thr	Ile	Asp	Lys	Val	Val	Lys	Ile	Tyr	Asp	Ala	Leu	Ile	Tyr	Met		
				565					570					575			
Gln	Val	Asn	Ser	Ser	Phe	Glu	Asp	His	Ile	Leu	Glu	Asp	Leu	Cys	Gly		
			580					585					590				
Met	Leu	Ser	Leu	Pro	Trp	Ile	Tyr	Ser	His	Ser	Asp	Asp	Gly	Cys	Leu		
		595					600					605					
Lys	Leu	Thr	Thr	Phe	Ala	Ala	Asn	Leu	Leu	Thr	Leu	Ser	Cys	Arg	Ile		
					615						620						
Ser	Asp	Ser	Tyr	Ser	Pro	Gln	Ala	Gln	Ser	Arg	Cys	Val	Phe	Leu	Leu		
625					630					635					640		
Thr	Leu	Phe	Pro	Arg	Arg	Ile	Phe	Leu	Glu	Trp	Arg	Thr	Ala	Val	Tyr		
				645					650					655			
Asn	Trp	Ala	Leu	Gln	Ser	Ser	His	Glu	Val	Ile	Arg	Ala	Ser	Cys	Val		
			660					665				670					
Ser	Gly	Phe	Phe	Ile	Leu	Leu	Gln	Gln	Gln	Asn	Ser	Cys	Asn	Arg	Val		

		675				680				685					
Pro	Lys	Ile	Leu	Ile	Asp	Lys	Val	Lys	Asp	Asp	Ser	Asp	Ile	Val	Lys
	690					695					700				
Lys	Glu	Phe	Ala	Ser	Ile	Leu	Gly	Gln	Leu	Val	Cys	Thr	Leu	His	Gly
705					710						715				720
Met	Phe	Tyr	Leu	Thr	Ser	Ser	Leu	Thr	Glu	Pro	Phe	Ser	Glu	His	Gly
				725					730					735	
His	Val	Asp	Leu	Phe	Cys	Arg	Asn	Leu	Lys	Ala	Thr	Ser	Gln	His	Glu
		740					745						750		
Cys	Ser	Ser	Ser	Gln	Leu	Lys	Ala	Ser	Val	Cys	Lys	Pro	Phe	Leu	Phe
		755					760					765			
Leu	Leu	Lys	Lys	Lys	Ile	Pro	Ser	Pro	Val	Lys	Leu	Ala	Phe	Ile	Asp
	770					775					780				
Asn	Leu	His	His	Leu	Cys	Lys	His	Leu	Asp	Phe	Arg	Glu	Asp	Glu	Thr
785					790					795					800
Asp	Val	Lys	Ala	Val	Leu	Gly	Thr	Leu	Leu	Asn	Leu	Met	Glu	Asp	Pro
				805					810					815	
Asp	Lys	Asp	Val	Arg	Val	Ala	Phe	Ser	Gly	Asn	Ile	Lys	His	Ile	Leu
			820					825					830		
Glu	Ser	Leu	Asp	Ser	Glu	Asp	Gly	Phe	Ile	Lys	Glu	Leu	Phe	Val	Leu
		835					840					845			
Arg	Met	Lys	Glu	Ala	Tyr	Thr	His	Ala	Gln	Ile	Ser	Arg	Asn	Asn	Glu
	850					855					860				
Leu	Lys	Asp	Thr	Leu	Ile	Leu	Thr	Thr	Gly	Asp	Ile	Gly	Arg	Ala	Ala
865					870					875					880
Lys	Gly	Asp	Leu	Val	Pro	Phe	Ala	Leu	Leu	His	Leu	Leu	His	Cys	Leu
				885					890					895	
Leu	Ser	Lys	Ser	Ala	Ser	Val	Ser	Gly	Ala	Ala	Tyr	Thr	Glu	Ile	Arg
			900					905					910		
Ala	Leu	Val	Ala	Ala	Lys	Ser	Val	Lys	Leu	Gln	Ser	Phe	Phe	Ser	Gln
		915					920					925			
Tyr	Lys	Lys	Pro	Ile	Cys	Gln	Phe	Leu	Val	Glu	Ser	Leu	His	Ser	Ser
	930					935					940				
Gln	Met	Thr	Ala	Leu	Pro	Asn	Thr	Pro	Cys	Gln	Asn	Ala	Asp	Val	Arg
945					950					955					960
Lys	Gln	Asp	Val	Ala	His	Gln	Arg	Glu	Met	Ala	Leu	Asn	Thr	Leu	Ser
				965					970					975	
Glu	Ile	Ala	Asn	Val	Phe	Asp	Phe	Pro	Asp	Leu	Asn	Arg	Phe	Leu	Thr
		980						985					990		
Arg	Thr	Leu	Gln	Val	Leu	Leu	Pro	Asp	Leu	Ala	Ala	Lys	Ala	Ser	Pro
		995					1000					1005			
Ala	Ala	Ser	Ala	Leu	Ile	Arg	Thr	Leu	Gly	Lys	Gln	Leu	Asn	Val	Asn
	1010					1015					1020				
Arg	Arg	Glu	Ile	Leu	Ile	Asn	Asn	Phe	Lys	Tyr	Ile	Phe	Ser	His	Leu
1025					1030					1035					1040
Val	Cys	Ser	Cys	Ser	Lys	Asp	Glu	Leu	Glu	Arg	Ala	Leu	His	Tyr	Leu
				1045					1050					1055	
Lys	Asn	Glu	Thr	Glu	Ile	Glu	Leu	Gly	Ser	Leu	Leu	Arg	Gln	Asp	Phe
			1060					1065					1070		
Gln	Gly	Leu	His	Asn	Glu	Leu	Leu	Leu	Arg	Ile	Gly	Glu	His	Tyr	Gln
	1075					1080					1085				
Gln	Val	Phe	Asn	Gly	Leu	Ser	Ile	Leu	Ala	Ser	Phe	Ala	Ser	Ser	Asp
	1090					1095					1100				
Asp	Pro	Tyr	Gln	Gly	Pro	Arg	Asp	Ile	Ile	Ser	Pro	Glu	Leu	Met	Ala

1105		1110		1115		1120
Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met						
	1125		1130		1135	
Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu						
	1140		1145		1150	
Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser						
	1155		1160		1165	
Ser Val Arg Val Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe						
	1170		1175		1180	
Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val						
1185		1190		1195		1200
Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile						
	1205		1210		1215	
Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala						
	1220		1225		1230	
Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe						
	1235		1240		1245	
Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile						
	1250		1255		1260	
Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp						
1265		1270		1275		1280
Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn						
	1285		1290		1295	
Val Asp Val Arg Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr						
	1300		1305		1310	
Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val						
	1315		1320		1325	
Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln						
	1330		1335		1340	
Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu						
1345		1350		1355		1360
Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr						
	1365		1370		1375	
Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe						
	1380		1385		1390	
Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala						
	1395		1400		1405	
Asp Asn Ser Arg Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu						
	1410		1415		1420	
Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His						
1425		1430		1435		1440
Gln Leu Trp Arg Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro						
	1445		1450		1455	
His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser						
	1460		1465		1470	
Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala						
	1475		1480		1485	
Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His						
	1490		1495		1500	
Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His						
1505		1510		1515		1520
Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val						
	1525		1530		1535	
Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile						

1970	1975	1980
Pro Glu Asn Glu Thr	Pro Pro Glu Gly Lys	Asn Met Leu Ile His Gly
1985	1990	1995
Arg Ala Met Leu Leu Val	Gly Arg Phe Met Glu Glu Thr	Ala Asn Phe
2005	2010	2015
Glu Ser Asn Ala Ile Met	Lys Lys Tyr Lys Asp Val Thr	Ala Cys Leu
2020	2025	2030
Pro Glu Trp Glu Asp Gly	His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys	
2035	2040	2045
Leu Met Pro Met Val Thr	Asp Asn Lys Met Glu Lys Gln Gly Asp Leu	
2050	2055	2060
Ile Arg Tyr Ile Val Leu	His Phe Gly Arg Ser Leu Gln Tyr Gly Asn	
2065	2070	2075
Gln Phe Ile Tyr Gln Ser	Met Pro Arg Met Leu Thr Leu Trp Leu Asp	
2085	2090	2095
Tyr Gly Thr Lys Ala Tyr	Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg	
2100	2105	2110
Val Gln Met Arg Asn Asp	Leu Gly Lys Ile Asn Lys Val Ile Thr Glu	
2115	2120	2125
His Thr Asn Tyr Leu Ala	Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln	
2130	2135	2140
Leu Ile Ser Arg Ile Cys	His Ser His Asp Glu Val Phe Val Val Leu	
2145	2150	2155
Met Glu Ile Ile Ala Lys	Val Phe Leu Ala Tyr Pro Gln Gln Ala Met	
2165	2170	2175
Trp Met Met Thr Ala Val	Ser Lys Ser Ser Tyr Pro Met Arg Val Asn	
2180	2185	2190
Arg Cys Lys Glu Ile Leu	Asn Lys Ala Ile His Met Lys Lys Ser Leu	
2195	2200	2205
Glu Lys Phe Val Gly Asp	Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu	
2210	2215	2220
Leu Cys Asn Lys Pro Val	Asp Gly Ser Ser Ser Thr Leu Ser Met Ser	
2225	2230	2235
Thr His Phe Lys Met Leu	Lys Lys Leu Val Glu Glu Ala Thr Phe Ser	
2245	2250	2255
Glu Ile Leu Ile Pro Leu	Gln Ser Val Met Ile Pro Thr Leu Pro Ser	
2260	2265	2270
Ile Leu Gly Thr His Ala	Asn His Ala Ser His Glu Pro Phe Pro Gly	
2275	2280	2285
His Trp Ala Tyr Ile Ala	Gly Phe Asp Asp Met Val Glu Ile Leu Ala	
2290	2295	2300
Ser Leu Gln Lys Pro Lys	Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys	
2305	2310	2315
Phe Tyr Ile Met Met Cys	Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys	
2325	2330	2335
Arg Leu Met Glu Phe Asn	Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp	
2340	2345	2350
Ala Glu Ser Arg Arg Arg	Glu Leu His Ile Arg Thr Tyr Ala Val Ile	
2355	2360	2365
Pro Leu Asn Asp Glu Cys	Gly Ile Ile Glu Trp Val Asn Asn Thr Ala	
2370	2375	2380
Gly Leu Arg Pro Ile Leu	Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr	
2385	2390	2395
Met Thr Gly Lys Glu Leu	Arg Gln Cys Met Leu Pro Lys Ser Ala Ala	

Asn	Pro	Ser	Gly	Lys	Asn	Arg	Glu	Thr	Val	Pro	Ile	Lys	Asp	Asn	Phe	
			100					105					110			
Glu	Leu	Glu	Val	Leu	Gln	Ala	Gln	Tyr	Lys	Glu	Leu	Lys	Glu	Lys	Met	
		115					120					125				
Lys	Val	Met	Glu	Glu	Glu	Val	Leu	Ile	Lys	Asn	Gly	Glu	Ile	Lys	Ile	
		130				135					140					
Leu	Arg	Asp	Ser	Leu	His	Gln	Thr	Glu	Ser	Val	Leu	Glu	Glu	Gln	Arg	
145					150					155					160	
Arg	Ser	His	Phe	Leu	Leu	Glu	Gln	Glu	Lys	Thr	Gln	Ala	Leu	Ser	Asp	
			165						170						175	
Lys	Glu	Lys	Glu	Phe	Ser	Lys	Lys	Leu	Gln	Ser	Leu	Gln	Ser	Glu	Leu	
			180					185					190			
Gln	Phe	Lys	Asp	Ala	Glu	Met	Asn	Glu	Leu	Arg	Thr	Lys	Leu	Gln	Thr	
		195					200					205				
Ser	Glu	Arg	Ala	Asn	Lys	Leu	Ala	Ala	Pro	Ser	Val	Ser	His	Val	Ser	
		210				215					220					
Pro	Arg	Lys	Asn	Pro	Ser	Val	Val	Ile	Lys	Pro	Glu	Ala	Cys	Ser	Pro	
225					230					235					240	
Gln	Phe	Gly	Lys	Thr	Ser	Phe	Pro	Thr	Lys	Glu	Ser	Phe	Ser	Ala	Asn	
				245					250						255	
Met	Ser	Leu	Pro	His	Pro	Cys	Gln	Thr	Glu	Ser	Gly	Tyr	Lys	Pro	Leu	
			260					265					270			
Val	Gly	Arg	Glu	Asp	Ser	Lys	Pro	His	Ser	Leu	Arg	Gly	Asp	Ser	Ile	
		275					280					285				
Lys	Gln	Glu	Glu	Ala	Gln	Lys	Ser	Phe	Val	Asp	Ser	Trp	Arg	Gln	Arg	
		290				295				300						
Ser	Asn	Thr	Gln	Gly	Ser	Ile	Leu	Ile	Asn	Leu	Leu	Leu	Lys	Gln	Pro	
305					310				315						320	
Leu	Ile	Pro	Gly	Ser	Ser	Leu	Ser	Leu	Cys	His	Leu	Leu	Ser	Ser	Ser	
				325					330					335		
Ser	Glu	Ser	Pro	Ala	Gly	Thr	Pro	Leu	Gln	Pro	Pro	Gly	Phe	Gly	Ser	
			340					345					350			
Thr	Leu	Ala	Gly	Met	Ser	Gly	Leu	Arg	Thr	Thr	Gly	Ser	Tyr	Asp	Gly	
		355					360					365				
Ser	Phe	Ser	Leu	Ser	Ala	Leu	Arg	Glu	Ala	Gln	Asn	Leu	Ala	Phe	Thr	
						375					380					
Gly	Leu	Asn	Leu	Val	Ala	Arg	Asn	Glu	Cys	Ser	Arg	Asp	Gly	Asp	Pro	
385					390					395					400	
Ala	Glu	Gly	Gly	Arg	Arg	Ala	Phe	Pro	Leu	Cys	Gln	Leu	Pro	Gly	Ala	
				405				410						415		
Val	His	Phe	Leu	Pro	Leu	Val	Gln	Phe	Phe	Ile	Gly	Leu	His	Cys	Gln	
			420					425					430			
Ala	Leu	Gln	Asp	Leu	Ala	Ala	Ala	Lys	Arg	Ser	Gly	Ala	Pro	Gly	Asp	
		435					440					445				
Ser	Pro	Thr	His	Ser	Ser	Cys	Val	Ser	Ser	Gly	Val	Glu	Thr	Asn	Pro	
		450				455					460					
Glu	Asp	Ser	Val	Cys	Ile	Leu	Glu	Gly	Phe	Ser	Val	Thr	Ala	Leu	Ser	
465					470					475					480	
Ile	Leu	Gln	His	Leu	Val	Cys	His	Ser	Gly	Ala	Val	Val	Ser	Leu	Leu	
				485					490					495		
Leu	Ser	Gly	Val	Gly	Ala	Asp	Ser	Ala	Ala	Gly	Glu	Gly	Asn	Arg	Ser	
			500					505					510			
Leu	Val	His	Arg	Leu	Ser	Asp	Gly	Asp	Met	Thr	Ser	Ala	Leu	Arg	Gly	
		515					520					525				

Val	Ala	Asp	Asp	Gln	Gly	Gln	His	Pro	Leu	Leu	Lys	Met	Leu	Leu	His
530						535					540				
Leu	Leu	Ala	Phe	Ser	Ser	Ala	Ala	Thr	Gly	His	Leu	Gln	Ala	Ser	Val
545					550					555					560
Leu	Thr	Gln	Cys	Leu	Lys	Val	Leu	Val	Lys	Leu	Ala	Glu	Asn	Thr	Ser
				565					570						575
Cys	Asp	Phe	Leu	Pro	Arg	Phe	Gln	Cys	Val	Phe	Gln	Val	Leu	Pro	Lys
			580					585					590		
Cys	Leu	Ser	Pro	Glu	Thr	Pro	Leu	Pro	Ser	Val	Leu	Leu	Ala	Val	Glu
		595					600					605			
Leu	Leu	Ser	Leu	Leu	Ala	Asp	His	Asp	Gln	Leu	Ala	Pro	Gln	Leu	Cys
610						615					620				
Ser	His	Ser	Glu	Gly	Cys	Leu	Leu	Leu	Leu	Leu	Tyr	Met	Tyr	Ile	Thr
625					630					635					640
Ser	Arg	Pro	Asp	Arg	Val	Ala	Leu	Glu	Thr	Gln	Trp	Leu	Gln	Leu	Glu
				645					650						655
Gln	Glu	Val	Val	Trp	Leu	Leu	Ala	Lys	Leu	Gly	Val	Gln	Ser	Pro	Leu
			660					665					670		
Pro	Pro	Val	Thr	Gly	Ser	Asn	Cys	Gln	Cys	Asn	Val	Glu	Val	Val	Arg
		675					680					685			
Ala	Leu	Thr	Val	Met	Leu	His	Arg	Gln	Trp	Leu	Thr	Val	Arg	Arg	Ala
690						695					700				
Gly	Gly	Pro	Pro	Arg	Thr	Asp	Gln	Gln	Arg	Arg	Thr	Val	Arg	Cys	Leu
705					710				715						720
Arg	Asp	Thr	Val	Leu	Leu	Leu	His	Gly	Leu	Ser	Gln	Lys	Asp	Lys	Leu
				725				730						735	
Phe	Met	Met	His	Cys	Val	Glu	Val	Leu	His	Gln	Phe	Asp	Gln	Val	Met
			740					745					750		
Pro	Gly	Val	Ser	Met	Leu	Ile	Arg	Gly	Leu	Pro	Asp	Val	Thr	Asp	Cys
		755					760					765			
Glu	Glu	Ala	Ala	Leu	Asp	Asp	Leu	Cys	Ala	Ala	Glu	Thr	Asp	Val	Glu
770					775						780				
Asp	Pro	Glu	Val	Glu	Cys	Gly									
785					790										

<210> 118

<211> 2492

<212> PRT

<213> Homo sapiens

<220>

<223> >ATRX|ENSG00000085224|ENST00000373344|7479

<400> 118

Met	Thr	Ala	Glu	Pro	Met	Ser	Glu	Ser	Lys	Leu	Asn	Thr	Leu	Val	Gln
1				5					10					15	
Lys	Leu	His	Asp	Phe	Leu	Ala	His	Ser	Ser	Glu	Glu	Ser	Glu	Glu	Thr
			20					25					30		
Ser	Ser	Pro	Pro	Arg	Leu	Ala	Met	Asn	Gln	Asn	Thr	Asp	Lys	Ile	Ser
		35					40					45			
Gly	Ser	Gly	Ser	Asn	Ser	Asp	Met	Met	Glu	Asn	Ser	Lys	Glu	Glu	Gly
50						55					60				
Thr	Ser	Ser	Ser	Glu	Lys	Ser	Lys	Ser	Ser	Gly	Ser	Ser	Arg	Ser	Lys

65					70				75					80	
Arg	Lys	Pro	Ser	Ile	Val	Thr	Lys	Tyr	Val	Glu	Ser	Asp	Asp	Glu	Lys
				85					90					95	
Pro	Leu	Asp	Asp	Glu	Thr	Val	Asn	Glu	Asp	Ala	Ser	Asn	Glu	Asn	Ser
			100					105					110		
Glu	Asn	Asp	Ile	Thr	Met	Gln	Ser	Leu	Pro	Lys	Gly	Thr	Val	Ile	Val
		115					120					125			
Gln	Pro	Glu	Pro	Val	Leu	Asn	Glu	Asp	Lys	Asp	Asp	Phe	Lys	Gly	Pro
	130					135				140					
Glu	Phe	Arg	Ser	Arg	Ser	Lys	Met	Lys	Thr	Glu	Asn	Leu	Lys	Lys	Arg
145				150						155					160
Gly	Glu	Asp	Gly	Leu	His	Gly	Ile	Val	Ser	Cys	Thr	Ala	Cys	Gly	Gln
				165					170					175	
Gln	Val	Asn	His	Phe	Gln	Lys	Asp	Ser	Ile	Tyr	Arg	His	Pro	Ser	Leu
			180				185						190		
Gln	Val	Leu	Ile	Cys	Lys	Asn	Cys	Phe	Lys	Tyr	Tyr	Met	Ser	Asp	Asp
	195					200						205			
Ile	Ser	Arg	Asp	Ser	Asp	Gly	Met	Asp	Glu	Gln	Cys	Arg	Trp	Cys	Ala
	210					215					220				
Glu	Gly	Gly	Asn	Leu	Ile	Cys	Cys	Asp	Phe	Cys	His	Asn	Ala	Phe	Cys
225				230						235					240
Lys	Lys	Cys	Ile	Leu	Arg	Asn	Leu	Gly	Arg	Lys	Glu	Leu	Ser	Thr	Ile
				245					250					255	
Met	Asp	Glu	Asn	Asn	Gln	Trp	Tyr	Cys	Tyr	Ile	Cys	His	Pro	Glu	Pro
			260					265					270		
Leu	Leu	Asp	Leu	Val	Thr	Ala	Cys	Asn	Ser	Val	Phe	Glu	Asn	Leu	Glu
		275					280					285			
Gln	Leu	Leu	Gln	Gln	Asn	Lys	Lys	Lys	Ile	Lys	Val	Asp	Ser	Glu	Lys
	290				295						300				
Ser	Asn	Lys	Val	Tyr	Glu	His	Thr	Ser	Arg	Phe	Ser	Pro	Lys	Lys	Thr
305				310						315					320
Ser	Ser	Asn	Cys	Asn	Gly	Glu	Glu	Lys	Lys	Leu	Asp	Asp	Ser	Cys	Ser
			325					330						335	
Gly	Ser	Val	Thr	Tyr	Ser	Tyr	Ser	Ala	Leu	Ile	Val	Pro	Lys	Glu	Met
			340					345					350		
Ile	Lys	Lys	Ala	Lys	Lys	Leu	Ile	Glu	Thr	Thr	Ala	Asn	Met	Asn	Ser
	355					360						365			
Ser	Tyr	Val	Lys	Phe	Leu	Lys	Gln	Ala	Thr	Asp	Asn	Ser	Glu	Ile	Ser
	370				375					380					
Ser	Ala	Thr	Lys	Leu	Arg	Gln	Leu	Lys	Ala	Phe	Lys	Ser	Val	Leu	Ala
385				390						395					400
Asp	Ile	Lys	Lys	Ala	His	Leu	Ala	Leu	Glu	Glu	Asp	Leu	Asn	Ser	Glu
				405					410					415	
Phe	Arg	Ala	Met	Asp	Ala	Val	Asn	Lys	Glu	Lys	Asn	Thr	Lys	Glu	His
			420					425					430		
Lys	Val	Ile	Asp	Ala	Lys	Phe	Glu	Thr	Lys	Ala	Arg	Lys	Gly	Glu	Lys
	435						440					445			
Pro	Cys	Ala	Leu	Glu	Lys	Lys	Asp	Ile	Ser	Lys	Ser	Glu	Ala	Lys	Leu
	450					455					460				
Ser	Arg	Lys	Gln	Val	Asp	Ser	Glu	His	Met	His	Gln	Asn	Val	Pro	Thr
465				470						475					480
Glu	Glu	Gln	Arg	Thr	Asn	Lys	Ser	Thr	Gly	Gly	Glu	His	Lys	Lys	Ser
				485					490					495	
Asp	Arg	Lys	Glu	Glu	Pro	Gln	Tyr	Glu	Pro	Ala	Asn	Thr	Ser	Glu	Asp

			500					505				510			
Leu	Asp	Met	Asp	Ile	Val	Ser	Val	Pro	Ser	Ser	Val	Pro	Glu	Asp	Ile
		515					520					525			
Phe	Glu	Asn	Leu	Glu	Thr	Ala	Met	Glu	Val	Gln	Ser	Ser	Val	Asp	His
	530					535					540				
Gln	Gly	Asp	Gly	Ser	Ser	Gly	Thr	Glu	Gln	Glu	Val	Glu	Ser	Ser	Ser
545				550						555					560
Val	Lys	Leu	Asn	Ile	Ser	Ser	Lys	Asp	Asn	Arg	Gly	Gly	Ile	Lys	Ser
			565					570						575	
Lys	Thr	Thr	Ala	Lys	Val	Thr	Lys	Glu	Leu	Tyr	Val	Lys	Leu	Thr	Pro
			580					585					590		
Val	Ser	Leu	Ser	Asn	Ser	Pro	Ile	Lys	Gly	Ala	Asp	Cys	Gln	Glu	Val
		595					600					605			
Pro	Gln	Asp	Lys	Asp	Gly	Tyr	Lys	Ser	Cys	Gly	Leu	Asn	Pro	Lys	Leu
	610					615					620				
Glu	Lys	Cys	Gly	Leu	Gly	Gln	Glu	Asn	Ser	Asp	Asn	Glu	His	Leu	Val
625				630						635					640
Glu	Asn	Glu	Val	Ser	Leu	Leu	Leu	Glu	Glu	Ser	Asp	Leu	Arg	Arg	Ser
			645					650						655	
Pro	Arg	Val	Lys	Thr	Thr	Pro	Leu	Arg	Arg	Pro	Thr	Glu	Thr	Asn	Pro
			660					665						670	
Val	Thr	Ser	Asn	Ser	Asp	Glu	Glu	Cys	Asn	Glu	Thr	Val	Lys	Glu	Lys
		675					680					685			
Gln	Lys	Leu	Ser	Val	Pro	Val	Arg	Lys	Lys	Asp	Lys	Arg	Asn	Ser	Ser
	690					695				700					
Asp	Ser	Ala	Ile	Asp	Asn	Pro	Lys	Pro	Asn	Lys	Leu	Pro	Lys	Ser	Lys
705				710						715					720
Gln	Ser	Glu	Thr	Val	Asp	Gln	Asn	Ser	Asp	Ser	Asp	Glu	Met	Leu	Ala
			725					730						735	
Ile	Leu	Lys	Glu	Val	Ser	Arg	Met	Ser	His	Ser	Ser	Ser	Ser	Asp	Thr
			740					745						750	
Asp	Ile	Asn	Glu	Ile	His	Thr	Asn	His	Lys	Thr	Leu	Tyr	Asp	Leu	Lys
		755					760					765			
Thr	Gln	Ala	Gly	Lys	Asp	Asp	Lys	Gly	Lys	Arg	Lys	Arg	Lys	Ser	Ser
	770					775					780				
Thr	Ser	Gly	Ser	Asp	Phe	Asp	Thr	Lys	Lys	Gly	Lys	Ser	Ala	Lys	Ser
785				790						795					800
Ser	Ile	Ile	Ser	Lys	Lys	Arg	Gln	Thr	Gln	Ser	Glu	Ser	Ser	Ser	Asn
			805					810						815	
Tyr	Asp	Ser	Glu	Leu	Glu	Lys	Glu	Ile	Lys	Ser	Met	Ser	Lys	Ile	Gly
			820					825					830		
Ala	Ala	Arg	Thr	Thr	Lys	Lys	Arg	Ile	Pro	Asn	Thr	Lys	Asp	Phe	Asp
		835					840					845			
Ser	Ser	Glu	Asp	Glu	Lys	His	Ser	Lys	Lys	Gly	Met	Asp	Asn	Gln	Gly
	850					855					860				
His	Lys	Asn	Leu	Lys	Thr	Ser	Gln	Glu	Gly	Ser	Ser	Asp	Asp	Ala	Glu
865				870						875					880
Arg	Lys	Gln	Glu	Arg	Glu	Thr	Phe	Ser	Ser	Ala	Glu	Gly	Thr	Val	Asp
			885					890						895	
Lys	Asp	Thr	Thr	Ile	Met	Glu	Leu	Arg	Asp	Arg	Leu	Pro	Lys	Lys	Gln
			900					905					910		
Gln	Ala	Ser	Ala	Ser	Thr	Asp	Gly	Val	Asp	Lys	Leu	Ser	Gly	Lys	Glu
	915					920						925			
Gln	Ser	Phe	Thr	Ser	Leu	Glu	Val	Arg	Lys	Val	Ala	Glu	Thr	Lys	Glu

930		935		940
Lys Ser Lys His Leu	Lys Thr Lys Thr Cys	Lys Lys Val Gln Asp Gly		
945	950	955		960
Leu Ser Asp Ile Ala	Glu Lys Phe Leu Lys	Lys Asp Gln Ser Asp Glu		
	965	970		975
Thr Ser Glu Asp Asp	Lys Lys Gln Ser Lys	Lys Gly Thr Glu Glu Lys		
	980	985		990
Lys Lys Pro Ser Asp	Phe Lys Lys Lys Val Ile	Lys Met Glu Gln Gln		
	995	1000		1005
Tyr Glu Ser Ser Ser	Asp Gly Thr Glu Lys	Leu Pro Glu Arg Glu Glu		
	1010	1015		1020
Ile Cys His Phe Pro	Lys Gly Ile Lys Gln Ile	Lys Asn Gly Thr Thr		
1025	1030	1035		1040
Asp Gly Glu Lys Lys	Ser Lys Lys Ile Arg	Asp Lys Thr Ser Lys Lys		
	1045	1050		1055
Lys Asp Glu Leu Ser	Asp Tyr Ala Glu Lys	Ser Thr Gly Lys Gly Asp		
	1060	1065		1070
Ser Cys Asp Ser Ser	Glu Asp Lys Lys Ser	Lys Asn Gly Ala Tyr Gly		
	1075	1080		1085
Arg Glu Lys Lys Arg	Cys Lys Leu Leu Gly	Lys Ser Ser Arg Lys Arg		
	1090	1095		1100
Gln Asp Cys Ser Ser	Ser Asp Thr Glu Lys	Tyr Ser Met Lys Glu Asp		
1105	1110	1115		1120
Gly Cys Asn Ser Ser	Asp Lys Arg Leu Lys	Arg Ile Glu Leu Arg Glu		
	1125	1130		1135
Arg Arg Asn Leu Ser	Ser Lys Arg Asn Thr	Lys Glu Ile Gln Ser Gly		
	1140	1145		1150
Ser Ser Ser Ser Asp	Ala Glu Glu Ser Ser	Glu Asp Asn Lys Lys Lys		
	1155	1160		1165
Lys Gln Arg Thr Ser	Ser Lys Lys Lys Ala	Val Ile Val Lys Glu Lys		
	1170	1175		1180
Lys Arg Asn Ser Leu	Arg Thr Ser Thr Lys	Arg Lys Gln Ala Asp Ile		
1185	1190	1195		1200
Thr Ser Ser Ser Ser	Ser Asp Ile Glu Asp	Asp Asp Gln Asn Ser Ile		
	1205	1210		1215
Gly Glu Gly Ser Ser	Asp Glu Gln Lys Ile	Lys Pro Val Thr Glu Asn		
	1220	1225		1230
Leu Val Leu Ser Ser	His Thr Gly Phe Cys	Gln Ser Ser Gly Asp Glu		
	1235	1240		1245
Ala Leu Ser Lys Ser	Val Pro Val Thr Val	Asp Asp Asp Asp Asp		
	1250	1255		1260
Asn Asp Pro Glu Asn	Arg Ile Ala Lys Lys	Met Leu Leu Glu Glu Ile		
1265	1270	1275		1280
Lys Ala Asn Leu Ser	Ser Asp Glu Asp Gly	Ser Ser Asp Asp Glu Pro		
	1285	1290		1295
Glu Glu Gly Lys Lys	Arg Thr Gly Lys Gln	Asn Glu Glu Asn Pro Gly		
	1300	1305		1310
Asp Glu Glu Ala Lys	Asn Gln Val Asn Ser	Glu Ser Asp Ser Asp Ser		
	1315	1320		1325
Glu Glu Ser Lys Lys	Pro Arg Tyr Arg His	Arg Leu Leu Arg His Lys		
	1330	1335		1340
Leu Thr Val Ser Asp	Gly Glu Ser Gly Glu	Glu Lys Lys Thr Lys Pro		
1345	1350	1355		1360
Lys Glu His Lys Glu	Val Lys Gly Arg Asn	Arg Arg Lys Val Ser Ser		

2225		2230		2235		2240									
Leu	Gln	Ile	His	Lys	Glu	His	Ile	Val	Gly	Tyr	His	Glu	His	Asp	Ser
				2245					2250					2255	
Leu	Leu	Asp	His	Lys	Glu	Glu	Glu	Glu	Leu	Thr	Glu	Glu	Glu	Arg	Lys
				2260				2265						2270	
Ala	Ala	Trp	Ala	Glu	Tyr	Glu	Ala	Glu	Lys	Lys	Gly	Leu	Thr	Met	Arg
				2275				2280						2285	
Phe	Asn	Ile	Pro	Thr	Gly	Thr	Asn	Leu	Pro	Pro	Val	Ser	Phe	Asn	Ser
				2290				2295						2300	
Gln	Thr	Pro	Tyr	Ile	Pro	Phe	Asn	Leu	Gly	Ala	Leu	Ser	Ala	Met	Ser
2305					2310					2315					2320
Asn	Gln	Gln	Leu	Glu	Asp	Leu	Ile	Asn	Gln	Gly	Arg	Glu	Lys	Val	Val
				2325						2330					2335
Glu	Ala	Thr	Asn	Ser	Val	Thr	Ala	Val	Arg	Ile	Gln	Pro	Leu	Glu	Asp
				2340					2345						2350
Ile	Ile	Ser	Ala	Val	Trp	Lys	Glu	Asn	Met	Asn	Leu	Ser	Glu	Ala	Gln
				2355					2360						2365
Val	Gln	Ala	Leu	Ala	Leu	Ser	Arg	Gln	Ala	Ser	Gln	Glu	Leu	Asp	Val
				2370				2375							2380
Lys	Arg	Arg	Glu	Ala	Ile	Tyr	Asn	Asp	Val	Leu	Thr	Lys	Gln	Gln	Met
2385					2390										2400
Leu	Ile	Ser	Cys	Val	Gln	Arg	Ile	Leu	Met	Asn	Arg	Arg	Leu	Gln	Gln
				2405						2410					2415
Gln	Tyr	Asn	Gln	Gln	Gln	Gln	Gln	Gln	Met	Thr	Tyr	Gln	Gln	Ala	Thr
				2420					2425						2430
Leu	Gly	His	Leu	Met	Met	Pro	Lys	Pro	Pro	Asn	Leu	Ile	Met	Asn	Pro
				2435				2440							2445
Ser	Asn	Tyr	Gln	Gln	Ile	Asp	Met	Arg	Gly	Met	Tyr	Gln	Pro	Val	Ala
				2450				2455							2460
Gly	Gly	Met	Gln	Pro	Pro	Pro	Leu	Gln	Arg	Ala	Pro	Pro	Pro	Met	Arg
2465					2470					2475					2480
Ser	Lys	Asn	Pro	Gly	Pro	Ser	Gln	Gly	Lys	Ser	Met				
				2485						2490					

<210> 119
 <211> 729
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >BAP1|ENSG00000163930|ENST00000460680|2190

<400> 119
 Met Asn Lys Gly Trp Leu Glu Leu Glu Ser Asp Pro Gly Leu Phe Thr
 1 5 10 15
 Leu Leu Val Glu Asp Phe Gly Val Lys Gly Val Gln Val Glu Glu Ile
 20 25 30
 Tyr Asp Leu Gln Ser Lys Cys Gln Gly Pro Val Tyr Gly Phe Ile Phe
 35 40 45
 Leu Phe Lys Trp Ile Glu Glu Arg Arg Ser Arg Arg Lys Val Ser Thr
 50 55 60
 Leu Val Asp Asp Thr Ser Val Ile Asp Asp Asp Ile Val Asn Asn Met
 65 70 75 80

Phe	Phe	Ala	His	Gln	Leu	Ile	Pro	Asn	Ser	Cys	Ala	Thr	His	Ala	Leu	85	90	95
Leu	Ser	Val	Leu	Leu	Asn	Cys	Ser	Ser	Val	Asp	Leu	Gly	Pro	Thr	Leu	100	105	110
Ser	Arg	Met	Lys	Asp	Phe	Thr	Lys	Gly	Phe	Ser	Pro	Glu	Ser	Lys	Gly	115	120	125
Tyr	Ala	Ile	Gly	Asn	Ala	Pro	Glu	Leu	Ala	Lys	Ala	His	Asn	Ser	His	130	135	140
Ala	Arg	Pro	Glu	Pro	Arg	His	Leu	Pro	Glu	Lys	Gln	Asn	Gly	Leu	Ser	145	150	155
Ala	Val	Arg	Thr	Met	Glu	Ala	Phe	His	Phe	Val	Ser	Tyr	Val	Pro	Ile	165	170	175
Thr	Gly	Arg	Leu	Phe	Glu	Leu	Asp	Gly	Leu	Lys	Val	Tyr	Pro	Ile	Asp	180	185	190
His	Gly	Pro	Trp	Gly	Glu	Asp	Glu	Glu	Trp	Thr	Asp	Lys	Ala	Arg	Arg	195	200	205
Val	Ile	Met	Glu	Arg	Ile	Gly	Leu	Ala	Thr	Ala	Gly	Glu	Pro	Tyr	His	210	215	220
Asp	Ile	Arg	Phe	Asn	Leu	Met	Ala	Val	Val	Pro	Asp	Arg	Arg	Ile	Lys	225	230	235
Tyr	Glu	Ala	Arg	Leu	His	Val	Leu	Lys	Val	Asn	Arg	Gln	Thr	Val	Leu	245	250	255
Glu	Ala	Leu	Gln	Gln	Leu	Ile	Arg	Val	Thr	Gln	Pro	Glu	Leu	Ile	Gln	260	265	270
Thr	His	Lys	Ser	Gln	Glu	Ser	Gln	Leu	Pro	Glu	Glu	Ser	Lys	Ser	Ala	275	280	285
Ser	Asn	Lys	Ser	Pro	Leu	Val	Leu	Glu	Ala	Asn	Arg	Ala	Pro	Ala	Ala	290	295	300
Ser	Glu	Gly	Asn	His	Thr	Asp	Gly	Ala	Glu	Glu	Ala	Ala	Gly	Ser	Cys	305	310	315
Ala	Gln	Ala	Pro	Ser	His	Ser	Pro	Pro	Asn	Lys	Pro	Lys	Leu	Val	Val	325	330	335
Lys	Pro	Pro	Gly	Ser	Ser	Leu	Asn	Gly	Val	His	Pro	Asn	Pro	Thr	Pro	340	345	350
Ile	Val	Gln	Arg	Leu	Pro	Ala	Phe	Leu	Asp	Asn	His	Asn	Tyr	Ala	Lys	355	360	365
Ser	Pro	Met	Gln	Glu	Glu	Glu	Asp	Leu	Ala	Ala	Gly	Val	Gly	Arg	Ser	370	375	380
Arg	Val	Pro	Val	Arg	Pro	Pro	Gln	Gln	Tyr	Ser	Asp	Asp	Glu	Asp	Asp	385	390	395
Tyr	Glu	Asp	Asp	Glu	Glu	Asp	Asp	Val	Gln	Asn	Thr	Asn	Ser	Ala	Leu	405	410	415
Arg	Tyr	Lys	Gly	Lys	Gly	Thr	Gly	Lys	Pro	Gly	Ala	Leu	Ser	Gly	Ser	420	425	430
Ala	Asp	Gly	Gln	Leu	Ser	Val	Leu	Gln	Pro	Asn	Thr	Ile	Asn	Val	Leu	435	440	445
Ala	Glu	Lys	Leu	Lys	Glu	Ser	Gln	Lys	Asp	Leu	Ser	Ile	Pro	Leu	Ser	450	455	460
Ile	Lys	Thr	Ser	Ser	Gly	Ala	Gly	Ser	Pro	Ala	Val	Ala	Val	Pro	Thr	465	470	475
His	Ser	Gln	Pro	Ser	Pro	Thr	Pro	Ser	Asn	Glu	Ser	Thr	Asp	Thr	Ala	485	490	495
Ser	Glu	Ile	Gly	Ser	Ala	Phe	Asn	Ser	Pro	Leu	Arg	Ser	Pro	Ile	Arg	500	505	510

Ser Ala Asn Pro Thr Arg Pro Ser Ser Pro Val Thr Ser His Ile Ser
 515 520 525
 Lys Val Leu Phe Gly Glu Asp Asp Ser Leu Leu Arg Val Asp Cys Ile
 530 535 540
 Arg Tyr Asn Arg Ala Val Arg Asp Leu Gly Pro Val Ile Ser Thr Gly
 545 550 555 560
 Leu Leu His Leu Ala Glu Asp Gly Val Leu Ser Pro Leu Ala Leu Thr
 565 570 575
 Glu Gly Gly Lys Gly Ser Ser Pro Ser Ile Arg Pro Ile Gln Gly Ser
 580 585 590
 Gln Gly Ser Ser Ser Pro Val Glu Lys Glu Val Val Glu Ala Thr Asp
 595 600 605
 Ser Arg Glu Lys Thr Gly Met Val Arg Pro Gly Glu Pro Leu Ser Gly
 610 615 620
 Glu Lys Tyr Ser Pro Lys Glu Leu Leu Ala Leu Leu Lys Cys Val Glu
 625 630 635 640
 Ala Glu Ile Ala Asn Tyr Glu Ala Cys Leu Lys Glu Glu Val Glu Lys
 645 650 655
 Arg Lys Lys Phe Lys Ile Asp Asp Gln Arg Arg Thr His Asn Tyr Asp
 660 665 670
 Glu Phe Ile Cys Thr Phe Ile Ser Met Leu Ala Gln Glu Gly Met Leu
 675 680 685
 Ala Asn Leu Val Glu Gln Asn Ile Ser Val Arg Arg Arg Gln Gly Val
 690 695 700
 Ser Ile Gly Arg Leu His Lys Gln Arg Lys Pro Asp Arg Arg Lys Arg
 705 710 715 720
 Ser Arg Pro Tyr Lys Ala Lys Arg Gln
 725

<210> 120

<211> 777

<212> PRT

<213> Homo sapiens

<220>

<223> >BARD1|ENSG00000138376|ENST00000260947|2334

<400> 120

Met Pro Asp Asn Arg Gln Pro Arg Asn Arg Gln Pro Arg Ile Arg Ser
 1 5 10 15
 Gly Asn Glu Pro Arg Ser Ala Pro Ala Met Glu Pro Asp Gly Arg Gly
 20 25 30
 Ala Trp Ala His Ser Arg Ala Ala Leu Asp Arg Leu Glu Lys Leu Leu
 35 40 45
 Arg Cys Ser Arg Cys Thr Asn Ile Leu Arg Glu Pro Val Cys Leu Gly
 50 55 60
 Gly Cys Glu His Ile Phe Cys Ser Asn Cys Val Ser Asp Cys Ile Gly
 65 70 75 80
 Thr Gly Cys Pro Val Cys Tyr Thr Pro Ala Trp Ile Gln Asp Leu Lys
 85 90 95
 Ile Asn Arg Gln Leu Asp Ser Met Ile Gln Leu Cys Ser Lys Leu Arg
 100 105 110
 Asn Leu Leu His Asp Asn Glu Leu Ser Asp Leu Lys Glu Asp Lys Pro

545	Asn	Thr	Gly	Gln	Arg	Arg	Asp	Gly	Pro	Leu	Val	Leu	Ile	Gly	Ser	Gly	560
					565					570						575	
Leu	Ser	Ser	Glu	Gln	Gln	Lys	Met	Leu	Ser	Glu	Leu	Ala	Val	Ile	Leu		
			580					585					590				
Lys	Ala	Lys	Lys	Tyr	Thr	Glu	Phe	Asp	Ser	Thr	Val	Thr	His	Val	Val		
		595					600					605					
Val	Pro	Gly	Asp	Ala	Val	Gln	Ser	Thr	Leu	Lys	Cys	Met	Leu	Gly	Ile		
	610					615					620						
Leu	Asn	Gly	Cys	Trp	Ile	Leu	Lys	Phe	Glu	Trp	Val	Lys	Ala	Cys	Leu		
625					630					635					640		
Arg	Arg	Lys	Val	Cys	Glu	Gln	Glu	Glu	Lys	Tyr	Glu	Ile	Pro	Glu	Gly		
				645					650					655			
Pro	Arg	Arg	Ser	Arg	Leu	Asn	Arg	Glu	Gln	Leu	Leu	Pro	Lys	Leu	Phe		
			660					665					670				
Asp	Gly	Cys	Tyr	Phe	Tyr	Leu	Trp	Gly	Thr	Phe	Lys	His	His	Pro	Lys		
		675					680					685					
Asp	Asn	Leu	Ile	Lys	Leu	Val	Thr	Ala	Gly	Gly	Gly	Gln	Ile	Leu	Ser		
	690					695				700							
Arg	Lys	Pro	Lys	Pro	Asp	Ser	Asp	Val	Thr	Gln	Thr	Ile	Asn	Thr	Val		
705					710					715					720		
Ala	Tyr	His	Ala	Arg	Pro	Asp	Ser	Asp	Gln	Arg	Phe	Cys	Thr	Gln	Tyr		
				725					730					735			
Ile	Ile	Tyr	Glu	Asp	Leu	Cys	Asn	Tyr	His	Pro	Glu	Arg	Val	Arg	Gln		
			740					745				750					
Gly	Lys	Val	Trp	Lys	Ala	Pro	Ser	Ser	Trp	Phe	Ile	Asp	Cys	Val	Met		
		755					760					765					
Ser	Phe	Glu	Leu	Leu	Pro	Leu	Asp	Ser									
	770					775											

<210> 121
 <211> 1417
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >BLM|ENSG00000197299|ENST00000355112|4254

<400> 121	Met	Ala	Ala	Val	Pro	Gln	Asn	Asn	Leu	Gln	Glu	Gln	Leu	Glu	Arg	His
1				5					10					15		
Ser	Ala	Arg	Thr	Leu	Asn	Asn	Lys	Leu	Ser	Leu	Ser	Lys	Pro	Lys	Phe	
		20						25				30				
Ser	Gly	Phe	Thr	Phe	Lys	Lys	Lys	Thr	Ser	Ser	Asp	Asn	Asn	Val	Ser	
	35					40					45					
Val	Thr	Asn	Val	Ser	Val	Ala	Lys	Thr	Pro	Val	Leu	Arg	Asn	Lys	Asp	
	50					55				60						
Val	Asn	Val	Thr	Glu	Asp	Phe	Ser	Phe	Ser	Glu	Pro	Leu	Pro	Asn	Thr	
65				70					75					80		
Thr	Asn	Gln	Gln	Arg	Val	Lys	Asp	Phe	Phe	Lys	Asn	Ala	Pro	Ala	Gly	
			85					90				95				
Gln	Glu	Thr	Gln	Arg	Gly	Gly	Ser	Lys	Ser	Leu	Leu	Pro	Asp	Phe	Leu	
			100				105					110				

Gln	Thr	Pro	Lys	Glu	Val	Val	Cys	Thr	Thr	Gln	Asn	Thr	Pro	Thr	Val
		115					120					125			
Lys	Lys	Ser	Arg	Asp	Thr	Ala	Leu	Lys	Lys	Leu	Glu	Phe	Ser	Ser	Ser
	130					135				140					
Pro	Asp	Ser	Leu	Ser	Thr	Ile	Asn	Asp	Trp	Asp	Asp	Met	Asp	Asp	Phe
145					150					155					160
Asp	Thr	Ser	Glu	Thr	Ser	Lys	Ser	Phe	Val	Thr	Pro	Pro	Gln	Ser	His
				165					170					175	
Phe	Val	Arg	Val	Ser	Thr	Ala	Gln	Lys	Ser	Lys	Lys	Gly	Lys	Arg	Asn
			180					185					190		
Phe	Phe	Lys	Ala	Gln	Leu	Tyr	Thr	Thr	Asn	Thr	Val	Lys	Thr	Asp	Leu
		195					200					205			
Pro	Pro	Pro	Ser	Ser	Glu	Ser	Glu	Gln	Ile	Asp	Leu	Thr	Glu	Glu	Gln
	210					215					220				
Lys	Asp	Asp	Ser	Glu	Trp	Leu	Ser	Ser	Asp	Val	Ile	Cys	Ile	Asp	Asp
225				230						235					240
Gly	Pro	Ile	Ala	Glu	Val	His	Ile	Asn	Glu	Asp	Ala	Gln	Glu	Ser	Asp
			245						250					255	
Ser	Leu	Lys	Thr	His	Leu	Glu	Asp	Glu	Arg	Asp	Asn	Ser	Glu	Lys	Lys
			260					265					270		
Lys	Asn	Leu	Glu	Glu	Ala	Glu	Leu	His	Ser	Thr	Glu	Lys	Val	Pro	Cys
		275					280					285			
Ile	Glu	Phe	Asp	Asp	Asp	Asp	Tyr	Asp	Thr	Asp	Phe	Val	Pro	Pro	Ser
	290					295					300				
Pro	Glu	Glu	Ile	Ile	Ser	Ala	Ser	Ser	Ser	Ser	Ser	Lys	Cys	Leu	Ser
305					310					315					320
Thr	Leu	Lys	Asp	Leu	Asp	Thr	Ser	Asp	Arg	Lys	Glu	Asp	Val	Leu	Ser
			325						330					335	
Thr	Ser	Lys	Asp	Leu	Leu	Ser	Lys	Pro	Glu	Lys	Met	Ser	Met	Gln	Glu
			340					345					350		
Leu	Asn	Pro	Glu	Thr	Ser	Thr	Asp	Cys	Asp	Ala	Arg	Gln	Ile	Ser	Leu
		355					360					365			
Gln	Gln	Gln	Leu	Ile	His	Val	Met	Glu	His	Ile	Cys	Lys	Leu	Ile	Asp
	370					375					380				
Thr	Ile	Pro	Asp	Asp	Lys	Leu	Lys	Leu	Leu	Asp	Cys	Gly	Asn	Glu	Leu
385					390					395					400
Leu	Gln	Gln	Arg	Asn	Ile	Arg	Arg	Lys	Leu	Leu	Thr	Glu	Val	Asp	Phe
			405						410					415	
Asn	Lys	Ser	Asp	Ala	Ser	Leu	Leu	Gly	Ser	Leu	Trp	Arg	Tyr	Arg	Pro
			420					425					430		
Asp	Ser	Leu	Asp	Gly	Pro	Met	Glu	Gly	Asp	Ser	Cys	Pro	Thr	Gly	Asn
		435					440					445			
Ser	Met	Lys	Glu	Leu	Asn	Phe	Ser	His	Leu	Pro	Ser	Asn	Ser	Val	Ser
	450					455					460				
Pro	Gly	As													

Arg	Glu	Thr	Gln	Pro	Ser	Tyr	Asp	Ile	Asp	Asn	Phe	Asp	Ile	Asp	Asp	545	550	555	560
Phe	Asp	Asp	Asp	Asp	Asp	Trp	Glu	Asp	Ile	Met	His	Asn	Leu	Ala	Ala	565	570	575	
Ser	Lys	Ser	Ser	Thr	Ala	Ala	Tyr	Gln	Pro	Ile	Lys	Glu	Gly	Arg	Pro	580	585	590	
Ile	Lys	Ser	Val	Ser	Glu	Arg	Leu	Ser	Ser	Ala	Lys	Thr	Asp	Cys	Leu	595	600	605	
Pro	Val	Ser	Ser	Thr	Ala	Gln	Asn	Ile	Asn	Phe	Ser	Glu	Ser	Ile	Gln	610	615	620	
Asn	Tyr	Thr	Asp	Lys	Ser	Ala	Gln	Asn	Leu	Ala	Ser	Arg	Asn	Leu	Lys	625	630	635	640
His	Glu	Arg	Phe	Gln	Ser	Leu	Ser	Phe	Pro	His	Thr	Lys	Glu	Met	Met	645	650	655	
Lys	Ile	Phe	His	Lys	Lys	Phe	Gly	Leu	His	Asn	Phe	Arg	Thr	Asn	Gln	660	665	670	
Leu	Glu	Ala	Ile	Asn	Ala	Ala	Leu	Leu	Gly	Glu	Asp	Cys	Phe	Ile	Leu	675	680	685	
Met	Pro	Thr	Gly	Gly	Gly	Lys	Ser	Leu	Cys	Tyr	Gln	Leu	Pro	Ala	Cys	690	695	700	
Val	Ser	Pro	Gly	Val	Thr	Val	Val	Ile	Ser	Pro	Leu	Arg	Ser	Leu	Ile	705	710	715	720
Val	Asp	Gln	Val	Gln	Lys	Leu	Thr	Ser	Leu	Asp	Ile	Pro	Ala	Thr	Tyr	725	730	735	
Leu	Thr	Gly	Asp	Lys	Thr	Asp	Ser	Glu	Ala	Thr	Asn	Ile	Tyr	Leu	Gln	740	745	750	
Leu	Ser	Lys	Lys	Asp	Pro	Ile	Ile	Lys	Leu	Leu	Tyr	Val	Thr	Pro	Glu	755	760	765	
Lys	Ile	Cys	Ala	Ser	Asn	Arg	Leu	Ile	Ser	Thr	Leu	Glu	Asn	Leu	Tyr	770	775	780	
Glu	Arg	Lys	Leu	Leu	Ala	Arg	Phe	Val	Ile	Asp	Glu	Ala	His	Cys	Val	785	790	795	800
Ser	Gln	Trp	Gly	His	Asp	Phe	Arg	Gln	Asp	Tyr	Lys	Arg	Met	Asn	Met	805	810	815	
Leu	Arg	Gln	Lys	Phe	Pro	Ser	Val	Pro	Val	Met	Ala	Leu	Thr	Ala	Thr	820	825	830	
Ala	Asn	Pro	Arg	Val	Gln	Lys	Asp	Ile	Leu	Thr	Gln	Leu	Lys	Ile	Leu	835	840	845	
Arg	Pro	Gln	Val	Phe	Ser	Met	Ser	Phe	Asn	Arg	His	Asn	Leu	Lys	Tyr	850	855	860	
Tyr	Val	Leu	Pro	Lys	Lys	Pro	Lys	Lys	Val	Ala	Phe	Asp	Cys	Leu	Glu	865	870	875	880
Trp	Ile	Arg	Lys	His	His	Pro	Tyr	Asp	Ser	Gly	Ile	Ile	Tyr	Cys	Leu	885	890	895	
Ser	Arg	Arg	Glu	Cys	Asp	Thr	Met	Ala	Asp	Thr	Leu	Gln	Arg	Asp	Gly	900	905	910	
Leu	Ala	Ala	Leu	Ala	Tyr	His	Ala	Gly	Leu	Ser	Asp	Ser	Ala	Arg	Asp	915	920	925	
Glu	Val	Gln	Gln	Lys	Trp	Ile	Asn	Gln	Asp	Gly	Cys	Gln	Val	Ile	Cys	930	935	940	
Ala	Thr	Ile	Ala	Phe	Gly	Met	Gly	Ile	Asp	Lys	Pro	Asp	Val	Arg	Phe	945	950	955	960
Val	Ile	His	Ala	Ser	Leu	Pro	Lys	Ser	Val	Glu	Gly	Tyr	Tyr	Gln	Glu	965	970	975	

Ser Gly Arg Ala Gly Arg Asp Gly Glu Ile Ser His Cys Leu Leu Phe
 980 985 990
 Tyr Thr Tyr His Asp Val Thr Arg Leu Lys Arg Leu Ile Met Met Glu
 995 1000 1005
 Lys Asp Gly Asn His His Thr Arg Glu Thr His Phe Asn Asn Leu Tyr
 1010 1015 1020
 Ser Met Val His Tyr Cys Glu Asn Ile Thr Glu Cys Arg Arg Ile Gln
 1025 1030 1035 1040
 Leu Leu Ala Tyr Phe Gly Glu Asn Gly Phe Asn Pro Asp Phe Cys Lys
 1045 1050 1055
 Lys His Pro Asp Val Ser Cys Asp Asn Cys Cys Lys Thr Lys Asp Tyr
 1060 1065 1070
 Lys Thr Arg Asp Val Thr Asp Asp Val Lys Ser Ile Val Arg Phe Val
 1075 1080 1085
 Gln Glu His Ser Ser Ser Gln Gly Met Arg Asn Ile Lys His Val Gly
 1090 1095 1100
 Pro Ser Gly Arg Phe Thr Met Asn Met Leu Val Asp Ile Phe Leu Gly
 1105 1110 1115 1120
 Ser Lys Ser Ala Lys Ile Gln Ser Gly Ile Phe Gly Lys Gly Ser Ala
 1125 1130 1135
 Tyr Ser Arg His Asn Ala Glu Arg Leu Phe Lys Lys Leu Ile Leu Asp
 1140 1145 1150
 Lys Ile Leu Asp Glu Asp Leu Tyr Ile Asn Ala Asn Asp Gln Ala Ile
 1155 1160 1165
 Ala Tyr Val Met Leu Gly Asn Lys Ala Gln Thr Val Leu Asn Gly Asn
 1170 1175 1180
 Leu Lys Val Asp Phe Met Glu Thr Glu Asn Ser Ser Ser Val Lys Lys
 1185 1190 1195 1200
 Gln Lys Ala Leu Val Ala Lys Val Ser Gln Arg Glu Glu Met Val Lys
 1205 1210 1215
 Lys Cys Leu Gly Glu Leu Thr Glu Val Cys Lys Ser Leu Gly Lys Val
 1220 1225 1230
 Phe Gly Val His Tyr Phe Asn Ile Phe Asn Thr Val Thr Leu Lys Lys
 1235 1240 1245
 Leu Ala Glu Ser Leu Ser Ser Asp Pro Glu Val Leu Leu Gln Ile Asp
 1250 1255 1260
 Gly Val Thr Glu Asp Lys Leu Glu Lys Tyr Gly Ala Glu Val Ile Ser
 1265 1270 1275 1280
 Val Leu Gln Lys Tyr Ser Glu Trp Thr Ser Pro Ala Glu Asp Ser Ser
 1285 1290 1295
 Pro Gly Ile Ser Leu Ser Ser Ser Arg Gly Pro Gly Arg Ser Ala Ala
 1300 1305 1310
 Glu Glu Leu Asp Glu Glu Ile Pro Val Ser Ser His Tyr Phe Ala Ser
 1315 1320 1325
 Lys Thr Arg Asn Glu Arg Lys Arg Lys Lys Met Pro Ala Ser Gln Arg
 1330 1335 1340
 Ser Lys Arg Arg Lys Thr Ala Ser Ser Gly Ser Lys Ala Lys Gly Gly
 1345 1350 1355 1360
 Ser Ala Thr Cys Arg Lys Ile Ser Ser Lys Thr Lys Ser Ser Ser Ile
 1365 1370 1375
 Ile Gly Ser Ser Ser Ala Ser His Thr Ser Gln Ala Thr Ser Gly Ala
 1380 1385 1390
 Asn Ser Lys Leu Gly Ile Met Ala Pro Pro Lys Pro Ile Asn Arg Pro
 1395 1400 1405

Phe Leu Lys Pro Ser Tyr Ala Phe Ser
1410 1415

<210> 122
<211> 766
<212> PRT
<213> Homo sapiens

<220>
<223> >BRAFI|ENSG00000157764|ENST00000288602|2301

<400> 122
Met Ala Ala Leu Ser Gly Gly Gly Gly Gly Gly Ala Glu Pro Gly Gln
1 5 10 15
Ala Leu Phe Asn Gly Asp Met Glu Pro Glu Ala Gly Ala Gly
20 25 30
Ala Ala Ala Ser Ser Ala Ala Asp Pro Ala Ile Pro Glu Glu Val Trp
35 40 45
Asn Ile Lys Gln Met Ile Lys Leu Thr Gln Glu His Ile Glu Ala Leu
50 55 60
Leu Asp Lys Phe Gly Gly Glu His Asn Pro Pro Ser Ile Tyr Leu Glu
65 70 75 80
Ala Tyr Glu Glu Tyr Thr Ser Lys Leu Asp Ala Leu Gln Gln Arg Glu
85 90 95
Gln Gln Leu Leu Glu Ser Leu Gly Asn Gly Thr Asp Phe Ser Val Ser
100 105 110
Ser Ser Ala Ser Met Asp Thr Val Thr Ser Ser Ser Ser Ser Ser Leu
115 120 125
Ser Val Leu Pro Ser Ser Leu Ser Val Phe Gln Asn Pro Thr Asp Val
130 135 140
Ala Arg Ser Asn Pro Lys Ser Pro Gln Lys Pro Ile Val Arg Val Phe
145 150 155 160
Leu Pro Asn Lys Gln Arg Thr Val Val Pro Ala Arg Cys Gly Val Thr
165 170 175
Val Arg Asp Ser Leu Lys Lys Ala Leu Met Met Arg Gly Leu Ile Pro
180 185 190
Glu Cys Cys Ala Val Tyr Arg Ile Gln Asp Gly Glu Lys Lys Pro Ile
195 200 205
Gly Trp Asp Thr Asp Ile Ser Trp Leu Thr Gly Glu Glu Leu His Val
210 215 220
Glu Val Leu Glu Asn Val Pro Leu Thr Thr His Asn Phe Val Arg Lys
225 230 235 240
Thr Phe Phe Thr Leu Ala Phe Cys Asp Phe Cys Arg Lys Leu Leu Phe
245 250 255
Gln Gly Phe Arg Cys Gln Thr Cys Gly Tyr Lys Phe His Gln Arg Cys
260 265 270
Ser Thr Glu Val Pro Leu Met Cys Val Asn Tyr Asp Gln Leu Asp Leu
275 280 285
Leu Phe Val Ser Lys Phe Phe Glu His His Pro Ile Pro Gln Glu Glu
290 295 300
Ala Ser Leu Ala Glu Thr Ala Leu Thr Ser Gly Ser Ser Pro Ser Ala
305 310 315 320
Pro Ala Ser Asp Ser Ile Gly Pro Gln Ile Leu Thr Ser Pro Ser Pro

755

760

765

<210> 123

<211> 1884

<212> PRT

<213> Homo sapiens

<220>

<223> >BRCA1|ENSG00000012048|ENST00000471181|5655

<400> 123

```

Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1          5          10          15
Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
20          25          30
Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
35          40          45
Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
50          55          60
Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65          70          75          80
Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
85          90          95
Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
100         105         110
Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
115         120         125
Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
130         135         140
Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145         150         155         160
Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
165         170         175
Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
180         185         190
Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
195         200         205
Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
210         215         220
Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225         230         235         240
Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245         250         255
His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260         265         270
Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275         280         285
Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290         295         300
Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305         310         315         320
Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325         330         335

```

Glu	Lys	Lys	Val	Asp	Leu	Asn	Ala	Asp	Pro	Leu	Cys	Glu	Arg	Lys	Glu	
			340					345					350			
Trp	Asn	Lys	Gln	Lys	Leu	Pro	Cys	Ser	Glu	Asn	Pro	Arg	Asp	Thr	Glu	
		355					360					365				
Asp	Val	Pro	Trp	Ile	Thr	Leu	Asn	Ser	Ser	Ile	Gln	Lys	Val	Asn	Glu	
	370					375					380					
Trp	Phe	Ser	Arg	Ser	Asp	Glu	Leu	Leu	Gly	Ser	Asp	Asp	Ser	His	Asp	
385					390				395						400	
Gly	Glu	Ser	Glu	Ser	Asn	Ala	Lys	Val	Ala	Asp	Val	Leu	Asp	Val	Leu	
			405					410						415		
Asn	Glu	Val	Asp	Glu	Tyr	Ser	Gly	Ser	Ser	Glu	Lys	Ile	Asp	Leu	Leu	
			420					425					430			
Ala	Ser	Asp	Pro	His	Glu	Ala	Leu	Ile	Cys	Lys	Ser	Glu	Arg	Val	His	
		435					440					445				
Ser	Lys	Ser	Val	Glu	Ser	Asn	Ile	Glu	Asp	Lys	Ile	Phe	Gly	Lys	Thr	
	450					455					460					
Tyr	Arg	Lys	Lys	Ala	Ser	Leu	Pro	Asn	Leu	Ser	His	Val	Thr	Glu	Asn	
465					470					475					480	
Leu	Ile	Ile	Gly	Ala	Phe	Val	Thr	Glu	Pro	Gln	Ile	Ile	Gln	Glu	Arg	
			485					490						495		
Pro	Leu	Thr	Asn	Lys	Leu	Lys	Arg	Lys	Arg	Arg	Pro	Thr	Ser	Gly	Leu	
			500					505						510		
His	Pro	Glu	Asp	Phe	Ile	Lys	Lys	Ala	Asp	Leu	Ala	Val	Gln	Lys	Thr	
		515					520					525				
Pro	Glu	Met	Ile	Asn	Gln	Gly	Thr	Asn	Gln	Thr	Glu	Gln	Asn	Gly	Gln	
	530					535					540					
Val	Met	Asn	Ile	Thr	Asn	Ser	Gly	His	Glu	Asn	Lys	Thr	Lys	Gly	Asp	
545					550				555						560	
Ser	Ile	Gln	Asn	Glu	Lys	Asn	Pro	Asn	Pro	Ile	Glu	Ser	Leu	Glu	Lys	
				565				570						575		
Glu	Ser	Ala	Phe	Lys	Thr	Lys	Ala	Glu	Pro	Ile	Ser	Ser	Ser	Ile	Ser	
			580					585						590		
Asn	Met	Glu	Leu	Glu	Leu	Asn	Ile	His	Asn	Ser	Lys	Ala	Pro	Lys	Lys	
		595					600					605				
Asn	Arg	Leu	Arg	Arg	Lys	Ser	Ser	Thr	Arg	His	Ile	His	Ala	Leu	Glu	
	610					615					620					
Leu	Val	Val	Ser	Arg	Asn	Leu	Ser	Pro	Pro	Asn	Cys	Thr	Glu	Leu	Gln	
625					630					635					640	
Ile	Asp	Ser	Cys	Ser	Ser	Ser	Glu	Glu	Ile	Lys	Lys	Lys	Lys	Tyr	Asn	
			645						650					655		
Gln	Met	Pro	Val	Arg	His	Ser	Arg	Asn	Leu	Gln	Leu	Met	Glu	Gly	Lys	
			660					665					670			
Glu	Pro	Ala	Thr	Gly	Ala	Lys	Lys	Ser	Asn	Lys	Pro	Asn	Glu	Gln	Thr	
		675					680					685				
Ser	Lys	Arg	His	Asp	Ser	Asp	Thr	Phe	Pro	Glu	Leu	Lys	Leu	Thr	Asn	
	690					695					700					
Ala	Pro	Gly	Ser	Phe	Thr	Lys	Cys	Ser	Asn	Thr	Ser	Glu	Leu	Lys	Glu	
705					710					715					720	
Phe	Val	Asn	Pro	Ser	Leu	Pro	Arg	Glu	Glu	Lys	Glu	Glu	Lys	Leu	Glu	
				725					730					735		
Thr	Val	Lys	Val	Ser	Asn	Asn	Ala	Glu	Asp	Pro	Lys	Asp	Leu	Met	Leu	
			740					745					750			
Ser	Gly	Glu	Arg	Val	Leu	Gln	Thr	Glu	Arg	Ser	Val	Glu	Ser	Ser	Ser	
		755					760					765				

Ile	Ser	Leu	Val	Pro	Gly	Thr	Asp	Tyr	Gly	Thr	Gln	Glu	Ser	Ile	Ser	770	775	780
Leu	Leu	Glu	Val	Ser	Thr	Leu	Gly	Lys	Ala	Lys	Thr	Glu	Pro	Asn	Lys	785	790	795
Cys	Val	Ser	Gln	Cys	Ala	Ala	Phe	Glu	Asn	Pro	Lys	Gly	Leu	Ile	His	805	810	815
Gly	Cys	Ser	Lys	Asp	Asn	Arg	Asn	Asp	Thr	Glu	Gly	Phe	Lys	Tyr	Pro	820	825	830
Leu	Gly	His	Glu	Val	Asn	His	Ser	Arg	Glu	Thr	Ser	Ile	Glu	Met	Glu	835	840	845
Glu	Ser	Glu	Leu	Asp	Ala	Gln	Tyr	Leu	Gln	Asn	Thr	Phe	Lys	Val	Ser	850	855	860
Lys	Arg	Gln	Ser	Phe	Ala	Pro	Phe	Ser	Asn	Pro	Gly	Asn	Ala	Glu	Glu	865	870	875
Glu	Cys	Ala	Thr	Phe	Ser	Ala	His	Ser	Gly	Ser	Leu	Lys	Lys	Gln	Ser	885	890	895
Pro	Lys	Val	Thr	Phe	Glu	Cys	Glu	Gln	Lys	Glu	Glu	Asn	Gln	Gly	Lys	900	905	910
Asn	Glu	Ser	Asn	Ile	Lys	Pro	Val	Gln	Thr	Val	Asn	Ile	Thr	Ala	Gly	915	920	925
Phe	Pro	Val	Val	Gly	Gln	Lys	Asp	Lys	Pro	Val	Asp	Asn	Ala	Lys	Cys	930	935	940
Ser	Ile	Lys	Gly	Gly	Ser	Arg	Phe	Cys	Leu	Ser	Ser	Gln	Phe	Arg	Gly	945	950	955
Asn	Glu	Thr	Gly	Leu	Ile	Thr	Pro	Asn	Lys	His	Gly	Leu	Leu	Gln	Asn	965	970	975
Pro	Tyr	Arg	Ile	Pro	Pro	Leu	Phe	Pro	Ile	Lys	Ser	Phe	Val	Lys	Thr	980	985	990
Lys	Cys	Lys	Lys	Asn	Leu	Leu	Glu	Glu	Asn	Phe	Glu	Glu	His	Ser	Met	995	1000	1005
Ser	Pro	Glu	Arg	Glu	Met	Gly	Asn	Glu	Asn	Ile	Pro	Ser	Thr	Val	Ser	1010	1015	1020
Thr	Ile	Ser	Arg	Asn	Asn	Ile	Arg	Glu	Asn	Val	Phe	Lys	Glu	Ala	Ser	1025	1030	1035
Ser	Ser	Asn	Ile	Asn	Glu	Val	Gly	Ser	Ser	Thr	Asn	Glu	Val	Gly	Ser	1045	1050	1055
Ser	Ile	Asn	Glu	Ile	Gly	Ser	Ser	Asp	Glu	Asn	Ile	Gln	Ala	Glu	Leu	1060	1065	1070
Gly	Arg	Asn	Arg	Gly	Pro	Lys	Leu	Asn	Ala	Met	Leu	Arg	Leu	Gly	Val	1075	1080	1085
Leu	Gln	Pro	Glu	Val	Tyr	Lys	Gln	Ser	Leu	Pro	Gly	Ser	Asn	Cys	Lys	1090	1095	1100
His	Pro	Glu	Ile	Lys	Lys	Gln	Glu	Tyr	Glu	Glu	Val	Val	Gln	Thr	Val	1105	1110	1115
Asn	Thr	Asp	Phe	Ser	Pro	Tyr	Leu	Ile	Ser	Asp	Asn	Leu	Glu	Gln	Pro	1125	1130	1135
Met	Gly	Ser	Ser	His	Ala	Ser	Gln	Val	Cys	Ser	Glu	Thr	Pro	Asp	Asp	1140	1145	1150
Leu	Leu	Asp	Asp	Gly	Glu	Ile	Lys	Glu	Asp	Thr	Ser	Phe	Ala	Glu	Asn	1155	1160	1165
Asp	Ile	Lys	Glu	Ser	Ser	Ala	Val	Phe	Ser	Lys	Ser	Val	Gln	Lys	Gly	1170	1175	1180
Glu	Leu	Ser	Arg	Ser	Pro	Ser	Pro	Phe	Thr	His	Thr	His	Leu	Ala	Gln	1185	1190	1195
																		1200

Gly	Tyr	Arg	Arg	Gly	Ala	Lys	Lys	Leu	Glu	Ser	Ser	Glu	Glu	Asn	Leu		
				1205					1210					1215			
Ser	Ser	Glu	Asp	Glu	Glu	Leu	Pro	Cys	Phe	Gln	His	Leu	Leu	Phe	Gly		
			1220					1225						1230			
Lys	Val	Asn	Asn	Ile	Pro	Ser	Gln	Ser	Thr	Arg	His	Ser	Thr	Val	Ala		
		1235					1240						1245				
Thr	Glu	Cys	Leu	Ser	Lys	Asn	Thr	Glu	Glu	Asn	Leu	Leu	Ser	Leu	Lys		
	1250					1255					1260						
Asn	Ser	Leu	Asn	Asp	Cys	Ser	Asn	Gln	Val	Ile	Leu	Ala	Lys	Ala	Ser		
1265				1270					1275						1280		
Gln	Glu	His	His	Leu	Ser	Glu	Glu	Thr	Lys	Cys	Ser	Ala	Ser	Leu	Phe		
			1285					1290						1295			
Ser	Ser	Gln	Cys	Ser	Glu	Leu	Glu	Asp	Leu	Thr	Ala	Asn	Thr	Asn	Thr		
		1300					1305						1310				
Gln	Asp	Pro	Phe	Leu	Ile	Gly	Ser	Ser	Lys	Gln	Met	Arg	His	Gln	Ser		
	1315					1320					1325						
Glu	Ser	Gln	Gly	Val	Gly	Leu	Ser	Asp	Lys	Glu	Leu	Val	Ser	Asp	Asp		
	1330				1335					1340							
Glu	Glu	Arg	Gly	Thr	Gly	Leu	Glu	Glu	Asn	Asn	Gln	Glu	Glu	Gln	Ser		
1345				1350					1355					1360			
Met	Asp	Ser	Asn	Leu	Gly	Glu	Ala	Ala	Ser	Gly	Cys	Glu	Ser	Glu	Thr		
			1365					1370						1375			
Ser	Val	Ser	Glu	Asp	Cys	Ser	Gly	Leu	Ser	Ser	Gln	Ser	Asp	Ile	Leu		
	1380						1385						1390				
Thr	Thr	Gln	Gln	Arg	Asp	Thr	Met	Gln	His	Asn	Leu	Ile	Lys	Leu	Gln		
	1395					1400					1405						
Gln	Glu	Met	Ala	Glu	Leu	Glu	Ala	Val	Leu	Glu	Gln	His	Gly	Ser	Gln		
	1410				1415				1420								
Pro	Ser	Asn	Ser	Tyr	Pro	Ser	Ile	Ile	Ser	Asp	Ser	Ser	Ala	Leu	Glu		
1425				1430					1435					1440			
Asp	Leu	Arg	Asn	Pro	Glu	Gln	Ser	Thr	Ser	Glu	Lys	Asp	Ser	His	Ile		
			1445					1450						1455			
His	Gly	Gln	Arg	Asn	Asn	Ser	Met	Phe	Ser	Lys	Arg	Pro	Arg	Glu	His		
		1460					1465						1470				
Ile	Ser	Val	Leu	Thr	Ser	Gln	Lys	Ser	Ser	Glu	Tyr	Pro	Ile	Ser	Gln		
	1475					1480						1485					
Asn	Pro	Glu	Gly	Leu	Ser	Ala	Asp	Lys	Phe	Glu	Val	Ser	Ala	Asp	Ser		
	1490				1495					1500							
Ser	Thr	Ser	Lys	Asn	Lys	Glu	Pro	Gly	Val	Glu	Arg	Ser	Ser	Pro	Ser		
1505				1510				1515						1520			
Lys	Cys	Pro	Ser	Leu	Asp	Asp	Arg	Trp	Tyr	Met	His	Ser	Cys	Ser	Gly		
			1525				1530						1535				
Ser	Leu	Gln	Asn	Arg	Asn	Tyr	Pro	Ser	Gln	Glu	Glu	Leu	Ile	Lys	Val		
		1540				1545							1550				
Val	Asp	Val	Glu	Glu	Gln	Gln	Leu	Glu	Glu	Ser	Gly	Pro	His	Asp	Leu		
	1555					1560					1565						
Thr	Glu	Thr	Ser	Tyr	Leu	Pro	Arg	Gln	Asp	Leu	Glu	Gly	Thr	Pro	Tyr		
	1570				1575					1580							
Leu	Glu	Ser	Gly	Ile	Ser	Leu	Phe	Ser	Asp	Asp	Pro	Glu	Ser	Asp	Pro		
1585				1590					1595					1600			
Ser	Glu	Asp	Arg	Ala	Pro	Glu	Ser	Ala	Arg	Val	Gly	Asn	Ile	Pro	Ser		
		1605						1610					1615				
Ser	Thr	Ser	Ala	Leu	Lys	Val	Pro	Gln	Leu	Lys	Val	Ala	Glu	Ser	Ala		
		1620				1625						1630					

Gln Ser Pro Ala Ala Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala
 1635 1640 1645
 Met Glu Glu Ser Val Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr
 1650 1655 1660
 Glu Arg Val Asn Lys Arg Met Ser Met Val Val Ser Gly Leu Thr Pro
 1665 1670 1675 1680
 Glu Glu Phe Met Leu Val Tyr Lys Phe Ala Arg Lys His His Ile Thr
 1685 1690 1695
 Leu Thr Asn Leu Ile Thr Glu Glu Thr Thr His Val Val Met Lys Thr
 1700 1705 1710
 Asp Ala Glu Phe Val Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile
 1715 1720 1725
 Ala Gly Gly Lys Trp Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile
 1730 1735 1740
 Lys Glu Arg Lys Met Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp
 1745 1750 1755 1760
 Val Val Asn Gly Arg Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser
 1765 1770 1775
 Gln Asp Arg Lys Ile Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro
 1780 1785 1790
 Phe Thr Asn Met Pro Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys
 1795 1800 1805
 Gly Ala Ser Val Val Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly
 1810 1815 1820
 Val His Pro Ile Val Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn
 1825 1830 1835 1840
 Gly Phe His Ala Ile Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg
 1845 1850 1855
 Glu Trp Val Leu Asp Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp
 1860 1865 1870
 Thr Tyr Leu Ile Pro Gln Ile Pro His Ser His Tyr
 1875 1880

<210> 124

<211> 3418

<212> PRT

<213> Homo sapiens

<220>

<223> >BRCA2|ENSG00000139618|ENST00000544455|10257

<400> 124

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys
 1 5 10 15
 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe
 20 25 30
 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu
 35 40 45
 Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr
 50 55 60
 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile
 65 70 75 80
 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys

945					950					955				960
Gln	His	Ile	Lys	Met	Thr	Leu	Gly	Gln	Asp	Leu	Lys	Ser	Asp	Ile
				965					970					975
Leu	Asn	Ile	Asp	Lys	Ile	Pro	Glu	Lys	Asn	Asn	Asp	Tyr	Met	Asn
			980					985					990	
Trp	Ala	Gly	Leu	Leu	Gly	Pro	Ile	Ser	Asn	His	Ser	Phe	Gly	Gly
		995					1000					1005		
Phe	Arg	Thr	Ala	Ser	Asn	Lys	Glu	Ile	Lys	Leu	Ser	Glu	His	Asn
	1010					1015					1020			
Lys	Lys	Ser	Lys	Met	Phe	Phe	Lys	Asp	Ile	Glu	Glu	Gln	Tyr	Pro
1025				1030					1035					1040
Ser	Leu	Ala	Cys	Val	Glu	Ile	Val	Asn	Thr	Leu	Ala	Leu	Asp	Asn
			1045					1050					1055	
Lys	Lys	Leu	Ser	Lys	Pro	Gln	Ser	Ile	Asn	Thr	Val	Ser	Ala	His
		1060					1065						1070	
Gln	Ser	Ser	Val	Val	Val	Ser	Asp	Cys	Lys	Asn	Ser	His	Ile	Thr
	1075					1080					1085			
Gln	Met	Leu	Phe	Ser	Lys	Gln	Asp	Phe	Asn	Ser	Asn	His	Asn	Leu
	1090					1095					1100			
Pro	Ser	Gln	Lys	Ala	Glu	Ile	Thr	Glu	Leu	Ser	Thr	Ile	Leu	Glu
1105				1110					1115					1120
Ser	Gly	Ser	Gln	Phe	Glu	Phe	Thr	Gln	Phe	Arg	Lys	Pro	Ser	Tyr
			1125					1130						1135
Leu	Gln	Lys	Ser	Thr	Phe	Glu	Val	Pro	Glu	Asn	Gln	Met	Thr	Ile
		1140					1145					1150		
Lys	Thr	Thr	Ser	Glu	Glu	Cys	Arg	Asp	Ala	Asp	Leu	His	Val	Ile
	1155					1160					1165			
Asn	Ala	Pro	Ser	Ile	Gly	Gln	Val	Asp	Ser	Ser	Lys	Gln	Phe	Glu
	1170				1175						1180			
Thr	Val	Glu	Ile	Lys	Arg	Lys	Phe	Ala	Gly	Leu	Leu	Lys	Asn	Asp
1185				1190					1195					1200
Asn	Lys	Ser	Ala	Ser	Gly	Tyr	Leu	Thr	Asp	Glu	Asn	Glu	Val	Gly
			1205					1210						1215
Arg	Gly	Phe	Tyr	Ser	Ala	His	Gly	Thr	Lys	Leu	Asn	Val	Ser	Thr
		1220					1225						1230	
Ala	Leu	Gln	Lys	Ala	Val	Lys	Leu	Phe	Ser	Asp	Ile	Glu	Asn	Ile
	1235					1240					1245			
Glu	Glu	Thr	Ser	Ala	Glu	Val	His	Pro	Ile	Ser	Leu	Ser	Ser	Ser
1250				1255						1260				
Cys	His	Asp	Ser	Val	Val	Ser	Met	Phe	Lys	Ile	Glu	Asn	His	Asn
1265				1270					1275					1280
Lys	Thr	Val	Ser	Glu	Lys	Asn	Asn	Lys	Cys	Gln	Leu	Ile	Leu	Gln
			1285					1290					1295	
Asn	Ile	Glu	Met	Thr	Thr	Gly	Thr	Phe	Val	Glu	Glu	Ile	Thr	Glu
	1300					1305						1310		
Tyr	Lys	Arg	Asn	Thr	Glu	Asn	Glu	Asp	Asn	Lys	Tyr	Thr	Ala	Ala
	1315					1320					1325			
Arg	Asn	Ser	His	Asn	Leu	Glu	Phe	Asp	Gly	Ser	Asp	Ser	Ser	Lys
	1330				1335				1340					
Asp	Thr	Val	Cys	Ile	His	Lys	Asp	Glu	Thr	Asp	Leu	Leu	Phe	Thr
1345				1350					1355					1360
Gln	His	Asn	Ile	Cys	Leu	Lys	Leu	Ser	Gly	Gln	Phe	Met	Lys	Glu
			1365					1370					1375	
Asn	Thr	Gln	Ile	Lys	Glu	Asp	Leu	Ser	Asp	Leu	Thr	Phe	Leu	Glu

1810	1815	1820
Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly		
1825	1830	1835
Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His		1840
	1845	1850
Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys		1855
	1860	1865
Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys		1870
	1875	1880
Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu		1885
	1890	1895
His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val		1900
1905	1910	1915
Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met		1920
	1925	1930
Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu		1935
	1940	1945
Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser		1950
	1955	1960
Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys		1965
	1970	1975
Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe		1980
1985	1990	1995
Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe		2000
	2005	2010
Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala		2015
	2020	2025
Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn		2030
	2035	2040
Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys		2045
	2050	2055
Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu		2060
2065	2070	2075
Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro		2080
	2085	2090
Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg		2095
	2100	2105
Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys		2110
	2115	2120
Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu		2125
	2130	2135
Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln		2140
2145	2150	2155
Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn		2160
	2165	2170
Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met		2175
	2180	2185
Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn		2190
	2195	2200
Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu		2205
	2210	2215
Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu		2220
2225	2230	2235
Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys		2240

				2245				2250				2255			
Pro	Glu	Asn	Glu	Glu	Met	Val	Leu	Ser	Asn	Ser	Arg	Ile	Gly	Lys	Arg
			2260					2265					2270		
Arg	Gly	Glu	Pro	Leu	Ile	Leu	Val	Gly	Glu	Pro	Ser	Ile	Lys	Arg	Asn
		2275					2280					2285			
Leu	Leu	Asn	Glu	Phe	Asp	Arg	Ile	Ile	Glu	Asn	Gln	Glu	Lys	Ser	Leu
	2290					2295				2300					
Lys	Ala	Ser	Lys	Ser	Thr	Pro	Asp	Gly	Thr	Ile	Lys	Asp	Arg	Arg	Leu
2305				2310					2315						2320
Phe	Met	His	His	Val	Ser	Leu	Glu	Pro	Ile	Thr	Cys	Val	Pro	Phe	Arg
			2325					2330					2335		
Thr	Thr	Lys	Glu	Arg	Gln	Glu	Ile	Gln	Asn	Pro	Asn	Phe	Thr	Ala	Pro
		2340					2345					2350			
Gly	Gln	Glu	Phe	Leu	Ser	Lys	Ser	His	Leu	Tyr	Glu	His	Leu	Thr	Leu
	2355					2360					2365				
Glu	Lys	Ser	Ser	Ser	Asn	Leu	Ala	Val	Ser	Gly	His	Pro	Phe	Tyr	Gln
	2370				2375					2380					
Val	Ser	Ala	Thr	Arg	Asn	Glu	Lys	Met	Arg	His	Leu	Ile	Thr	Thr	Gly
2385				2390					2395						2400
Arg	Pro	Thr	Lys	Val	Phe	Val	Pro	Pro	Phe	Lys	Thr	Lys	Ser	His	Phe
			2405					2410						2415	
His	Arg	Val	Glu	Gln	Cys	Val	Arg	Asn	Ile	Asn	Leu	Glu	Glu	Asn	Arg
		2420					2425					2430			
Gln	Lys	Gln	Asn	Ile	Asp	Gly	His	Gly	Ser	Asp	Asp	Ser	Lys	Asn	Lys
	2435				2440					2445					
Ile	Asn	Asp	Asn	Glu	Ile	His	Gln	Phe	Asn	Lys	Asn	Asn	Ser	Asn	Gln
	2450				2455					2460					
Ala	Val	Ala	Val	Thr	Phe	Thr	Lys	Cys	Glu	Glu	Glu	Pro	Leu	Asp	Leu
2465				2470					2475						2480
Ile	Thr	Ser	Leu	Gln	Asn	Ala	Arg	Asp	Ile	Gln	Asp	Met	Arg	Ile	Lys
			2485					2490					2495		
Lys	Lys	Gln	Arg	Gln	Arg	Val	Phe	Pro	Gln	Pro	Gly	Ser	Leu	Tyr	Leu
		2500					2505					2510			
Ala	Lys	Thr	Ser	Thr	Leu	Pro	Arg	Ile	Ser	Leu	Lys	Ala	Ala	Val	Gly
	2515						2520				2525				
Gly	Gln	Val	Pro	Ser	Ala	Cys	Ser	His	Lys	Gln	Leu	Tyr	Thr	Tyr	Gly
	2530				2535					2540					
Val	Ser	Lys	His	Cys	Ile	Lys	Ile	Asn	Ser	Lys	Asn	Ala	Glu	Ser	Phe
2545				2550					2555						2560
Gln	Phe	His	Thr	Glu	Asp	Tyr	Phe	Gly	Lys	Glu	Ser	Leu	Trp	Thr	Gly
			2565					2570					2575		
Lys	Gly	Ile	Gln	Leu	Ala	Asp	Gly	Gly	Trp	Leu	Ile	Pro	Ser	Asn	Asp
		2580					2585					2590			
Gly	Lys	Ala	Gly	Lys	Glu	Glu	Phe	Tyr	Arg	Ala	Leu	Cys	Asp	Thr	Pro
	2595					2600					2605				
Gly	Val	Asp	Pro	Lys	Leu	Ile	Ser	Arg	Ile	Trp	Val	Tyr	Asn	His	Tyr
	2610				2615					2620					
Arg	Trp	Ile	Ile	Trp	Lys	Leu	Ala	Ala	Met	Glu	Cys	Ala	Phe	Pro	Lys
2625				2630					2635						2640
Glu	Phe	Ala	Asn	Arg	Cys	Leu	Ser	Pro	Glu	Arg	Val	Leu	Leu	Gln	Leu
			2645					2650					2655		
Lys	Tyr	Arg	Tyr	Asp	Thr	Glu	Ile	Asp	Arg	Ser	Arg	Arg	Ser	Ala	Ile
		2660					2665				2670				
Lys	Lys	Ile	Met	Glu	Arg	Asp	Asp	Thr	Ala	Ala	Lys	Thr	Leu	Val	Leu

3105					3110					3115				3120
Ala	Ala	Ser	Asn	Leu	Gln	Trp	Arg	Pro	Glu	Ser	Lys	Ser	Gly	Leu
					3125					3130				3135
Thr	Leu	Phe	Ala	Gly	Asp	Phe	Ser	Val	Phe	Ser	Ala	Ser	Pro	Lys
					3140					3145				3150
Gly	His	Phe	Gln	Glu	Thr	Phe	Asn	Lys	Met	Lys	Asn	Thr	Val	Glu
					3155					3160				3165
Ile	Asp	Ile	Leu	Cys	Asn	Glu	Ala	Glu	Asn	Lys	Leu	Met	His	Ile
					3170					3175				3180
His	Ala	Asn	Asp	Pro	Lys	Trp	Ser	Thr	Pro	Thr	Lys	Asp	Cys	Thr
					3185					3190				3195
Gly	Pro	Tyr	Thr	Ala	Gln	Ile	Ile	Pro	Gly	Thr	Gly	Asn	Lys	Leu
					3205					3210				3215
Met	Ser	Ser	Pro	Asn	Cys	Glu	Ile	Tyr	Tyr	Gln	Ser	Pro	Leu	Ser
					3220					3225				3230
Cys	Met	Ala	Lys	Arg	Lys	Ser	Val	Ser	Thr	Pro	Val	Ser	Ala	Gln
					3235					3240				3245
Thr	Ser	Lys	Ser	Cys	Lys	Gly	Glu	Lys	Glu	Ile	Asp	Asp	Gln	Lys
					3250					3255				3260
Cys	Lys	Lys	Arg	Arg	Ala	Leu	Asp	Phe	Leu	Ser	Arg	Leu	Pro	Leu
					3265					3270				3275
Pro	Pro	Val	Ser	Pro	Ile	Cys	Thr	Phe	Val	Ser	Pro	Ala	Ala	Gln
					3285					3290				3295
Ala	Phe	Gln	Pro	Pro	Arg	Ser	Cys	Gly	Thr	Lys	Tyr	Glu	Thr	Pro
					3300					3305				3310
Lys	Lys	Lys	Glu	Leu	Asn	Ser	Pro	Gln	Met	Thr	Pro	Phe	Lys	Lys
					3315					3320				3325
Asn	Glu	Ile	Ser	Leu	Leu	Glu	Ser	Asn	Ser	Ile	Ala	Asp	Glu	Glu
					3330					3335				3340
Ala	Leu	Ile	Asn	Thr	Gln	Ala	Leu	Leu	Ser	Gly	Ser	Thr	Gly	Glu
					3345					3350				3355
Gln	Phe	Ile	Ser	Val	Ser	Glu	Ser	Thr	Arg	Thr	Ala	Pro	Thr	Ser
					3365					3370				3375
Glu	Asp	Tyr	Leu	Arg	Leu	Lys	Arg	Arg	Cys	Thr	Thr	Ser	Leu	Ile
					3380					3385				3390
Glu	Gln	Glu	Ser	Ser	Gln	Ala	Ser	Thr	Glu	Glu	Cys	Glu	Lys	Asn
					3395					3400				3405
Gln	Asp	Thr	Ile	Thr	Thr	Lys	Lys	Tyr	Ile					
					3410				3415					

<210> 125

<211> 1249

<212> PRT

<213> Homo sapiens

<220>

<223> >BRIP1|ENSG00000136492|ENST00000259008|3750

<400> 125

Met	Ser	Ser	Met	Trp	Ser	Glu	Tyr	Thr	Ile	Gly	Gly	Val	Lys	Ile	Tyr
1				5					10					15	
Phe	Pro	Tyr	Lys	Ala	Tyr	Pro	Ser	Gln	Leu	Ala	Met	Met	Asn	Ser	Ile
				20				25					30		

Leu	Arg	Gly	Leu	Asn	Ser	Lys	Gln	His	Cys	Leu	Leu	Glu	Ser	Pro	Thr		
		35					40					45					
Gly	Ser	Gly	Lys	Ser	Leu	Ala	Leu	Leu	Cys	Ser	Ala	Leu	Ala	Trp	Gln		
	50					55					60						
Gln	Ser	Leu	Ser	Gly	Lys	Pro	Ala	Asp	Glu	Gly	Val	Ser	Glu	Lys	Ala		
65				70						75					80		
Glu	Val	Gln	Leu	Ser	Cys	Cys	Cys	Ala	Cys	His	Ser	Lys	Asp	Phe	Thr		
			85						90					95			
Asn	Asn	Asp	Met	Asn	Gln	Gly	Thr	Ser	Arg	His	Phe	Asn	Tyr	Pro	Ser		
			100					105					110				
Thr	Pro	Pro	Ser	Glu	Arg	Asn	Gly	Thr	Ser	Ser	Thr	Cys	Gln	Asp	Ser		
		115					120					125					
Pro	Glu	Lys	Thr	Thr	Leu	Ala	Ala	Lys	Leu	Ser	Ala	Lys	Lys	Gln	Ala		
		130				135					140						
Ser	Ile	Tyr	Arg	Asp	Glu	Asn	Asp	Asp	Phe	Gln	Val	Glu	Lys	Lys	Arg		
145				150						155					160		
Ile	Arg	Pro	Leu	Glu	Thr	Thr	Gln	Gln	Ile	Arg	Lys	Arg	His	Cys	Phe		
			165						170					175			
Gly	Thr	Glu	Val	His	Asn	Leu	Asp	Ala	Lys	Val	Asp	Ser	Gly	Lys	Thr		
			180					185					190				
Val	Lys	Leu	Asn	Ser	Pro	Leu	Glu	Lys	Ile	Asn	Ser	Phe	Ser	Pro	Gln		
		195					200					205					
Lys	Pro	Pro	Gly	His	Cys	Ser	Arg	Cys	Cys	Cys	Ser	Thr	Lys	Gln	Gly		
	210					215					220						
Asn	Ser	Gln	Glu	Ser	Ser	Asn	Thr	Ile	Lys	Lys	Asp	His	Thr	Gly	Lys		
225				230					235					240			
Ser	Lys	Ile	Pro	Lys	Ile	Tyr	Phe	Gly	Thr	Arg	Thr	His	Lys	Gln	Ile		
			245					250					255				
Ala	Gln	Ile	Thr	Arg	Glu	Leu	Arg	Arg	Thr	Ala	Tyr	Ser	Gly	Val	Pro		
			260					265					270				
Met	Thr	Ile	Leu	Ser	Ser	Arg	Asp	His	Thr	Cys	Val	His	Pro	Glu	Val		
		275					280					285					
Val	Gly	Asn	Phe	Asn	Arg	Asn	Glu	Lys	Cys	Met	Glu	Leu	Leu	Asp	Gly		
	290					295					300						
Lys	Asn	Gly	Lys	Ser	Cys	Tyr	Phe	Tyr	His	Gly	Val	His	Lys	Ile	Ser		
305					310					315					320		
Asp	Gln	His	Thr	Leu	Gln	Thr	Phe	Gln	Gly	Met	Cys	Lys	Ala	Trp	Asp		
			325						330					335			
Ile	Glu	Glu	Leu	Val	Ser	Leu	Gly	Lys	Lys	Leu	Lys	Ala	Cys	Pro	Tyr		
			340				345						350				
Tyr	Thr	Ala	Arg	Glu	Leu	Ile	Gln	Asp	Ala	Asp	Ile	Ile	Phe	Cys	Pro		
		355					360						365				
Tyr	Asn	Tyr	Leu	Leu	Asp	Ala	Gln	Ile	Arg	Glu	Ser	Met	Asp	Leu	Asn		
					375						380						
Leu	Lys	Glu	Gln	Val	Val	Ile	Leu	Asp	Glu	Ala	His	Asn	Ile	Glu	Asp		
385				390						395				400			
Cys	Ala	Arg	Glu	Ser	Ala	Ser	Tyr	Ser	Val	Thr	Glu	Val	Gln	Leu	Arg		
			405						410					415			
Phe	Ala	Arg	Asp	Glu	Leu	Asp	Ser	Met	Val	Asn	Asn	Asn	Ile	Arg	Lys		
			420					425					430				
Lys	Asp	His	Glu	Pro	Leu	Arg	Ala	Val	Cys	Cys	Ser	Leu	Ile	Asn	Trp		
		435					440					445					
Leu	Glu	Ala	Asn	Ala	Glu	Tyr	Leu	Val	Glu	Arg	Asp	Tyr	Glu	Ser	Ala		
	450					455					460						

Cys 465	Lys	Ile	Trp	Ser	Gly 470	Asn	Glu	Met	Leu	Leu	Thr	Leu	His	Lys	Met 480
Gly	Ile	Thr	Thr	Ala 485	Thr	Phe	Pro	Ile	Leu	Gln	Gly	His	Phe	Ser	Ala 495
Val	Leu	Gln	Lys 500	Glu	Glu	Lys	Ile	Ser 505	Pro	Ile	Tyr	Gly	Lys	Glu	Glu
Ala	Arg	Glu	Val	Pro	Val	Ile	Ser 520	Ala	Ser	Thr	Gln	Ile	Met	Leu	Lys
Gly	Leu	Phe	Met	Val	Leu	Asp 535	Tyr	Leu	Phe	Arg	Gln	Asn	Ser	Arg	Phe
Ala 545	Asp	Asp	Tyr	Lys	Ile	Ala	Ile	Gln	Gln	Thr	Tyr	Ser	Trp	Thr	Asn 560
Gln	Ile	Asp	Ile	Ser 565	Asp	Lys	Asn	Gly	Leu	Leu	Val	Leu	Pro	Lys	Asn 575
Lys	Lys	Arg	Ser 580	Arg	Gln	Lys	Thr	Ala 585	Val	His	Val	Leu	Asn	Phe	Trp
Cys	Leu	Asn	Pro	Ala	Val	Ala	Phe 600	Ser	Asp	Ile	Asn	Gly	Lys	Val	Gln
Thr	Ile	Val	Leu	Thr	Ser	Gly 615	Thr	Leu	Ser	Pro	Met	Lys	Ser	Phe	Ser
Ser 625	Glu	Leu	Gly	Val	Thr	Phe	Thr	Ile	Gln	Leu	Glu	Ala	Asn	His	Ile 640
Ile	Lys	Asn	Ser	Gln 645	Val	Trp	Val	Gly	Thr	Ile	Gly	Ser	Gly	Pro	Lys 655
Gly	Arg	Asn	Leu	Cys 660	Ala	Thr	Phe	Gln	Asn	Thr	Glu	Thr	Phe	Glu	Phe
Gln	Asp	Glu	Val	Gly	Ala	Leu	Leu	Leu	Ser	Val	Cys	Gln	Thr	Val	Ser
Gln	Gly	Ile	Leu	Cys	Phe	Leu	Pro	Ser	Tyr	Lys	Leu	Leu	Glu	Lys	Leu
Lys 705	Glu	Arg	Trp	Leu	Ser	Thr	Gly	Leu	Trp	His	Asn	Leu	Glu	Leu	Val 720
Lys	Thr	Val	Ile	Val	Glu	Pro	Gln	Gly	Gly	Glu	Lys	Thr	Asn	Phe	Asp
Glu	Leu	Leu	Gln	Val	Tyr	Tyr	Asp	Ala 745	Ile	Lys	Tyr	Lys	Gly	Glu	Lys
Asp	Gly	Ala	Leu	Leu	Val	Ala	Val	Cys 760	Arg	Gly	Lys	Val	Ser	Glu	Gly
Leu	Asp	Phe	Ser	Asp	Asp	Asn	Ala	Arg	Ala	Val	Ile	Thr	Ile	Gly	Ile
Pro	Phe	Pro	Asn	Val	Lys	Asp	Leu	Gln	Val	Glu	Leu	Lys	Arg	Gln	Tyr
Asn	Asp	His	His	Ser	Lys	Leu	Arg	Gly	Leu	Leu	Pro	Gly	Arg	Gln	Trp
Tyr	Glu	Ile	Gln	Ala	Tyr	Arg	Ala	Leu	Asn	Gln	Ala	Leu	Gly	Arg	Cys
Ile	Arg	His	Arg	Asn	Asp	Trp	Gly	Ala	Leu	Ile	Leu	Val	Asp	Asp	Arg
Phe	Arg	Asn	Asn	Pro	Ser	Arg	Tyr	Ile	Ser	Gly	Leu	Ser	Lys	Trp	Val
Arg 865	Gln	Gln	Ile	Gln	His	His	Ser	Thr	Phe	Glu	Ser	Ala	Leu	Glu	Ser 880
Leu	Ala	Glu	Phe	Ser	Lys	Lys	His	Gln	Lys	Val	Leu	Asn	Val	Ser	Ile

Lys Asp Arg Thr Asn Ile Gln Asp Asn Glu Ser Thr Leu Glu Val Thr
 900 905 910
 Ser Leu Lys Tyr Ser Thr Ser Pro Tyr Leu Leu Glu Ala Ala Ser His
 915 920 925
 Leu Ser Pro Glu Asn Phe Val Glu Asp Glu Ala Lys Ile Cys Val Gln
 930 935 940
 Glu Leu Gln Cys Pro Lys Ile Ile Thr Lys Asn Ser Pro Leu Pro Ser
 945 950 955 960
 Ser Ile Ile Ser Arg Lys Glu Lys Asn Asp Pro Val Phe Leu Glu Glu
 965 970 975
 Ala Gly Lys Ala Glu Lys Ile Val Ile Ser Arg Ser Thr Ser Pro Thr
 980 985 990
 Phe Asn Lys Gln Thr Lys Arg Val Ser Trp Ser Ser Phe Asn Ser Leu
 995 1000 1005
 Gly Gln Tyr Phe Thr Gly Lys Ile Pro Lys Ala Thr Pro Glu Leu Gly
 1010 1015 1020
 Ser Ser Glu Asn Ser Ala Ser Ser Pro Pro Arg Phe Lys Thr Glu Lys
 1025 1030 1035 1040
 Met Glu Ser Lys Thr Val Leu Pro Phe Thr Asp Lys Cys Glu Ser Ser
 1045 1050 1055
 Asn Leu Thr Val Asn Thr Ser Phe Gly Ser Cys Pro Gln Ser Glu Thr
 1060 1065 1070
 Ile Ile Ser Ser Leu Lys Ile Asp Ala Thr Leu Thr Arg Lys Asn His
 1075 1080 1085
 Ser Glu His Pro Leu Cys Ser Glu Glu Ala Leu Asp Pro Asp Ile Glu
 1090 1095 1100
 Leu Ser Leu Val Ser Glu Glu Asp Lys Gln Ser Thr Ser Asn Arg Asp
 1105 1110 1115 1120
 Phe Glu Thr Glu Ala Glu Asp Glu Ser Ile Tyr Phe Thr Pro Glu Leu
 1125 1130 1135
 Tyr Asp Pro Glu Asp Thr Asp Glu Glu Lys Asn Asp Leu Ala Glu Thr
 1140 1145 1150
 Asp Arg Gly Asn Arg Leu Ala Asn Asn Ser Asp Cys Ile Leu Ala Lys
 1155 1160 1165
 Asp Leu Phe Glu Ile Arg Thr Ile Lys Glu Val Asp Ser Ala Arg Glu
 1170 1175 1180
 Val Lys Ala Glu Asp Cys Ile Asp Thr Lys Leu Asn Gly Ile Leu His
 1185 1190 1195 1200
 Ile Glu Glu Ser Lys Ile Asp Asp Ile Asp Gly Asn Val Lys Thr Thr
 1205 1210 1215
 Trp Ile Asn Glu Leu Glu Leu Gly Lys Thr His Glu Ile Glu Ile Lys
 1220 1225 1230
 Asn Phe Lys Pro Ser Pro Ser Lys Asn Lys Gly Met Phe Pro Gly Phe
 1235 1240 1245
 Lys

<210> 126

<211> 295

<212> PRT

<213> Homo sapiens

<220>

<223> >CCND1|ENSG00000110092|ENST00000227507|888

<400> 126

Met	Glu	His	Gln	Leu	Leu	Cys	Cys	Glu	Val	Glu	Thr	Ile	Arg	Arg	Ala	
1				5					10						15	
Tyr	Pro	Asp	Ala	Asn	Leu	Leu	Asn	Asp	Arg	Val	Leu	Arg	Ala	Met	Leu	
			20					25						30		
Lys	Ala	Glu	Glu	Thr	Cys	Ala	Pro	Ser	Val	Ser	Tyr	Phe	Lys	Cys	Val	
		35					40					45				
Gln	Lys	Glu	Val	Leu	Pro	Ser	Met	Arg	Lys	Ile	Val	Ala	Thr	Trp	Met	
	50					55					60					
Leu	Glu	Val	Cys	Glu	Glu	Gln	Lys	Cys	Glu	Glu	Glu	Val	Phe	Pro	Leu	
65					70					75					80	
Ala	Met	Asn	Tyr	Leu	Asp	Arg	Phe	Leu	Ser	Leu	Glu	Pro	Val	Lys	Lys	
				85					90					95		
Ser	Arg	Leu	Gln	Leu	Leu	Gly	Ala	Thr	Cys	Met	Phe	Val	Ala	Ser	Lys	
			100					105					110			
Met	Lys	Glu	Thr	Ile	Pro	Leu	Thr	Ala	Glu	Lys	Leu	Cys	Ile	Tyr	Thr	
		115					120						125			
Asp	Asn	Ser	Ile	Arg	Pro	Glu	Glu	Leu	Leu	Gln	Met	Glu	Leu	Leu	Leu	
	130					135					140					
Val	Asn	Lys	Leu	Lys	Trp	Asn	Leu	Ala	Ala	Met	Thr	Pro	His	Asp	Phe	
145					150					155					160	
Ile	Glu	His	Phe	Leu	Ser	Lys	Met	Pro	Glu	Ala	Glu	Glu	Asn	Lys	Gln	
				165					170					175		
Ile	Ile	Arg	Lys	His	Ala	Gln	Thr	Phe	Val	Ala	Leu	Cys	Ala	Thr	Asp	
			180					185					190			
Val	Lys	Phe	Ile	Ser	Asn	Pro	Pro	Ser	Met	Val	Ala	Ala	Gly	Ser	Val	
		195					200					205				
Val	Ala	Ala	Val	Gln	Gly	Leu	Asn	Leu	Arg	Ser	Pro	Asn	Asn	Phe	Leu	
	210					215					220					
Ser	Tyr	Tyr	Arg	Leu	Thr	Arg	Phe	Leu	Ser	Arg	Val	Ile	Lys	Cys	Asp	
225					230					235					240	
Pro	Asp	Cys	Leu	Arg	Ala	Cys	Gln	Glu	Gln	Ile	Glu	Ala	Leu	Leu	Glu	
				245					250					255		
Ser	Ser	Leu	Arg	Gln	Ala	Gln	Gln	Asn	Met	Asp	Pro	Lys	Ala	Ala	Glu	
			260					265					270			
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Val	Asp	Leu	Ala	Cys	Thr	Pro	Thr	
		275					280					285				
Asp	Val	Arg	Asp	Val	Asp	Ile										
	290					295										

<210> 127

<211> 410

<212> PRT

<213> Homo sapiens

<220>

<223> >CCNE1|ENSG00000105173|ENST00000262643|1233

<400> 127

Met	Pro	Arg	Glu	Arg	Arg	Glu	Arg	Asp	Ala	Lys	Glu	Arg	Asp	Thr	Met	
1				5					10					15		

Lys	Glu	Asp	Gly	Gly	Ala	Glu	Phe	Ser	Ala	Arg	Ser	Arg	Lys	Arg	Lys			
			20					25					30					
Ala	Asn	Val	Thr	Val	Phe	Leu	Gln	Asp	Pro	Asp	Glu	Glu	Met	Ala	Lys			
		35					40					45						
Ile	Asp	Arg	Thr	Ala	Arg	Asp	Gln	Cys	Gly	Ser	Gln	Pro	Trp	Asp	Asn			
	50					55					60							
Asn	Ala	Val	Cys	Ala	Asp	Pro	Cys	Ser	Leu	Ile	Pro	Thr	Pro	Asp	Lys			
65					70					75					80			
Glu	Asp	Asp	Asp	Arg	Val	Tyr	Pro	Asn	Ser	Thr	Cys	Lys	Pro	Arg	Ile			
				85				90						95				
Ile	Ala	Pro	Ser	Arg	Gly	Ser	Pro	Leu	Pro	Val	Leu	Ser	Trp	Ala	Asn			
			100					105					110					
Arg	Glu	Glu	Val	Trp	Lys	Ile	Met	Leu	Asn	Lys	Glu	Lys	Thr	Tyr	Leu			
		115					120					125						
Arg	Asp	Gln	His	Phe	Leu	Glu	Gln	His	Pro	Leu	Leu	Gln	Pro	Lys	Met			
	130					135					140							
Arg	Ala	Ile	Leu	Leu	Asp	Trp	Leu	Met	Glu	Val	Cys	Glu	Val	Tyr	Lys			
145					150					155					160			
Leu	His	Arg	Glu	Thr	Phe	Tyr	Leu	Ala	Gln	Asp	Phe	Phe	Asp	Arg	Tyr			
				165					170					175				
Met	Ala	Thr	Gln	Glu	Asn	Val	Val	Lys	Thr	Leu	Leu	Gln	Leu	Ile	Gly			
			180					185					190					
Ile	Ser	Ser	Leu	Phe	Ile	Ala	Ala	Lys	Leu	Glu	Glu	Ile	Tyr	Pro	Pro			
	195					200						205						
Lys	Leu	His	Gln	Phe	Ala	Tyr	Val	Thr	Asp	Gly	Ala	Cys	Ser	Gly	Asp			
	210					215				220								
Glu	Ile	Leu	Thr	Met	Glu	Leu	Met	Ile	Met	Lys	Ala	Leu	Lys	Trp	Arg			
225					230				235						240			
Leu	Ser	Pro	Leu	Thr	Ile	Val	Ser	Trp	Leu	Asn	Val	Tyr	Met	Gln	Val			
				245				250						255				
Ala	Tyr	Leu	Asn	Asp	Leu	His	Glu	Val	Leu	Leu	Pro	Gln	Tyr	Pro	Gln			
			260				265						270					
Gln	Ile	Phe	Ile	Gln	Ile	Ala	Glu	Leu	Leu	Asp	Leu	Cys	Val	Leu	Asp			
		275					280					285						
Val	Asp	Cys	Leu	Glu	Phe	Pro	Tyr	Gly	Ile	Leu	Ala	Ala	Ser	Ala	Leu			
	290					295					300							
Tyr	His	Phe	Ser	Ser	Ser	Glu	Leu	Met	Gln	Lys	Val	Ser	Gly	Tyr	Gln			
305					310					315					320			
Trp	Cys	Asp	Ile	Glu	Asn	Cys	Val	Lys	Trp	Met	Val	Pro	Phe	Ala	Met			
				325				330						335				
Val	Ile	Arg	Glu	Thr	Gly	Ser	Ser	Lys	Leu	Lys	His	Phe	Arg	Gly	Val			
			340					345					350					
Ala	Asp	Glu	Asp	Ala	His	Asn	Ile	Gln	Thr	His	Arg	Asp	Ser	Leu	Asp			
		355				360						365						
Leu	Leu	Asp	Lys	Ala	Arg	Ala	Lys	Lys	Ala	Met	Leu	Ser	Glu	Gln	Asn			
	370					375					380							
Arg	Ala	Ser	Pro	Leu	Pro	Ser	Gly	Leu	Leu	Thr	Pro	Pro	Gln	Ser	Gly			
385					390					395					400			
Lys	Lys	Gln	Ser	Ser	Gly	Pro	Glu	Met	Ala									
				405					410									

<210> 128

<211> 404

<212> PRT

<213> Homo sapiens

<220>

<223> >CCNE2|ENSG00000175305|ENST00000520509|1215

<400> 128

Met	Ser	Arg	Arg	Ser	Ser	Arg	Leu	Gln	Ala	Lys	Gln	Gln	Pro	Gln	Pro	
1				5					10					15		
Ser	Gln	Thr	Glu	Ser	Pro	Gln	Glu	Ala	Gln	Ile	Ile	Gln	Ala	Lys	Lys	
			20					25					30			
Arg	Lys	Thr	Thr	Gln	Asp	Val	Lys	Lys	Arg	Arg	Glu	Glu	Val	Thr	Lys	
			35				40					45				
Lys	His	Gln	Tyr	Glu	Ile	Arg	Asn	Cys	Trp	Pro	Pro	Val	Leu	Ser	Gly	
	50					55					60					
Gly	Ile	Ser	Pro	Cys	Ile	Ile	Ile	Glu	Thr	Pro	His	Lys	Glu	Ile	Gly	
65					70					75					80	
Thr	Ser	Asp	Phe	Ser	Arg	Phe	Thr	Asn	Tyr	Arg	Phe	Lys	Asn	Leu	Phe	
				85					90					95		
Ile	Asn	Pro	Ser	Pro	Leu	Pro	Asp	Leu	Ser	Trp	Gly	Cys	Ser	Lys	Glu	
			100					105					110			
Val	Trp	Leu	Asn	Met	Leu	Lys	Lys	Glu	Ser	Arg	Tyr	Val	His	Asp	Lys	
		115					120					125				
His	Phe	Glu	Val	Leu	His	Ser	Asp	Leu	Glu	Pro	Gln	Met	Arg	Ser	Ile	
	130					135					140					
Leu	Leu	Asp	Trp	Leu	Leu	Glu	Val	Cys	Glu	Val	Tyr	Thr	Leu	His	Arg	
145					150					155					160	
Glu	Thr	Phe	Tyr	Leu	Ala	Gln	Asp	Phe	Phe	Asp	Arg	Phe	Met	Leu	Thr	
				165					170					175		
Gln	Lys	Asp	Ile	Asn	Lys	Asn	Met	Leu	Gln	Leu	Ile	Gly	Ile	Thr	Ser	
			180					185					190			
Leu	Phe	Ile	Ala	Ser	Lys	Leu	Glu	Glu	Ile	Tyr	Ala	Pro	Lys	Leu	Gln	
	195						200					205				
Glu	Phe	Ala	Tyr	Val	Thr	Asp	Gly	Ala	Cys	Ser	Glu	Glu	Asp	Ile	Leu	
	210					215					220					
Arg	Met	Glu	Leu	Ile	Ile	Leu	Lys	Ala	Leu	Lys	Trp	Glu	Leu	Cys	Pro	
225					230					235					240	
Val	Thr	Ile	Ile	Ser	Trp	Leu	Asn	Leu	Phe	Leu	Gln	Val	Asp	Ala	Leu	
				245					250					255		
Lys	Asp	Ala	Pro	Lys	Val	Leu	Leu	Pro	Gln	Tyr	Ser	Gln	Glu	Thr	Phe	
		260						265					270			
Ile	Gln	Ile	Ala	Gln	Leu	Leu	Asp	Leu	Cys	Ile	Leu	Ala	Ile	Asp	Ser	
		275					280						285			
Leu	Glu	Phe	Gln	Tyr	Arg	Ile	Leu	Thr	Ala	Ala	Ala	Leu	Cys	His	Phe	
	290					295					300					
Thr	Ser	Ile	Glu	Val	Val	Lys	Lys	Ala	Ser	Gly	Leu	Glu	Trp	Asp	Ser	
305					310					315					320	
Ile	Ser	Glu	Cys	Val	Asp	Trp	Met	Val	Pro	Phe	Val	Asn	Val	Val	Lys	
				325					330					335		
Ser	Thr	Ser	Pro	Val	Lys	Leu	Lys	Thr	Phe	Lys	Lys	Ile	Pro	Met	Glu	
			340					345					350			
Asp	Arg	His	Asn	Ile	Gln	Thr	His	Thr	Asn	Tyr	Leu	Ala	Met	Leu	Glu	
		355					360					365				
Glu	Val	Asn	Tyr	Ile	Asn	Thr	Phe	Arg	Lys	Gly	Gly	Gln	Leu	Ser	Pro	

370	375	380
Val Cys Asn Gly Gly Ile Met Thr Pro Pro Lys Ser Thr Glu Lys Pro		
385	390	395
Pro Gly Lys His		400

<210> 129
 <211> 574
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >CDC7|ENSG00000097046|ENST00000428239|1725

<400> 129

Met Glu Ala Ser Leu Gly Ile Gln Met Asp Glu Pro Met Ala Phe Ser	
1 5 10 15	
Pro Gln Arg Asp Arg Phe Gln Ala Glu Gly Ser Leu Lys Lys Asn Glu	
20 25 30	
Gln Asn Phe Lys Leu Ala Gly Val Lys Lys Asp Ile Glu Lys Leu Tyr	
35 40 45	
Glu Ala Val Pro Gln Leu Ser Asn Val Phe Lys Ile Glu Asp Lys Ile	
50 55 60	
Gly Glu Gly Thr Phe Ser Ser Val Tyr Leu Ala Thr Ala Gln Leu Gln	
65 70 75 80	
Val Gly Pro Glu Glu Lys Ile Ala Leu Lys His Leu Ile Pro Thr Ser	
85 90 95	
His Pro Ile Arg Ile Ala Ala Glu Leu Gln Cys Leu Thr Val Ala Gly	
100 105 110	
Gly Gln Asp Asn Val Met Gly Val Lys Tyr Cys Phe Arg Lys Asn Asp	
115 120 125	
His Val Val Ile Ala Met Pro Tyr Leu Glu His Glu Ser Phe Leu Asp	
130 135 140	
Ile Leu Asn Ser Leu Ser Phe Gln Glu Val Arg Glu Tyr Met Leu Asn	
145 150 155 160	
Leu Phe Lys Ala Leu Lys Arg Ile His Gln Phe Gly Ile Val His Arg	
165 170 175	
Asp Val Lys Pro Ser Asn Phe Leu Tyr Asn Arg Arg Leu Lys Lys Tyr	
180 185 190	
Ala Leu Val Asp Phe Gly Leu Ala Gln Gly Thr His Asp Thr Lys Ile	
195 200 205	
Glu Leu Leu Lys Phe Val Gln Ser Glu Ala Gln Gln Glu Arg Cys Ser	
210 215 220	
Gln Asn Lys Ser His Ile Ile Thr Gly Asn Lys Ile Pro Leu Ser Gly	
225 230 235 240	
Pro Val Pro Lys Glu Leu Asp Gln Gln Ser Thr Thr Lys Ala Ser Val	
245 250 255	
Lys Arg Pro Tyr Thr Asn Ala Gln Ile Gln Ile Lys Gln Gly Lys Asp	
260 265 270	
Gly Lys Glu Gly Ser Val Gly Leu Ser Val Gln Arg Ser Val Phe Gly	
275 280 285	
Glu Arg Asn Phe Asn Ile His Ser Ser Ile Ser His Glu Ser Pro Ala	
290 295 300	

Val	Lys	Leu	Met	Lys	Gln	Ser	Lys	Thr	Val	Asp	Val	Leu	Ser	Arg	Lys
305					310					315					320
Leu	Ala	Thr	Lys	Lys	Lys	Ala	Ile	Ser	Thr	Lys	Val	Met	Asn	Ser	Ala
				325					330						335
Val	Met	Arg	Lys	Thr	Ala	Ser	Ser	Cys	Pro	Ala	Ser	Leu	Thr	Cys	Asp
			340					345						350	
Cys	Tyr	Ala	Thr	Asp	Lys	Val	Cys	Ser	Ile	Cys	Leu	Ser	Arg	Arg	Gln
		355					360					365			
Gln	Val	Ala	Pro	Arg	Ala	Gly	Thr	Pro	Gly	Phe	Arg	Ala	Pro	Glu	Val
	370					375					380				
Leu	Thr	Lys	Cys	Pro	Asn	Gln	Thr	Thr	Ala	Ile	Asp	Met	Trp	Ser	Ala
385					390						395				400
Gly	Val	Ile	Phe	Leu	Ser	Leu	Leu	Ser	Gly	Arg	Tyr	Pro	Phe	Tyr	Lys
				405					410						415
Ala	Ser	Asp	Asp	Leu	Thr	Ala	Leu	Ala	Gln	Ile	Met	Thr	Ile	Arg	Gly
			420					425						430	
Ser	Arg	Glu	Thr	Ile	Gln	Ala	Ala	Lys	Thr	Phe	Gly	Lys	Ser	Ile	Leu
		435					440						445		
Cys	Ser	Lys	Glu	Val	Pro	Ala	Gln	Asp	Leu	Arg	Lys	Leu	Cys	Glu	Arg
	450					455					460				
Leu	Arg	Gly	Met	Asp	Ser	Ser	Thr	Pro	Lys	Leu	Thr	Ser	Asp	Ile	Gln
465					470					475					480
Gly	His	Ala	Ser	His	Gln	Pro	Ala	Ile	Ser	Glu	Lys	Thr	Asp	His	Lys
				485					490						495
Ala	Ser	Cys	Leu	Val	Gln	Thr	Pro	Pro	Gly	Gln	Tyr	Ser	Gly	Asn	Ser
			500					505						510	
Phe	Lys	Lys	Gly	Asp	Ser	Asn	Ser	Cys	Glu	His	Cys	Phe	Asp	Glu	Tyr
		515					520					525			
Asn	Thr	Asn	Leu	Glu	Gly	Trp	Asn	Glu	Val	Pro	Asp	Glu	Ala	Tyr	Asp
	530					535					540				
Leu	Leu	Asp	Lys	Leu	Leu	Asp	Leu	Asn	Pro	Ala	Ser	Arg	Ile	Thr	Ala
545					550					555					560
Glu	Glu	Ala	Leu	Leu	His	Pro	Phe	Phe	Lys	Asp	Met	Ser	Leu		
				565					570						

<210> 130
 <211> 1490
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >CDK12|ENSG00000167258|ENST00000447079|4473

Met	Pro	Asn	Ser	Glu	Arg	His	Gly	Gly	Lys	Lys	Asp	Gly	Ser	Gly	Gly
1				5					10					15	
Ala	Ser	Gly	Thr	Leu	Gln	Pro	Ser	Ser	Gly	Gly	Gly	Ser	Ser	Asn	Ser
			20					25					30		
Arg	Glu	Arg	His	Arg	Leu	Val	Ser	Lys	His	Lys	Arg	His	Lys	Ser	Lys
		35					40					45			
His	Ser	Lys	Asp	Met	Gly	Leu	Val	Thr	Pro	Glu	Ala	Ala	Ser	Leu	Gly
	50					55					60				
Thr	Val	Ile	Lys	Pro	Leu	Val	Glu	Tyr	Asp	Asp	Ile	Ser	Ser	Asp	Ser

65					70					75					80
Asp	Thr	Phe	Ser	Asp	Asp	Met	Ala	Phe	Lys	Leu	Asp	Arg	Arg	Glu	Asn
				85					90					95	
Asp	Glu	Arg	Arg	Gly	Ser	Asp	Arg	Ser	Asp	Arg	Leu	His	Lys	His	Arg
			100					105					110		
His	His	Gln	His	Arg	Arg	Ser	Arg	Asp	Leu	Leu	Lys	Ala	Lys	Gln	Thr
		115					120					125			
Glu	Lys	Glu	Lys	Ser	Gln	Glu	Val	Ser	Ser	Lys	Ser	Gly	Ser	Met	Lys
	130					135				140					
Asp	Arg	Ile	Ser	Gly	Ser	Ser	Lys	Arg	Ser	Asn	Glu	Glu	Thr	Asp	Asp
145				150						155					160
Tyr	Gly	Lys	Ala	Gln	Val	Ala	Lys	Ser	Ser	Ser	Lys	Glu	Ser	Arg	Ser
			165					170						175	
Ser	Lys	Leu	His	Lys	Glu	Lys	Thr	Arg	Lys	Glu	Arg	Glu	Leu	Lys	Ser
		180					185					190			
Gly	His	Lys	Asp	Arg	Ser	Lys	Ser	His	Arg	Lys	Arg	Glu	Thr	Pro	Lys
	195					200					205				
Ser	Tyr	Lys	Thr	Val	Asp	Ser	Pro	Lys	Arg	Arg	Ser	Arg	Ser	Pro	His
	210				215						220				
Arg	Lys	Trp	Ser	Asp	Ser	Ser	Lys	Gln	Asp	Asp	Ser	Pro	Ser	Gly	Ala
225				230						235					240
Ser	Tyr	Gly	Gln	Asp	Tyr	Asp	Leu	Ser	Pro	Ser	Arg	Ser	His	Thr	Ser
			245						250					255	
Ser	Asn	Tyr	Asp	Ser	Tyr	Lys	Lys	Ser	Pro	Gly	Ser	Thr	Ser	Arg	Arg
		260						265				270			
Gln	Ser	Val	Ser	Pro	Pro	Tyr	Lys	Glu	Pro	Ser	Ala	Tyr	Gln	Ser	Ser
	275					280					285				
Thr	Arg	Ser	Pro	Ser	Pro	Tyr	Ser	Arg	Arg	Gln	Arg	Ser	Val	Ser	Pro
	290					295				300					
Tyr	Ser	Arg	Arg	Arg	Ser	Ser	Ser	Tyr	Glu	Arg	Ser	Gly	Ser	Tyr	Ser
305					310					315					320
Gly	Arg	Ser	Pro	Ser	Pro	Tyr	Gly	Arg	Arg	Arg	Ser	Ser	Ser	Pro	Phe
			325					330						335	
Leu	Ser	Lys	Arg	Ser	Leu	Ser	Arg	Ser	Pro	Leu	Pro	Ser	Arg	Lys	Ser
		340					345						350		
Met	Lys	Ser	Arg	Ser	Arg	Ser	Pro	Ala	Tyr	Ser	Arg	His	Ser	Ser	Ser
	355					360						365			
His	Ser	Lys	Lys	Lys	Arg	Ser	Ser	Ser	Arg	Ser	Arg	His	Ser	Ser	Ile
	370				375						380				
Ser	Pro	Val	Arg	Leu	Pro	Leu	Asn	Ser	Ser	Leu	Gly	Ala	Glu	Leu	Ser
385				390						395					400
Arg	Lys	Lys	Lys	Glu	Arg	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Lys	Met	Asp
			405					410						415	
Gly	Lys	Glu	Ser	Lys	Gly	Ser	Pro	Val	Phe	Leu	Pro	Arg	Lys	Glu	Asn
		420					425						430		
Ser	Ser	Val	Glu	Ala	Lys	Asp	Ser	Gly	Leu	Glu	Ser	Lys	Lys	Leu	Pro
		435				440						445			
Arg	Ser	Val	Lys	Leu	Glu	Lys	Ser	Ala	Pro	Asp	Thr	Glu	Leu	Val	Asn
	450					455					460				
Val	Thr	His	Leu	Asn	Thr	Glu	Val	Lys	Asn	Ser	Ser	Asp	Thr	Gly	Lys
465				470						475					480
Val	Lys	Leu	Asp	Glu	Asn	Ser	Glu	Lys	His	Leu	Val	Lys	Asp	Leu	Lys
			485					490					495		
Ala	Gln	Gly	Thr	Arg	Asp	Ser	Lys	Pro	Ile	Ala	Leu	Lys	Glu	Glu	Ile

930						935						940					
Leu	Glu	Leu	Ile	Ser	Arg	Leu	Cys	Gly	Ser	Pro	Cys	Pro	Ala	Val	Trp		
945					950					955					960		
Pro	Asp	Val	Ile	Lys	Leu	Pro	Tyr	Phe	Asn	Thr	Met	Lys	Pro	Lys	Lys		
				965					970						975		
Gln	Tyr	Arg	Arg	Arg	Leu	Arg	Glu	Glu	Phe	Ser	Phe	Ile	Pro	Ser	Ala		
				980					985					990			
Ala	Leu	Asp	Leu	Leu	Asp	His	Met	Leu	Thr	Leu	Asp	Pro	Ser	Lys	Arg		
		995					1000					1005					
Cys	Thr	Ala	Glu	Gln	Thr	Leu	Gln	Ser	Asp	Phe	Leu	Lys	Asp	Val	Glu		
1010						1015						1020					
Leu	Ser	Lys	Met	Ala	Pro	Pro	Asp	Leu	Pro	His	Trp	Gln	Asp	Cys	His		
1025					1030					1035					1040		
Glu	Leu	Trp	Ser	Lys	Lys	Arg	Arg	Arg	Gln	Arg	Gln	Ser	Gly	Val	Val		
				1045					1050						1055		
Val	Glu	Glu	Pro	Pro	Pro	Ser	Lys	Thr	Ser	Arg	Lys	Glu	Thr	Thr	Ser		
			1060					1065						1070			
Gly	Thr	Ser	Thr	Glu	Pro	Val	Lys	Asn	Ser	Ser	Pro	Ala	Pro	Pro	Gln		
		1075					1080					1085					
Pro	Ala	Pro	Gly	Lys	Val	Glu	Ser	Gly	Ala	Gly	Asp	Ala	Ile	Gly	Leu		
1090						1095					1100						
Ala	Asp	Ile	Thr	Gln	Gln	Leu	Asn	Gln	Ser	Glu	Leu	Ala	Val	Leu	Leu		
1105				1110						1115					1120		
Asn	Leu	Leu	Gln	Ser	Gln	Thr	Asp	Leu	Ser	Ile	Pro	Gln	Met	Ala	Gln		
				1125					1130					1135			
Leu	Leu	Asn	Ile	His	Ser	Asn	Pro	Glu	Met	Gln	Gln	Gln	Leu	Glu	Ala		
		1140					1145					1150					
Leu	Asn	Gln	Ser	Ile	Ser	Ala	Leu	Thr	Glu	Ala	Thr	Ser	Gln	Gln	Gln		
		1155					1160					1165					
Asp	Ser	Glu	Thr	Met	Ala	Pro	Glu	Glu	Ser	Leu	Lys	Glu	Ala	Pro	Ser		
1170					1175						1180						
Ala	Pro	Val	Ile	Leu	Pro	Ser	Ala	Glu	Gln	Thr	Thr	Leu	Glu	Ala	Ser		
1185				1190						1195					1200		
Ser	Thr	Pro	Ala	Asp	Met	Gln	Asn	Ile	Leu	Ala	Val	Leu	Leu	Ser	Gln		
				1205					1210					1215			
Leu	Met	Lys	Thr	Gln	Glu	Pro	Ala	Gly	Ser	Leu	Glu	Glu	Asn	Asn	Ser		
		1220						1225					1230				
Asp	Lys	Asn	Ser	Gly	Pro	Gln	Gly	Pro	Arg	Arg	Thr	Pro	Thr	Met	Pro		
1235							1240					1245					
Gln	Glu	Glu	Ala	Ala	Ala	Cys	Pro	Pro	His	Ile	Leu	Pro	Pro	Glu	Lys		
1250					1255						1260						
Arg	Pro	Pro	Glu	Pro	Pro	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro		
1265				1270					1275						1280		
Leu	Val	Glu	Gly	Asp	Leu	Ser	Ser	Ala	Pro	Gln	Glu	Leu	Asn	Pro	Ala		
				1285					1290					1295			
Val	Thr	Ala	Ala	Leu	Leu	Gln	Leu	Leu	Ser	Gln	Pro	Glu	Ala	Glu	Pro		
		1300					1305						1310				
Pro	Gly	His	Leu	Pro	His	Glu	His	Gln	Ala	Leu	Arg	Pro	Met	Glu	Tyr		
		1315				1320					1325						
Ser	Thr	Arg	Pro	Arg	Pro	Asn	Arg	Thr	Tyr	Gly	Asn	Thr	Asp	Gly	Pro		
1330					1335						1340						
Glu	Thr	Gly	Phe	Ser	Ala	Ile	Asp	Thr	Asp	Glu	Arg	Asn	Ser	Gly	Pro		
1345				1350					1355					1360			
Ala	Leu	Thr	Glu	Ser	Leu	Val	Gln	Thr	Leu	Val	Lys	Asn	Arg	Thr	Phe		

Asp	Gln	Pro	Ser	Asp	Ser	Cys	Gln	Glu	Tyr	Ser	Asp	Trp	Lys	Glu	Lys
210						215					220				
Lys	Thr	Tyr	Leu	Asn	Pro	Trp	Lys	Lys	Ile	Asp	Ser	Ala	Pro	Leu	Ala
225					230					235					240
Leu	Leu	His	Lys	Ile	Leu	Val	Glu	Asn	Pro	Ser	Ala	Arg	Ile	Thr	Ile
				245						250					255
Pro	Asp	Ile	Lys	Lys	Asp	Arg	Trp	Tyr	Asn	Lys	Pro	Leu	Lys	Lys	Gly
			260					265					270		
Ala	Lys	Arg	Pro	Arg	Val	Thr	Ser	Gly	Gly	Val	Ser	Glu	Ser	Pro	Ser
		275					280					285			
Gly	Phe	Ser	Lys	His	Ile	Gln	Ser	Asn	Leu	Asp	Phe	Ser	Pro	Val	Asn
290						295					300				
Ser	Ala	Ser	Ser	Glu	Glu	Asn	Val	Lys	Tyr	Ser	Ser	Ser	Gln	Pro	Glu
305					310					315					320
Pro	Arg	Thr	Gly	Leu	Ser	Leu	Trp	Asp	Thr	Ser	Pro	Ser	Tyr	Ile	Asp
				325					330						335
Lys	Leu	Val	Gln	Gly	Ile	Ser	Phe	Ser	Gln	Pro	Thr	Cys	Pro	Asp	His
			340					345					350		
Met	Leu	Leu	Asn	Ser	Gln	Leu	Leu	Gly	Thr	Pro	Gly	Ser	Ser	Gln	Asn
		355					360					365			
Pro	Trp	Gln	Arg	Leu	Val	Lys	Arg	Met	Thr	Arg	Phe	Phe	Thr	Lys	Leu
		370				375					380				
Asp	Ala	Asp	Lys	Ser	Tyr	Gln	Cys	Leu	Lys	Glu	Thr	Cys	Glu	Lys	Leu
385					390					395					400
Gly	Tyr	Gln	Trp	Lys	Lys	Ser	Cys	Met	Asn	Gln	Val	Thr	Ile	Ser	Thr
				405					410					415	
Thr	Asp	Arg	Arg	Asn	Asn	Lys	Leu	Ile	Phe	Lys	Val	Asn	Leu	Leu	Glu
			420					425					430		
Met	Asp	Asp	Lys	Ile	Leu	Val	Asp	Phe	Arg	Leu	Ser	Lys	Gly	Asp	Gly
		435					440					445			
Leu	Glu	Phe	Lys	Arg	His	Phe	Leu	Lys	Ile	Lys	Gly	Lys	Leu	Ile	Asp
	450					455					460				
Ile	Val	Ser	Ser	Gln	Lys	Ile	Trp	Leu	Pro	Ala	Thr				
465					470					475					

<210> 132
 <211> 586
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >CHEK2|ENSG00000183765|ENST00000382580|1761

<400> 132
 Met Ser Arg Glu Ser Asp Val Glu Ala Gln Gln Ser His Gly Ser Ser
 1 5 10 15
 Ala Cys Ser Gln Pro His Gly Ser Val Thr Gln Ser Gln Gly Ser Ser
 20 25 30
 Ser Gln Ser Gln Gly Ile Ser Ser Ser Thr Ser Thr Met Pro Asn
 35 40 45
 Ser Ser Gln Ser Ser His Ser Ser Gly Thr Leu Ser Ser Leu Glu
 50 55 60
 Thr Val Ser Thr Gln Glu Leu Tyr Ser Ile Pro Glu Asp Gln Glu Pro

65					70					75					80
Glu	Asp	Gln	Glu	Pro	Glu	Glu	Pro	Thr	Pro	Ala	Pro	Trp	Ala	Arg	Leu
				85					90					95	
Trp	Ala	Leu	Gln	Asp	Gly	Phe	Ala	Asn	Leu	Glu	Thr	Glu	Ser	Gly	His
			100					105					110		
Val	Thr	Gln	Ser	Asp	Leu	Glu	Leu	Leu	Leu	Ser	Ser	Asp	Pro	Pro	Ala
		115					120					125			
Ser	Ala	Ser	Gln	Ser	Ala	Gly	Ile	Arg	Gly	Val	Arg	His	His	Pro	Arg
	130					135					140				
Pro	Val	Cys	Ser	Leu	Lys	Cys	Val	Asn	Asp	Asn	Tyr	Trp	Phe	Gly	Arg
145					150				155						160
Asp	Lys	Ser	Cys	Glu	Tyr	Cys	Phe	Asp	Glu	Pro	Leu	Leu	Lys	Arg	Thr
				165				170						175	
Asp	Lys	Tyr	Arg	Thr	Tyr	Ser	Lys	Lys	His	Phe	Arg	Ile	Phe	Arg	Glu
			180					185					190		
Val	Gly	Pro	Lys	Asn	Ser	Tyr	Ile	Ala	Tyr	Ile	Glu	Asp	His	Ser	Gly
		195					200					205			
Asn	Gly	Thr	Phe	Val	Asn	Thr	Glu	Leu	Val	Gly	Lys	Gly	Lys	Arg	Arg
	210					215					220				
Pro	Leu	Asn	Asn	Asn	Ser	Glu	Ile	Ala	Leu	Ser	Leu	Ser	Arg	Asn	Lys
225					230					235					240
Val	Phe	Val	Phe	Phe	Asp	Leu	Thr	Val	Asp	Asp	Gln	Ser	Val	Tyr	Pro
				245					250					255	
Lys	Ala	Leu	Arg	Asp	Glu	Tyr	Ile	Met	Ser	Lys	Thr	Leu	Gly	Ser	Gly
			260					265					270		
Ala	Cys	Gly	Glu	Val	Lys	Leu	Ala	Phe	Glu	Arg	Lys	Thr	Cys	Lys	Lys
		275					280					285			
Val	Ala	Ile	Lys	Ile	Ile	Ser	Lys	Arg	Lys	Phe	Ala	Ile	Gly	Ser	Ala
	290					295					300				
Arg	Glu	Ala	Asp	Pro	Ala	Leu	Asn	Val	Glu	Thr	Glu	Ile	Glu	Ile	Leu
305					310					315					320
Lys	Lys	Leu	Asn	His	Pro	Cys	Ile	Ile	Lys	Ile	Lys	Asn	Phe	Phe	Asp
			325						330					335	
Ala	Glu	Asp	Tyr	Tyr	Ile	Val	Leu	Glu	Leu	Met	Glu	Gly	Gly	Glu	Leu
		340						345				350			
Phe	Asp	Lys	Val	Val	Gly	Asn	Lys	Arg	Leu	Lys	Glu	Ala	Thr	Cys	Lys
		355				360						365			
Leu	Tyr	Phe	Tyr	Gln	Met	Leu	Leu	Ala	Val	Gln	Tyr	Leu	His	Glu	Asn
	370				375						380				
Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	Glu	Asn	Val	Leu	Leu	Ser	Ser
385					390				395						400
Gln	Glu	Glu	Asp	Cys	Leu	Ile	Lys	Ile	Thr	Asp	Phe	Gly	His	Ser	Lys
			405						410					415	
Ile	Leu	Gly	Glu	Thr	Ser	Leu	Met	Arg	Thr	Leu	Cys	Gly	Thr	Pro	Thr
		420						425					430		
Tyr	Leu	Ala	Pro	Glu	Val	Leu	Val	Ser	Val	Gly	Thr	Ala	Gly	Tyr	Asn
		435					440					445			
Arg	Ala	Val	Asp	Cys	Trp	Ser	Leu	Gly	Val	Ile	Leu	Phe	Ile	Cys	Leu
	450					455					460				
Ser	Gly	Tyr	Pro	Pro	Phe	Ser	Glu	His	Arg	Thr	Gln	Val	Ser	Leu	Lys
465					470				475						480
Asp	Gln	Ile	Thr	Ser	Gly	Lys	Tyr	Asn	Phe	Ile	Pro	Glu	Val	Trp	Ala
			485						490					495	
Glu	Val	Ser	Glu	Lys	Ala	Leu	Asp	Leu	Val	Lys	Lys	Leu	Leu	Val	Val

Gln	Ala	Leu	Gln	Phe	Thr	Asp	Phe	Val	Glu	Asn	Asp	Lys	Leu	Val	Gly			
			260					265					270					
Val	Ala	Leu	Arg	Leu	Ala	Asn	Asn	Ser	Glu	His	Ile	Asn	Leu	Pro	Leu			
		275				280						285						
Pro	Glu	Asn	Asp	Phe	Ser	Asp	Cys	Glu	Ile	Ser	Tyr	Ser	Pro	Leu	Gln			
	290					295					300							
Ser	Asp	Glu	Asp	Thr	His	Asp	Ile	Asp	Glu	Lys	Pro	Asp	Asp	Ser	Gln			
305				310						315					320			
Glu	Gln	Leu	Phe	Phe	Thr	Glu	Ser	Ser	Lys	Asp	Gly	Ser	Leu	Glu	Glu			
			325						330					335				
Asp	Asp	Asp	Ser	Cys	Gly	Phe	Phe	Lys	Lys	Arg	His	Gly	Pro	Leu	Leu			
			340					345					350					
Lys	Asp	Gln	Asp	Glu	Ser	Cys	Pro	Lys	Val	Asn	Ser	Phe	Leu	Thr	Arg			
		355				360						365						
Asp	Lys	Tyr	Asp	Glu	Gly	Leu	Tyr	Arg	Phe	Asn	Ser	Leu	Asn	Asp	Leu			
	370				375						380							
Ser	Gln	Pro	Ile	Ser	Gln	Asn	Asn	Glu	Ser	Thr	Leu	Pro	Tyr	Asp	Leu			
385					390					395					400			
Ala	Cys	Thr	Gly	Gly	Asp	Phe	Val	Leu	Phe	Pro	Pro	Ala	Leu	Ala	Gly			
			405						410					415				
Lys	Leu	Ala	Ala	Ser	Val	His	Gln	Ala	Thr	Lys	Ala	Lys	Pro	Asp	Glu			
		420						425					430					
Pro	Glu	Phe	His	Ser	Ala	Gln	Ser	Asn	Lys	Gln	Lys	Gln	Val	Ile	Glu			
	435					440						445						
Glu	Ser	Ser	Val	Tyr	Asn	Gln	Val	Ser	Leu	Pro	Leu	Val	Lys	Ser	Leu			
	450				455					460								
Met	Leu	Lys	Pro	Phe	Glu	Ser	Gln	Val	Glu	Gly	Tyr	Leu	Ser	Ser	Gln			
465				470					475						480			
Pro	Thr	Gln	Asn	Thr	Ile	Arg	Lys	Leu	Ser	Ser	Glu	Asn	Leu	Asn	Ala			
			485					490						495				
Lys	Asn	Asn	Thr	Asn	Ser	Ala	Cys	Phe	Cys	Arg	Lys	Ala	Leu	Glu	Gly			
		500					505						510					
Val	Pro	Val	Gly	Lys	Ala	Thr	Ile	Leu	Asn	Thr	Glu	Asn	Leu	Ser	Ser			
		515					520					525						
Thr	Pro	Ala	Pro	Lys	Tyr	Leu	Lys	Ile	Leu	Pro	Ser	Gly	Leu	Lys	Tyr			
	530				535						540							
Asn	Ala	Arg	His	Pro	Ser	Thr	Lys	Val	Met	Lys	Gln	Met	Asp	Ile	Gly			
545				550					555						560			
Val	Tyr	Phe	Gly	Leu	Pro	Pro	Lys	Arg	Lys	Glu	Glu	Lys	Leu	Leu	Gly			
			565					570						575				
Glu	Ser	Ala	Leu	Glu	Gly	Ile	Asn	Leu	Asn	Pro	Val	Pro	Ser	Pro	Asn			
		580					585						590					
Gln	Lys	Arg	Ser	Ser	Gln	Cys	Lys	Arg	Lys	Ala	Glu	Lys	Ser	Leu	Ser			
		595				600						605						
Asp	Leu	Glu	Phe	Asp	Ala	Ser	Thr	Leu	His	Glu	Ser	Gln	Leu	Ser	Val			
	610				615					620								
Glu	Leu	Ser	Ser	Glu	Arg	Ser	Gln	Arg	Gln	Lys	Lys	Arg	Cys	Arg	Lys			
625				630						635					640			
Ser	Asn	Ser	Leu	Gln	Glu	Gly	Ala	Cys	Gln	Lys	Arg	Ser	Asp	His	Leu			
			645					650						655				
Ile	Asn	Thr	Glu	Ser	Glu	Ala	Val	Asn	Leu	Ser	Lys	Val	Lys	Val	Phe			
		660					665						670					
Thr	Lys	Ser	Ala	His	Gly	Gly	Leu	Gln	Arg	Gly	Asn	Lys	Lys	Ile	Pro			
		675					680					685						

<223> >DCLRE1B|ENSG00000118655|ENST00000369563|1599

<400> 134

Met	Asn	Gly	Val	Leu	Ile	Pro	His	Thr	Pro	Ile	Ala	Val	Asp	Phe	Trp
1				5					10					15	
Ser	Leu	Arg	Arg	Ala	Gly	Thr	Ala	Arg	Leu	Phe	Phe	Leu	Ser	His	Met
			20					25					30		
His	Ser	Asp	His	Thr	Val	Gly	Leu	Ser	Ser	Thr	Trp	Ala	Arg	Pro	Leu
		35				40					45				
Tyr	Cys	Ser	Pro	Ile	Thr	Ala	His	Leu	Leu	His	Arg	His	Leu	Gln	Val
	50					55				60					
Ser	Lys	Gln	Trp	Ile	Gln	Ala	Leu	Glu	Val	Gly	Glu	Ser	His	Val	Leu
65					70					75					80
Pro	Leu	Asp	Glu	Ile	Gly	Gln	Glu	Thr	Met	Thr	Val	Thr	Leu	Leu	Asp
					85				90					95	
Ala	Asn	His	Cys	Pro	Gly	Ser	Val	Met	Phe	Leu	Phe	Glu	Gly	Tyr	Phe
			100					105					110		
Gly	Thr	Ile	Leu	Tyr	Thr	Gly	Asp	Phe	Arg	Tyr	Thr	Pro	Ser	Met	Leu
		115				120						125			
Lys	Glu	Pro	Ala	Leu	Thr	Leu	Gly	Lys	Gln	Ile	His	Thr	Leu	Tyr	Leu
	130					135					140				
Asp	Asn	Thr	Asn	Cys	Asn	Pro	Ala	Leu	Val	Leu	Pro	Ser	Arg	Gln	Glu
145					150					155					160
Ala	Ala	His	Gln	Ile	Val	Gln	Leu	Ile	Arg	Lys	His	Pro	Gln	His	Asn
			165						170					175	
Ile	Lys	Ile	Gly	Leu	Tyr	Ser	Leu	Gly	Lys	Glu	Ser	Leu	Leu	Glu	Gln
			180					185					190		
Leu	Ala	Leu	Glu	Phe	Gln	Thr	Trp	Val	Val	Leu	Ser	Pro	Arg	Arg	Leu
		195					200					205			
Glu	Leu	Val	Gln	Leu	Leu	Gly	Leu	Ala	Asp	Val	Phe	Thr	Val	Glu	Glu
	210					215					220				
Lys	Ala	Gly	Arg	Ile	His	Ala	Val	Asp	His	Met	Glu	Ile	Cys	His	Ser
225					230					235					240
Asn	Met	Leu	Arg	Trp	Asn	Gln	Thr	His	Pro	Thr	Ile	Ala	Ile	Leu	Pro
			245						250					255	
Thr	Ser	Arg	Lys	Ile	His	Ser	Ser	His	Pro	Asp	Ile	His	Val	Ile	Pro
			260					265					270		
Tyr	Ser	Asp	His	Ser	Ser	Tyr	Ser	Glu	Leu	Arg	Ala	Phe	Val	Ala	Ala
		275					280					285			
Leu	Lys	Pro	Cys	Gln	Val	Val	Pro	Ile	Val	Ser	Arg	Arg	Pro	Cys	Gly
	290					295				300					
Gly	Phe	Gln	Asp	Ser	Leu	Ser	Pro	Arg	Ile	Ser	Val	Pro	Leu	Ile	Pro
305					310					315					320
Asp	Ser	Val	Gln	Gln	Tyr	Met	Ser	Ser	Ser	Ser	Arg	Lys	Pro	Ser	Leu
					325				330					335	
Leu	Trp	Leu	Leu	Glu	Arg	Arg	Leu	Lys	Arg	Pro	Arg	Thr	Gln	Gly	Val
			340					345					350		
Val	Phe	Glu	Ser	Pro	Glu	Glu	Ser	Ala	Asp	Gln	Ser	Gln	Ala	Asp	Arg
		355					360					365			
Asp	Ser	Lys	Lys	Ala	Lys	Lys	Glu	Lys	Leu	Ser	Pro	Trp	Pro	Ala	Asp
	370					375					380				
Leu	Glu	Lys	Gln	Pro	Ser	His	His	Pro	Leu	Arg	Ile	Lys	Lys	Gln	Leu
385					390					395					400
Phe	Pro	Asp	Leu	Tyr	Ser	Lys	Glu	Trp	Asn	Lys	Ala	Val	Pro	Phe	Cys

				405				410					415			
Glu	Ser	Gln	Lys	Arg	Val	Thr	Met	Leu	Thr	Ala	Pro	Leu	Gly	Phe	Ser	
			420					425					430			
Val	His	Leu	Arg	Ser	Thr	Asp	Glu	Glu	Phe	Ile	Ser	Gln	Lys	Thr	Arg	
		435					440					445				
Glu	Glu	Ile	Gly	Leu	Gly	Ser	Pro	Leu	Val	Pro	Met	Gly	Asp	Asp	Asp	
	450				455					460						
Gly	Gly	Pro	Glu	Ala	Thr	Gly	Asn	Gln	Ser	Ala	Trp	Met	Gly	His	Gly	
465				470				475							480	
Ser	Pro	Leu	Ser	His	Ser	Ser	Lys	Gly	Thr	Pro	Leu	Leu	Ala	Thr	Glu	
			485					490					495			
Phe	Arg	Gly	Leu	Ala	Leu	Lys	Tyr	Leu	Leu	Thr	Pro	Val	Asn	Phe	Phe	
		500						505					510			
Gln	Ala	Gly	Tyr	Ser	Ser	Arg	Arg	Phe	Asp	Gln	Gln	Val	Glu	Lys	Tyr	
	515						520					525				
His	Lys	Pro	Cys													
	530															

<210> 135

<211> 692

<212> PRT

<213> Homo sapiens

<220>

<223> >DCLRE1C|ENSG00000152457|ENST00000378278|2079

<400> 135

Met	Ser	Ser	Phe	Glu	Gly	Gln	Met	Ala	Glu	Tyr	Pro	Thr	Ile	Ser	Ile	
1				5				10					15			
Asp	Arg	Phe	Asp	Arg	Glu	Asn	Leu	Arg	Ala	Arg	Ala	Tyr	Phe	Leu	Ser	
		20						25				30				
His	Cys	His	Lys	Asp	His	Met	Lys	Gly	Leu	Arg	Ala	Pro	Thr	Leu	Lys	
	35					40					45					
Arg	Arg	Leu	Glu	Cys	Ser	Leu	Lys	Val	Tyr	Leu	Tyr	Cys	Ser	Pro	Val	
	50				55					60						
Thr	Lys	Glu	Leu	Leu	Leu	Thr	Ser	Pro	Lys	Tyr	Arg	Phe	Trp	Lys	Lys	
65			70						75					80		
Arg	Ile	Ile	Ser	Ile	Glu	Ile	Glu	Thr	Pro	Thr	Gln	Ile	Ser	Leu	Val	
		85					90					95				
Asp	Glu	Ala	Ser	Gly	Glu	Lys	Glu	Glu	Ile	Val	Val	Thr	Leu	Leu	Pro	
	100						105					110				
Ala	Gly	His	Cys	Pro	Gly	Ser	Val	Met	Phe	Leu	Phe	Gln	Gly	Asn	Asn	
	115					120						125				
Gly	Thr	Val	Leu	Tyr	Thr	Gly	Asp	Phe	Arg	Leu	Ala	Gln	Gly	Glu	Ala	
	130				135					140						
Ala	Arg	Met	Glu	Leu	Leu	His	Ser	Gly	Gly	Arg	Val	Lys	Asp	Ile	Gln	
145				150					155					160		
Ser	Val	Tyr	Leu	Asp	Thr	Thr	Phe	Cys	Asp	Pro	Arg	Phe	Tyr	Gln	Ile	
		165						170					175			
Pro	Ser	Arg	Glu	Glu	Cys	Leu	Ser	Gly	Val	Leu	Glu	Leu	Val	Arg	Ser	
	180						185					190				
Trp	Ile	Thr	Arg	Ser	Pro	Tyr	His	Val	Val	Trp	Leu	Asn	Cys	Lys	Ala	
	195					200						205				

Ala	Tyr	Gly	Tyr	Glu	Tyr	Leu	Phe	Thr	Asn	Leu	Ser	Glu	Glu	Leu	Gly
210						215					220				
Val	Gln	Val	His	Val	Asn	Lys	Leu	Asp	Met	Phe	Arg	Asn	Met	Pro	Glu
225					230					235					240
Ile	Leu	His	His	Leu	Thr	Thr	Asp	Arg	Asn	Thr	Gln	Ile	His	Ala	Cys
				245					250					255	
Arg	His	Pro	Lys	Ala	Glu	Glu	Tyr	Phe	Gln	Trp	Ser	Lys	Leu	Pro	Cys
			260					265					270		
Gly	Ile	Thr	Ser	Arg	Asn	Arg	Ile	Pro	Leu	His	Ile	Ile	Ser	Ile	Lys
		275					280					285			
Pro	Ser	Thr	Met	Trp	Phe	Gly	Glu	Arg	Ser	Arg	Lys	Thr	Asn	Val	Ile
		290				295					300				
Val	Arg	Thr	Gly	Glu	Ser	Ser	Tyr	Arg	Ala	Cys	Phe	Ser	Phe	His	Ser
305					310					315					320
Ser	Tyr	Ser	Glu	Ile	Lys	Asp	Phe	Leu	Ser	Tyr	Leu	Cys	Pro	Val	Asn
				325					330					335	
Ala	Tyr	Pro	Asn	Val	Ile	Pro	Val	Gly	Thr	Thr	Met	Asp	Lys	Val	Val
			340					345					350		
Glu	Ile	Leu	Lys	Pro	Leu	Cys	Arg	Ser	Ser	Gln	Ser	Thr	Glu	Pro	Lys
		355					360					365			
Tyr	Lys	Pro	Leu	Gly	Lys	Leu	Lys	Arg	Ala	Arg	Thr	Val	His	Arg	Asp
		370				375					380				
Ser	Glu	Glu	Glu	Asp	Asp	Tyr	Leu	Phe	Asp	Asp	Pro	Leu	Pro	Ile	Pro
385					390					395					400
Leu	Arg	His	Lys	Val	Pro	Tyr	Pro	Glu	Thr	Phe	His	Pro	Glu	Val	Phe
				405					410					415	
Ser	Met	Thr	Ala	Val	Ser	Glu	Lys	Gln	Pro	Glu	Lys	Leu	Arg	Gln	Thr
			420					425					430		
Pro	Gly	Cys	Cys	Arg	Ala	Glu	Cys	Met	Gln	Ser	Ser	Arg	Phe	Thr	Asn
		435					440					445			
Phe	Val	Asp	Cys	Glu	Glu	Ser	Asn	Ser	Glu	Ser	Glu	Glu	Glu	Val	Gly
	450					455					460				
Ile	Pro	Ala	Ser	Leu	Gln	Gly	Asp	Leu	Gly	Ser	Val	Leu	His	Leu	Gln
465					470					475					480
Lys	Ala	Asp	Gly	Asp	Val	Pro	Gln	Trp	Glu	Val	Phe	Phe	Lys	Arg	Asn
				485					490					495	
Asp	Glu	Ile	Thr	Asp	Glu	Ser	Leu	Glu	Asn	Phe	Pro	Ser	Ser	Thr	Val
			500					505					510		
Ala	Gly	Gly	Ser	Gln	Ser	Pro	Lys	Leu	Phe	Ser	Asp	Ser	Asp	Gly	Glu
		515					520				525				
Ser	Thr	His	Ile	Ser	Ser	Gln	Asn	Ser	Ser	Gln	Ser	Thr	His	Ile	Thr
		530				535					540				
Glu	Gln	Gly	Ser	Gln	Gly	Trp	Asp	Ser	Gln	Ser	Asp	Thr	Val	Leu	Leu
545					550					555					560
Ser	Ser	Gln	Glu	Arg	Asn	Ser	Gly	Asp	Ile	Thr	Ser	Leu	Asp	Lys	Ala
				565					570					575	
Asp	Tyr	Arg	Pro	Thr	Ile	Lys	Glu	Asn	Ile	Pro	Ala	Ser	Leu	Met	Glu
			580					585					590		
Gln	Asn	Val	Ile	Cys	Pro	Lys	Asp	Thr	Tyr	Ser	Asp	Leu	Lys	Ser	Arg
		595					600					605			
Asp	Lys	Asp	Val	Thr	Ile	Val	Pro	Ser	Thr	Gly	Glu	Pro	Thr	Thr	Leu
	610					615					620				
Ser	Ser	Glu	Thr	His	Ile	Pro	Glu	Glu	Lys	Ser	Leu	Leu	Asn	Leu	Ser
625					630					635					640

705					710					715					720
Gln	Phe	Ser	Ala	Asn	Thr	Gly	Pro	Ala	His	Tyr	Met	Thr	Glu	Gly	His
				725					730					735	
Leu	Thr	Met	Arg	Gln	Gly	Ala	Asp	Arg	Glu	Glu	Ser	Pro	Met	Thr	Gly
			740					745					750		
Val	Cys	Val	Gln	Gln	Ser	Pro	Val	Ala	Ser	Ser					
		755					760								

<210> 137
 <211> 1210
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >EGFR|ENSG00000146648|ENST00000275493|3633

<400> 137

Met	Arg	Pro	Ser	Gly	Thr	Ala	Gly	Ala	Ala	Leu	Leu	Ala	Leu	Leu	Ala
1				5				10					15		
Ala	Leu	Cys	Pro	Ala	Ser	Arg	Ala	Leu	Glu	Glu	Lys	Lys	Val	Cys	Gln
			20				25						30		
Gly	Thr	Ser	Asn	Lys	Leu	Thr	Gln	Leu	Gly	Thr	Phe	Glu	Asp	His	Phe
		35				40					45				
Leu	Ser	Leu	Gln	Arg	Met	Phe	Asn	Asn	Cys	Glu	Val	Val	Leu	Gly	Asn
	50				55					60					
Leu	Glu	Ile	Thr	Tyr	Val	Gln	Arg	Asn	Tyr	Asp	Leu	Ser	Phe	Leu	Lys
65				70					75					80	
Thr	Ile	Gln	Glu	Val	Ala	Gly	Tyr	Val	Leu	Ile	Ala	Leu	Asn	Thr	Val
				85				90						95	
Glu	Arg	Ile	Pro	Leu	Glu	Asn	Leu	Gln	Ile	Ile	Arg	Gly	Asn	Met	Tyr
			100				105						110		
Tyr	Glu	Asn	Ser	Tyr	Ala	Leu	Ala	Val	Leu	Ser	Asn	Tyr	Asp	Ala	Asn
		115				120						125			
Lys	Thr	Gly	Leu	Lys	Glu	Leu	Pro	Met	Arg	Asn	Leu	Gln	Glu	Ile	Leu
	130					135					140				
His	Gly	Ala	Val	Arg	Phe	Ser	Asn	Asn	Pro	Ala	Leu	Cys	Asn	Val	Glu
145				150					155						160
Ser	Ile	Gln	Trp	Arg	Asp	Ile	Val	Ser	Ser	Asp	Phe	Leu	Ser	Asn	Met
			165				170							175	
Ser	Met	Asp	Phe	Gln	Asn	His	Leu	Gly	Ser	Cys	Gln	Lys	Cys	Asp	Pro
		180					185						190		
Ser	Cys	Pro	Asn	Gly	Ser	Cys	Trp	Gly	Ala	Gly	Glu	Glu	Asn	Cys	Gln
		195				200						205			
Lys	Leu	Thr	Lys	Ile	Ile	Cys	Ala	Gln	Gln	Cys	Ser	Gly	Arg	Cys	Arg
	210					215					220				
Gly	Lys	Ser	Pro	Ser	Asp	Cys	Cys	His	Asn	Gln	Cys	Ala	Ala	Gly	Cys
225				230					235					240	
Thr	Gly	Pro	Arg	Glu	Ser	Asp	Cys	Leu	Val	Cys	Arg	Lys	Phe	Arg	Asp
			245					250						255	
Glu	Ala	Thr	Cys	Lys	Asp	Thr	Cys	Pro	Pro	Leu	Met	Leu	Tyr	Asn	Pro
		260					265						270		
Thr	Thr	Tyr	Gln	Met	Asp	Val	Asn	Pro	Glu	Gly	Lys	Tyr	Ser	Phe	Gly
		275					280					285			

Ala	Thr	Cys	Val	Lys	Lys	Cys	Pro	Arg	Asn	Tyr	Val	Val	Thr	Asp	His
290						295					300				
Gly	Ser	Cys	Val	Arg	Ala	Cys	Gly	Ala	Asp	Ser	Tyr	Glu	Met	Glu	Glu
305					310					315					320
Asp	Gly	Val	Arg	Lys	Cys	Lys	Lys	Cys	Glu	Gly	Pro	Cys	Arg	Lys	Val
				325					330					335	
Cys	Asn	Gly	Ile	Gly	Ile	Gly	Glu	Phe	Lys	Asp	Ser	Leu	Ser	Ile	Asn
			340					345					350		
Ala	Thr	Asn	Ile	Lys	His	Phe	Lys	Asn	Cys	Thr	Ser	Ile	Ser	Gly	Asp
		355				360						365			
Leu	His	Ile	Leu	Pro	Val	Ala	Phe	Arg	Gly	Asp	Ser	Phe	Thr	His	Thr
370						375					380				
Pro	Pro	Leu	Asp	Pro	Gln	Glu	Leu	Asp	Ile	Leu	Lys	Thr	Val	Lys	Glu
385					390						395				400
Ile	Thr	Gly	Phe	Leu	Leu	Ile	Gln	Ala	Trp	Pro	Glu	Asn	Arg	Thr	Asp
				405					410					415	
Leu	His	Ala	Phe	Glu	Asn	Leu	Glu	Ile	Ile	Arg	Gly	Arg	Thr	Lys	Gln
			420					425					430		
His	Gly	Gln	Phe	Ser	Leu	Ala	Val	Val	Ser	Leu	Asn	Ile	Thr	Ser	Leu
			435				440					445			
Gly	Leu	Arg	Ser	Leu	Lys	Glu	Ile	Ser	Asp	Gly	Asp	Val	Ile	Ile	Ser
						455					460				
Gly	Asn	Lys	Asn	Leu	Cys	Tyr	Ala	Asn	Thr	Ile	Asn	Trp	Lys	Lys	Leu
465					470					475					480
Phe	Gly	Thr	Ser	Gly	Gln	Lys	Thr	Lys	Ile	Ile	Ser	Asn	Arg	Gly	Glu
				485					490					495	
Asn	Ser	Cys	Lys	Ala	Thr	Gly	Gln	Val	Cys	His	Ala	Leu	Cys	Ser	Pro
			500					505					510		
Glu	Gly	Cys	Trp	Gly	Pro	Glu	Pro	Arg	Asp	Cys	Val	Ser	Cys	Arg	Asn
			515				520					525			
Val	Ser	Arg	Gly	Arg	Glu	Cys	Val	Asp	Lys	Cys	Asn	Leu	Leu	Glu	Gly
						535					540				
Glu	Pro	Arg	Glu	Phe	Val	Glu	Asn	Ser	Glu	Cys	Ile	Gln	Cys	His	Pro
545					550					555					560
Glu	Cys	Leu	Pro	Gln	Ala	Met	Asn	Ile	Thr	Cys	Thr	Gly	Arg	Gly	Pro
				565					570					575	
Asp	Asn	Cys	Ile	Gln	Cys	Ala	His	Tyr	Ile	Asp	Gly	Pro	His	Cys	Val
			580					585					590		
Lys	Thr	Cys	Pro	Ala	Gly	Val	Met	Gly	Glu	Asn	Asn	Thr	Leu	Val	Trp
			595				600					605			
Lys	Tyr	Ala	Asp	Ala	Gly	His	Val	Cys	His	Leu	Cys	His	Pro	Asn	Cys
						615					620				
Thr	Tyr	Gly	Cys	Thr	Gly	Pro	Gly	Leu	Glu	Gly	Cys	Pro	Thr	Asn	Gly
625					630					635					640
Pro	Lys	Ile	Pro	Ser	Ile	Ala	Thr	Gly	Met	Val	Gly	Ala	Leu	Leu	Leu
				645					650					655	
Leu	Leu	Val	Val	Ala	Leu	Gly	Ile	Gly	Leu	Phe	Met	Arg	Arg	Arg	His
			660					665					670		
Ile	Val	Arg	Lys	Arg	Thr	Leu	Arg	Arg	Leu	Leu	Gln	Glu	Arg	Glu	Leu
			675				680					685			
Val	Glu	Pro	Leu	Thr	Pro	Ser	Gly	Glu	Ala	Pro	Asn	Gln	Ala	Leu	Leu
						695					700				
Arg	Ile	Leu	Lys	Glu	Thr	Glu	Phe	Lys	Lys	Ile	Lys	Val	Leu	Gly	Ser
705					710					715					720

Gly	Ala	Phe	Gly	Thr	Val	Tyr	Lys	Gly	Leu	Trp	Ile	Pro	Glu	Gly	Glu	
				725					730					735		
Lys	Val	Lys	Ile	Pro	Val	Ala	Ile	Lys	Glu	Leu	Arg	Glu	Ala	Thr	Ser	
			740					745					750			
Pro	Lys	Ala	Asn	Lys	Glu	Ile	Leu	Asp	Glu	Ala	Tyr	Val	Met	Ala	Ser	
		755					760					765				
Val	Asp	Asn	Pro	His	Val	Cys	Arg	Leu	Leu	Gly	Ile	Cys	Leu	Thr	Ser	
	770					775					780					
Thr	Val	Gln	Leu	Ile	Thr	Gln	Leu	Met	Pro	Phe	Gly	Cys	Leu	Leu	Asp	
785					790					795					800	
Tyr	Val	Arg	Glu	His	Lys	Asp	Asn	Ile	Gly	Ser	Gln	Tyr	Leu	Leu	Asn	
				805					810					815		
Trp	Cys	Val	Gln	Ile	Ala	Lys	Gly	Met	Asn	Tyr	Leu	Glu	Asp	Arg	Arg	
			820					825					830			
Leu	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu	Val	Lys	Thr	Pro	
		835					840					845				
Gln	His	Val	Lys	Ile	Thr	Asp	Phe	Gly	Leu	Ala	Lys	Leu	Leu	Gly	Ala	
	850					855					860					
Glu	Glu	Lys	Glu	Tyr	His	Ala	Glu	Gly	Gly	Lys	Val	Pro	Ile	Lys	Trp	
865					870				875						880	
Met	Ala	Leu	Glu	Ser	Ile	Leu	His	Arg	Ile	Tyr	Thr	His	Gln	Ser	Asp	
				885					890					895		
Val	Trp	Ser	Tyr	Gly	Val	Thr	Val	Trp	Glu	Leu	Met	Thr	Phe	Gly	Ser	
			900					905					910			
Lys	Pro	Tyr	Asp	Gly	Ile	Pro	Ala	Ser	Glu	Ile	Ser	Ser	Ile	Leu	Glu	
		915					920					925				
Lys	Gly	Glu	Arg	Leu	Pro	Gln	Pro	Pro	Ile	Cys	Thr	Ile	Asp	Val	Tyr	
	930					935					940					
Met	Ile	Met	Val	Lys	Cys	Trp	Met	Ile	Asp	Ala	Asp	Ser	Arg	Pro	Lys	
945					950					955					960	
Phe	Arg	Glu	Leu	Ile	Ile	Glu	Phe	Ser	Lys	Met	Ala	Arg	Asp	Pro	Gln	
				965					970					975		
Arg	Tyr	Leu	Val	Ile	Gln	Gly	Asp	Glu	Arg	Met	His	Leu	Pro	Ser	Pro	
			980					985					990			
Thr	Asp	Ser	Asn	Phe	Tyr	Arg	Ala	Leu	Met	Asp	Glu	Glu	Asp	Met	Asp	
		995					1000					1005				
Asp	Val	Val	Asp	Ala	Asp	Glu	Tyr	Leu	Ile	Pro	Gln	Gln	Gly	Phe	Phe	
	1010					1015					1020					
Ser	Ser	Pro	Ser	Thr	Ser	Arg	Thr	Pro	Leu	Leu	Ser	Ser	Leu	Ser	Ala	
1025					1030				1035						1040	
Thr	Ser	Asn	Asn	Ser	Thr	Val	Ala	Cys	Ile	Asp	Arg	Asn	Gly	Leu	Gln	
			1045						1050				1055			
Ser	Cys	Pro	Ile	Lys	Glu	Asp	Ser	Phe	Leu	Gln	Arg	Tyr	Ser	Ser	Asp	
			1060					1065					1070			
Pro	Thr	Gly	Ala	Leu	Thr	Glu	Asp	Ser	Ile	Asp	Asp	Thr	Phe	Leu	Pro	
		1075					1080					1085				
Val	Pro	Glu	Tyr	Ile	Asn	Gln	Ser	Val	Pro	Lys	Arg	Pro	Ala	Gly	Ser	
	1090					1095					1100					
Val	Gln	Asn	Pro	Val	Tyr	His	Asn	Gln	Pro	Leu	Asn	Pro	Ala	Pro	Ser	
1105					1110					1115					1120	
Arg	Asp	Pro	His	Tyr	Gln	Asp	Pro	His	Ser	Thr	Ala	Val	Gly	Asn	Pro	
			1125						1130					1135		
Glu	Tyr	Leu	Asn	Thr	Val	Gln	Pro	Thr	Cys	Val	Asn	Ser	Thr	Phe	Asp	
			1140					1145					1150			

Ser Pro Ala His Trp Ala Gln Lys Gly Ser His Gln Ile Ser Leu Asp
1155 1160 1165
Asn Pro Asp Tyr Gln Gln Asp Phe Phe Pro Lys Glu Ala Lys Pro Asn
1170 1175 1180
Gly Ile Phe Lys Gly Ser Thr Ala Glu Asn Ala Glu Tyr Leu Arg Val
1185 1190 1195 1200
Ala Pro Gln Ser Ser Glu Phe Ile Gly Ala
1205 1210

<210> 138
<211> 1255
<212> PRT
<213> Homo sapiens

<220>
<223> >ERBB2|ENSG00000141736|ENST00000269571|3768

<400> 138
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15
Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30
Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45
Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80
Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95
Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110
Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125
Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140
Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160
Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175
Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190
His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205
Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220
Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240
Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255
His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270
Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg

Page 429

705					710					715					720
Arg	Lys	Val	Lys	Val	Leu	Gly	Ser	Gly	Ala	Phe	Gly	Thr	Val	Tyr	Lys
				725					730					735	
Gly	Ile	Trp	Ile	Pro	Asp	Gly	Glu	Asn	Val	Lys	Ile	Pro	Val	Ala	Ile
			740					745					750		
Lys	Val	Leu	Arg	Glu	Asn	Thr	Ser	Pro	Lys	Ala	Asn	Lys	Glu	Ile	Leu
		755					760					765			
Asp	Glu	Ala	Tyr	Val	Met	Ala	Gly	Val	Gly	Ser	Pro	Tyr	Val	Ser	Arg
	770					775					780				
Leu	Leu	Gly	Ile	Cys	Leu	Thr	Ser	Thr	Val	Gln	Leu	Val	Thr	Gln	Leu
785					790				795						800
Met	Pro	Tyr	Gly	Cys	Leu	Leu	Asp	His	Val	Arg	Glu	Asn	Arg	Gly	Arg
			805						810					815	
Leu	Gly	Ser	Gln	Asp	Leu	Leu	Asn	Trp	Cys	Met	Gln	Ile	Ala	Lys	Gly
			820					825					830		
Met	Ser	Tyr	Leu	Glu	Asp	Val	Arg	Leu	Val	His	Arg	Asp	Leu	Ala	Ala
	835						840					845			
Arg	Asn	Val	Leu	Val	Lys	Ser	Pro	Asn	His	Val	Lys	Ile	Thr	Asp	Phe
	850					855					860				
Gly	Leu	Ala	Arg	Leu	Leu	Asp	Ile	Asp	Glu	Thr	Glu	Tyr	His	Ala	Asp
865					870				875						880
Gly	Gly	Lys	Val	Pro	Ile	Lys	Trp	Met	Ala	Leu	Glu	Ser	Ile	Leu	Arg
			885						890					895	
Arg	Arg	Phe	Thr	His	Gln	Ser	Asp	Val	Trp	Ser	Tyr	Gly	Val	Thr	Val
		900						905					910		
Trp	Glu	Leu	Met	Thr	Phe	Gly	Ala	Lys	Pro	Tyr	Asp	Gly	Ile	Pro	Ala
	915					920						925			
Arg	Glu	Ile	Pro	Asp	Leu	Leu	Glu	Lys	Gly	Glu	Arg	Leu	Pro	Gln	Pro
	930					935					940				
Pro	Ile	Cys	Thr	Ile	Asp	Val	Tyr	Met	Ile	Met	Val	Lys	Cys	Trp	Met
945					950					955					960
Ile	Asp	Ser	Glu	Cys	Arg	Pro	Arg	Phe	Arg	Glu	Leu	Val	Ser	Glu	Phe
			965						970					975	
Ser	Arg	Met	Ala	Arg	Asp	Pro	Gln	Arg	Phe	Val	Val	Ile	Gln	Asn	Glu
		980						985					990		
Asp	Leu	Gly	Pro	Ala	Ser	Pro	Leu	Asp	Ser	Thr	Phe	Tyr	Arg	Ser	Leu
	995						1000					1005			
Leu	Glu	Asp	Asp	Asp	Met	Gly	Asp	Leu	Val	Asp	Ala	Glu	Glu	Tyr	Leu
	1010					1015				1020					
Val	Pro	Gln	Gln	Gly	Phe	Phe	Cys	Pro	Asp	Pro	Ala	Pro	Gly	Ala	Gly
1025					1030				1035						1040
Gly	Met	Val	His	His	Arg	His	Arg	Ser	Ser	Ser	Thr	Arg	Ser	Gly	Gly
			1045						1050					1055	
Gly	Asp	Leu	Thr	Leu	Gly	Leu	Glu	Pro	Ser	Glu	Glu	Glu	Ala	Pro	Arg
		1060						1065					1070		
Ser	Pro	Leu	Ala	Pro	Ser	Glu	Gly	Ala	Gly	Ser	Asp	Val	Phe	Asp	Gly
	1075						1080				1085				
Asp	Leu	Gly	Met	Gly	Ala	Ala	Lys	Gly	Leu	Gln	Ser	Leu	Pro	Thr	His
	1090					1095					1100				
Asp	Pro	Ser	Pro	Leu	Gln	Arg	Tyr	Ser	Glu	Asp	Pro	Thr	Val	Pro	Leu
	1105				1110					1115					1120
Pro	Ser	Glu	Thr	Asp	Gly	Tyr	Val	Ala	Pro	Leu	Thr	Cys	Ser	Pro	Gln
			1125						1130					1135	
Pro	Glu	Tyr	Val	Asn	Gln	Pro	Asp	Val	Arg	Pro	Gln	Pro	Pro	Ser	Pro

Gln	Cys	Cys	His	Asp	Glu	Cys	Ala	Gly	Gly	Cys	Ser	Gly	Pro	Gln	Asp	225	230	235	240
Thr	Asp	Cys	Phe	Ala	Cys	Arg	His	Phe	Asn	Asp	Ser	Gly	Ala	Cys	Val		245	250	255
Pro	Arg	Cys	Pro	Gln	Pro	Leu	Val	Tyr	Asn	Lys	Leu	Thr	Phe	Gln	Leu		260	265	270
Glu	Pro	Asn	Pro	His	Thr	Lys	Tyr	Gln	Tyr	Gly	Gly	Val	Cys	Val	Ala		275	280	285
Ser	Cys	Pro	His	Asn	Phe	Val	Val	Asp	Gln	Thr	Ser	Cys	Val	Arg	Ala		290	295	300
Cys	Pro	Pro	Asp	Lys	Met	Glu	Val	Asp	Lys	Asn	Gly	Leu	Lys	Met	Cys		305	310	315
Glu	Pro	Cys	Gly	Gly	Leu	Cys	Pro	Lys	Ala	Cys	Glu	Gly	Thr	Gly	Ser		320	325	330
Gly	Ser	Arg	Phe	Gln	Thr	Val	Asp	Ser	Ser	Asn	Ile	Asp	Gly	Phe	Val		335	340	345
Asn	Cys	Thr	Lys	Ile	Leu	Gly	Asn	Leu	Asp	Phe	Leu	Ile	Thr	Gly	Leu		350	355	360
Asn	Gly	Asp	Pro	Trp	His	Lys	Ile	Pro	Ala	Leu	Asp	Pro	Glu	Lys	Leu		365	370	375
Asn	Val	Phe	Arg	Thr	Val	Arg	Glu	Ile	Thr	Gly	Tyr	Leu	Asn	Ile	Gln		380	385	390
Ser	Trp	Pro	Pro	His	Met	His	Asn	Phe	Ser	Val	Phe	Ser	Asn	Leu	Thr		395	400	405
Thr	Ile	Gly	Gly	Arg	Ser	Leu	Tyr	Asn	Arg	Gly	Phe	Ser	Leu	Leu	Ile		410	415	420
Met	Lys	Asn	Leu	Asn	Val	Thr	Ser	Leu	Gly	Phe	Arg	Ser	Leu	Lys	Glu		425	430	435
Ile	Ser	Ala	Gly	Arg	Ile	Tyr	Ile	Ser	Ala	Asn	Arg	Gln	Leu	Cys	Tyr		440	445	450
His	His	Ser	Leu	Asn	Trp	Thr	Lys	Val	Leu	Arg	Gly	Pro	Thr	Glu	Glu		455	460	465
Arg	Leu	Asp	Ile	Lys	His	Asn	Arg	Pro	Arg	Arg	Asp	Cys	Val	Ala	Glu		470	475	480
Gly	Lys	Val	Cys	Asp	Pro	Leu	Cys	Ser	Ser	Gly	Gly	Cys	Trp	Gly	Pro		485	490	495
Gly	Pro	Gly	Gln	Cys	Leu	Ser	Cys	Arg	Asn	Tyr	Ser	Arg	Gly	Gly	Val		500	505	510
Cys	Val	Thr	His	Cys	Asn	Phe	Leu	Asn	Gly	Glu	Pro	Arg	Glu	Phe	Ala		515	520	525
His	Glu	Ala	Glu	Cys	Phe	Ser	Cys	His	Pro	Glu	Cys	Gln	Pro	Met	Glu		530	535	540
Gly	Thr	Ala	Thr	Cys	Asn	Gly	Ser	Gly	Ser	Asp	Thr	Cys	Ala	Gln	Cys		545	550	555
Ala	His	Phe	Arg	Asp	Gly	Pro	His	Cys	Val	Ser	Ser	Cys	Pro	His	Gly		560	565	570
Val	Leu	Gly	Ala	Lys	Gly	Pro	Ile	Tyr	Lys	Tyr	Pro	Asp	Val	Gln	Asn		575	580	585
Glu	Cys	Arg	Pro	Cys	His	Glu	Asn	Cys	Thr	Gln	Gly	Cys	Lys	Gly	Pro		590	595	600
Glu	Leu	Gln	Asp	Cys	Leu	Gly	Gln	Thr	Leu	Val	Leu	Ile	Gly	Lys	Thr		605	610	615
His	Leu	Thr	Met	Ala	Leu	Thr	Val	Ile	Ala	Gly	Leu	Val	Val	Ile	Phe		620	625	630
																	635	640	645
																	650	655	

Met	Met	Leu	Gly	Gly	Thr	Phe	Leu	Tyr	Trp	Arg	Gly	Arg	Arg	Ile	Gln	660	665	670
Asn	Lys	Arg	Ala	Met	Arg	Arg	Tyr	Leu	Glu	Arg	Gly	Glu	Ser	Ile	Glu	675	680	685
Pro	Leu	Asp	Pro	Ser	Glu	Lys	Ala	Asn	Lys	Val	Leu	Ala	Arg	Ile	Phe	690	695	700
Lys	Glu	Thr	Glu	Leu	Arg	Lys	Leu	Lys	Val	Leu	Gly	Ser	Gly	Val	Phe	705	710	715
Gly	Thr	Val	His	Lys	Gly	Val	Trp	Ile	Pro	Glu	Gly	Glu	Ser	Ile	Lys	725	730	735
Ile	Pro	Val	Cys	Ile	Lys	Val	Ile	Glu	Asp	Lys	Ser	Gly	Arg	Gln	Ser	740	745	750
Phe	Gln	Ala	Val	Thr	Asp	His	Met	Leu	Ala	Ile	Gly	Ser	Leu	Asp	His	755	760	765
Ala	His	Ile	Val	Arg	Leu	Leu	Gly	Leu	Cys	Pro	Gly	Ser	Ser	Leu	Gln	770	775	780
Leu	Val	Thr	Gln	Tyr	Leu	Pro	Leu	Gly	Ser	Leu	Leu	Asp	His	Val	Arg	785	790	795
Gln	His	Arg	Gly	Ala	Leu	Gly	Pro	Gln	Leu	Leu	Leu	Asn	Trp	Gly	Val	805	810	815
Gln	Ile	Ala	Lys	Gly	Met	Tyr	Tyr	Leu	Glu	Glu	His	Gly	Met	Val	His	820	825	830
Arg	Asn	Leu	Ala	Ala	Arg	Asn	Val	Leu	Leu	Lys	Ser	Pro	Ser	Gln	Val	835	840	845
Gln	Val	Ala	Asp	Phe	Gly	Val	Ala	Asp	Leu	Leu	Pro	Pro	Asp	Asp	Lys	850	855	860
Gln	Leu	Leu	Tyr	Ser	Glu	Ala	Lys	Thr	Pro	Ile	Lys	Trp	Met	Ala	Leu	865	870	875
Glu	Ser	Ile	His	Phe	Gly	Lys	Tyr	Thr	His	Gln	Ser	Asp	Val	Trp	Ser	885	890	895
Tyr	Gly	Val	Thr	Val	Trp	Glu	Leu	Met	Thr	Phe	Gly	Ala	Glu	Pro	Tyr	900	905	910
Ala	Gly	Leu	Arg	Leu	Ala	Glu	Val	Pro	Asp	Leu	Leu	Glu	Lys	Gly	Glu	915	920	925
Arg	Leu	Ala	Gln	Pro	Gln	Ile	Cys	Thr	Ile	Asp	Val	Tyr	Met	Val	Met	930	935	940
Val	Lys	Cys	Trp	Met	Ile	Asp	Glu	Asn	Ile	Arg	Pro	Thr	Phe	Lys	Glu	945	950	955
Leu	Ala	Asn	Glu	Phe	Thr	Arg	Met	Ala	Arg	Asp	Pro	Pro	Arg	Tyr	Leu	965	970	975
Val	Ile	Lys	Arg	Glu	Ser	Gly	Pro	Gly	Ile	Ala	Pro	Gly	Pro	Glu	Pro	980	985	990
His	Gly	Leu	Thr	Asn	Lys	Lys	Leu	Glu	Glu	Val	Glu	Leu	Glu	Pro	Glu	995	1000	1005
Leu	Asp	Leu	Asp	Leu	Asp	Leu	Glu	Ala	Glu	Glu	Asp	Asn	Leu	Ala	Thr	1010	1015	1020
Thr	Thr	Leu	Gly	Ser	Ala	Leu	Ser	Leu	Pro	Val	Gly	Thr	Leu	Asn	Arg	1025	1030	1035
Pro	Arg	Gly	Ser	Gln	Ser	Leu	Leu	Ser	Pro	Ser	Ser	Gly	Tyr	Met	Pro	1045	1050	1055
Met	Asn	Gln	Gly	Asn	Leu	Gly	Glu	Ser	Cys	Gln	Glu	Ser	Ala	Val	Ser	1060	1065	1070
Gly	Ser	Ser	Glu	Arg	Cys	Pro	Arg	Pro	Val	Ser	Leu	His	Pro	Met	Pro	1075	1080	1085

Arg Gly Cys Leu Ala Ser Glu Ser Ser Glu Gly His Val Thr Gly Ser
 1090 1095 1100
 Glu Ala Glu Leu Gln Glu Lys Val Ser Met Cys Arg Ser Arg Ser Arg
 1105 1110 1115 1120
 Ser Arg Ser Pro Arg Pro Arg Gly Asp Ser Ala Tyr His Ser Gln Arg
 1125 1130 1135
 His Ser Leu Leu Thr Pro Val Thr Pro Leu Ser Pro Pro Gly Leu Glu
 1140 1145 1150
 Glu Glu Asp Val Asn Gly Tyr Val Met Pro Asp Thr His Leu Lys Gly
 1155 1160 1165
 Thr Pro Ser Ser Arg Glu Gly Thr Leu Ser Ser Val Gly Leu Ser Ser
 1170 1175 1180
 Val Leu Gly Thr Glu Glu Glu Asp Glu Asp Glu Glu Tyr Glu Tyr Met
 1185 1190 1195 1200
 Asn Arg Arg Arg Arg His Ser Pro Pro His Pro Pro Arg Pro Ser Ser
 1205 1210 1215
 Leu Glu Glu Leu Gly Tyr Glu Tyr Met Asp Val Gly Ser Asp Leu Ser
 1220 1225 1230
 Ala Ser Leu Gly Ser Thr Gln Ser Cys Pro Leu His Pro Val Pro Ile
 1235 1240 1245
 Met Pro Thr Ala Gly Thr Thr Pro Asp Glu Asp Tyr Glu Tyr Met Asn
 1250 1255 1260
 Arg Gln Arg Asp Gly Gly Gly Pro Gly Gly Asp Tyr Ala Ala Met Gly
 1265 1270 1275 1280
 Ala Cys Pro Ala Ser Glu Gln Gly Tyr Glu Glu Met Arg Ala Phe Gln
 1285 1290 1295
 Gly Pro Gly His Gln Ala Pro His Val His Tyr Ala Arg Leu Lys Thr
 1300 1305 1310
 Leu Arg Ser Leu Glu Ala Thr Asp Ser Ala Phe Asp Asn Pro Asp Tyr
 1315 1320 1325
 Trp His Ser Arg Leu Phe Pro Lys Ala Asn Ala Gln Arg Thr
 1330 1335 1340

<210> 140

<211> 760

<212> PRT

<213> Homo sapiens

<220>

<223> >ERCC2|ENSG00000104884|ENST00000391945|2283

<400> 140

Met Lys Leu Asn Val Asp Gly Leu Leu Val Tyr Phe Pro Tyr Asp Tyr
 1 5 10 15
 Ile Tyr Pro Glu Gln Phe Ser Tyr Met Arg Glu Leu Lys Arg Thr Leu
 20 25 30
 Asp Ala Lys Gly His Gly Val Leu Glu Met Pro Ser Gly Thr Gly Lys
 35 40 45
 Thr Val Ser Leu Leu Ala Leu Ile Met Ala Tyr Gln Arg Ala Tyr Pro
 50 55 60
 Leu Glu Val Thr Lys Leu Ile Tyr Cys Ser Arg Thr Val Pro Glu Ile
 65 70 75 80
 Glu Lys Val Ile Glu Glu Leu Arg Lys Leu Leu Asn Phe Tyr Glu Lys

Gln	Asp	Phe	Leu	Val	Ala	Ile	Ala	Glu	Pro	Val	Cys	Arg	Pro	Thr	His	
			100					105					110			
Val	His	Glu	Tyr	Lys	Leu	Thr	Ala	Tyr	Ser	Leu	Tyr	Ala	Ala	Val	Ser	
		115					120					125				
Val	Gly	Leu	Gln	Thr	Ser	Asp	Ile	Thr	Glu	Tyr	Leu	Arg	Lys	Leu	Ser	
	130					135					140					
Lys	Thr	Gly	Val	Pro	Asp	Gly	Ile	Met	Gln	Phe	Ile	Lys	Leu	Cys	Thr	
145					150					155					160	
Val	Ser	Tyr	Gly	Lys	Val	Lys	Leu	Val	Leu	Lys	His	Asn	Arg	Tyr	Phe	
			165					170						175		
Val	Glu	Ser	Cys	His	Pro	Asp	Val	Ile	Gln	His	Leu	Leu	Gln	Asp	Pro	
			180					185					190			
Val	Ile	Arg	Glu	Cys	Arg	Leu	Arg	Asn	Ser	Glu	Gly	Glu	Ala	Thr	Glu	
		195					200					205				
Leu	Ile	Thr	Glu	Thr	Phe	Thr	Ser	Lys	Ser	Ala	Ile	Ser	Lys	Thr	Ala	
	210					215					220					
Glu	Ser	Ser	Gly	Gly	Pro	Ser	Thr	Ser	Arg	Val	Thr	Asp	Pro	Gln	Gly	
225					230					235					240	
Lys	Ser	Asp	Ile	Pro	Met	Asp	Leu	Phe	Asp	Phe	Tyr	Glu	Gln	Met	Asp	
			245					250						255		
Lys	Asp	Glu	Glu	Glu	Glu	Glu	Glu	Thr	Gln	Thr	Val	Ser	Phe	Glu	Val	
		260						265					270			
Lys	Gln	Glu	Met	Ile	Glu	Glu	Leu	Gln	Lys	Arg	Cys	Ile	His	Leu	Glu	
	275						280					285				
Tyr	Pro	Leu	Leu	Ala	Glu	Tyr	Asp	Phe	Arg	Asn	Asp	Ser	Val	Asn	Pro	
	290					295					300					
Asp	Ile	Asn	Ile	Asp	Leu	Lys	Pro	Thr	Ala	Val	Leu	Arg	Pro	Tyr	Gln	
305					310					315					320	
Glu	Lys	Ser	Leu	Arg	Lys	Met	Phe	Gly	Asn	Gly	Arg	Ala	Arg	Ser	Gly	
			325					330						335		
Val	Ile	Val	Leu	Pro	Cys	Gly	Ala	Gly	Lys	Ser	Leu	Val	Gly	Val	Thr	
		340						345					350			
Ala	Ala	Cys	Thr	Val	Arg	Lys	Arg	Cys	Leu	Val	Leu	Gly	Asn	Ser	Ala	
		355					360					365				
Val	Ser	Val	Glu	Gln	Trp	Lys	Ala	Gln	Phe	Lys	Met	Trp	Ser	Thr	Ile	
	370					375					380					
Asp	Asp	Ser	Gln	Ile	Cys	Arg	Phe	Thr	Ser	Asp	Ala	Lys	Asp	Lys	Pro	
385					390					395					400	
Ile	Gly	Cys	Ser	Val	Ala	Ile	Ser	Thr	Tyr	Ser	Met	Leu	Gly	His	Thr	
			405					410						415		
Thr	Lys	Arg	Ser	Trp	Glu	Ala	Glu	Arg	Val	Met	Glu	Trp	Leu	Lys	Thr	
		420						425					430			
Gln	Glu	Trp	Gly	Leu	Met	Ile	Leu	Asp	Glu	Val	His	Thr	Ile	Pro	Ala	
	435						440					445				
Lys	Met	Phe	Arg	Arg	Val	Leu	Thr	Ile	Val	Gln	Ala	His	Cys	Lys	Leu	
	450					455					460					
Gly	Leu	Thr	Ala	Thr	Leu	Val	Arg	Glu	Asp	Asp	Lys	Ile	Val	Asp	Leu	
465					470					475					480	
Asn	Phe	Leu	Ile	Gly	Pro	Lys	Leu	Tyr	Glu	Ala	Asn	Trp	Met	Glu	Leu	
			485					490						495		
Gln	Asn	Asn	Gly	Tyr	Ile	Ala	Lys	Val	Gln	Cys	Ala	Glu	Val	Trp	Cys	
		500						505					510			
Pro	Met	Ser	Pro	Glu	Phe	Tyr	Arg	Glu	Tyr	Val	Ala	Ile	Lys	Thr	Lys	
		515					520					525				

Lys	Arg	Ile	Leu	Leu	Tyr	Thr	Met	Asn	Pro	Asn	Lys	Phe	Arg	Ala	Cys
530						535					540				
Gln	Phe	Leu	Ile	Lys	Phe	His	Glu	Arg	Arg	Asn	Asp	Lys	Ile	Ile	Val
545					550					555					560
Phe	Ala	Asp	Asn	Val	Phe	Ala	Leu	Lys	Glu	Tyr	Ala	Ile	Arg	Leu	Asn
			565						570					575	
Lys	Pro	Tyr	Ile	Tyr	Gly	Pro	Thr	Ser	Gln	Gly	Glu	Arg	Met	Gln	Ile
		580						585					590		
Leu	Gln	Asn	Phe	Lys	His	Asn	Pro	Lys	Ile	Asn	Thr	Ile	Phe	Ile	Ser
	595					600						605			
Lys	Val	Gly	Asp	Thr	Ser	Phe	Asp	Leu	Pro	Glu	Ala	Asn	Val	Leu	Ile
610						615					620				
Gln	Ile	Ser	Ser	His	Gly	Gly	Ser	Arg	Arg	Gln	Glu	Ala	Gln	Arg	Leu
625					630					635					640
Gly	Arg	Val	Leu	Arg	Ala	Lys	Lys	Gly	Met	Val	Ala	Glu	Glu	Tyr	Asn
			645						650					655	
Ala	Phe	Phe	Tyr	Ser	Leu	Val	Ser	Gln	Asp	Thr	Gln	Glu	Met	Ala	Tyr
			660					665					670		
Ser	Thr	Lys	Arg	Gln	Arg	Phe	Leu	Val	Asp	Gln	Gly	Tyr	Ser	Phe	Lys
		675					680					685			
Val	Ile	Thr	Lys	Leu	Ala	Gly	Met	Glu	Glu	Glu	Asp	Leu	Ala	Phe	Ser
690						695					700				
Thr	Lys	Glu	Glu	Gln	Gln	Gln	Leu	Leu	Gln	Lys	Val	Leu	Ala	Ala	Thr
705					710					715					720
Asp	Leu	Asp	Ala	Glu	Glu	Glu	Val	Val	Ala	Gly	Glu	Phe	Gly	Ser	Arg
			725						730					735	
Ser	Ser	Gln	Ala	Ser	Arg	Arg	Phe	Gly	Thr	Met	Ser	Ser	Met	Ser	Gly
		740						745					750		
Ala	Asp	Asp	Thr	Val	Tyr	Met	Glu	Tyr	His	Ser	Ser	Arg	Ser	Lys	Ala
	755					760						765			
Pro	Ser	Lys	His	Val	His	Pro	Leu	Phe	Lys	Arg	Phe	Arg	Lys		
770						775					780				

<210> 142
 <211> 916
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >ERCC4|ENSG00000175595|ENST00000311895|2751

Met	Glu	Ser	Gly	Gln	Pro	Ala	Arg	Arg	Ile	Ala	Met	Ala	Pro	Leu	Leu
1				5					10					15	
Glu	Tyr	Glu	Arg	Gln	Leu	Val	Leu	Glu	Leu	Leu	Asp	Thr	Asp	Gly	Leu
			20					25					30		
Val	Val	Cys	Ala	Arg	Gly	Leu	Gly	Ala	Asp	Arg	Leu	Leu	Tyr	His	Phe
		35					40					45			
Leu	Gln	Leu	His	Cys	His	Pro	Ala	Cys	Leu	Val	Leu	Val	Leu	Asn	Thr
	50					55					60				
Gln	Pro	Ala	Glu	Glu	Glu	Tyr	Phe	Ile	Asn	Gln	Leu	Lys	Ile	Glu	Gly
65					70					75					80
Val	Glu	His	Leu	Pro	Arg	Arg	Val	Thr	Asn	Glu	Ile	Thr	Ser	Asn	Ser

Page 439

<212> PRT
 <213> Homo sapiens

<220>
 <223> >ERCC5|ENSG00000134899|ENST00000355739|3561

<400> 143
 Met Gly Val Gln Gly Leu Trp Lys Leu Leu Glu Cys Ser Gly Arg Gln
 1 5 10 15
 Val Ser Pro Glu Ala Leu Glu Gly Lys Ile Leu Ala Val Asp Ile Ser
 20 25 30
 Ile Trp Leu Asn Gln Ala Leu Lys Gly Val Arg Asp Arg His Gly Asn
 35 40 45
 Ser Ile Glu Asn Pro His Leu Leu Thr Leu Phe His Arg Leu Cys Lys
 50 55 60
 Leu Leu Phe Phe Arg Ile Arg Pro Ile Phe Val Phe Asp Gly Asp Ala
 65 70 75 80
 Pro Leu Leu Lys Lys Gln Thr Leu Val Lys Arg Arg Gln Arg Lys Asp
 85 90 95
 Leu Ala Ser Ser Asp Ser Arg Lys Thr Thr Glu Lys Leu Leu Lys Thr
 100 105 110
 Phe Leu Lys Arg Gln Ala Ile Lys Thr Ala Phe Arg Ser Lys Arg Asp
 115 120 125
 Glu Ala Leu Pro Ser Leu Thr Gln Val Arg Arg Glu Asn Asp Leu Tyr
 130 135 140
 Val Leu Pro Pro Leu Gln Glu Glu Glu Lys His Ser Ser Glu Glu Glu
 145 150 155 160
 Asp Glu Lys Glu Trp Gln Glu Arg Met Asn Gln Lys Gln Ala Leu Gln
 165 170 175
 Glu Glu Phe Phe His Asn Pro Gln Ala Ile Asp Ile Glu Ser Glu Asp
 180 185 190
 Phe Ser Ser Leu Pro Pro Glu Val Lys His Glu Ile Leu Thr Asp Met
 195 200 205
 Lys Glu Phe Thr Lys Arg Arg Arg Thr Leu Phe Glu Ala Met Pro Glu
 210 215 220
 Glu Ser Asp Asp Phe Ser Gln Tyr Gln Leu Lys Gly Leu Leu Lys Lys
 225 230 235 240
 Asn Tyr Leu Asn Gln His Ile Glu His Val Gln Lys Glu Met Asn Gln
 245 250 255
 Gln His Ser Gly His Ile Arg Arg Gln Tyr Glu Asp Glu Gly Gly Phe
 260 265 270
 Leu Lys Glu Val Glu Ser Arg Arg Val Val Ser Glu Asp Thr Ser His
 275 280 285
 Tyr Ile Leu Ile Lys Gly Ile Gln Ala Lys Thr Val Ala Glu Val Asp
 290 295 300
 Ser Glu Ser Leu Pro Ser Ser Ser Lys Met His Gly Met Ser Phe Asp
 305 310 315 320
 Val Lys Ser Ser Pro Cys Glu Lys Leu Lys Thr Glu Lys Glu Pro Asp
 325 330 335
 Ala Thr Pro Pro Ser Pro Arg Thr Leu Leu Ala Met Gln Ala Ala Leu
 340 345 350
 Leu Gly Ser Ser Ser Glu Glu Glu Leu Glu Ser Glu Asn Arg Arg Gln
 355 360 365

Ala 370	Arg 375	Gly 380	Arg 385	Asn 390	Ala 395	Pro 400	Ala 405	Ala 410	Val 415	Asp 420	Glu 425	Gly 430	Ser 435	Ile 440	Ser 445
Pro 385	Arg 390	Thr 395	Leu 400	Ser 405	Ala 410	Ile 415	Lys 420	Arg 425	Ala 430	Leu 435	Asp 440	Asp 445	Asp 450	Glu 455	Asp 460
Val 380	Lys 385	Val 390	Cys 395	Ala 400	Gly 405	Asp 410	Asp 415	Val 420	Gln 425	Thr 430	Gly 435	Gly 440	Pro 445	Gly 450	Ala 455
Glu 375	Glu 380	Met 385	Arg 390	Ile 395	Asn 400	Ser 405	Ser 410	Thr 415	Glu 420	Asn 425	Ser 430	Asp 435	Glu 440	Gly 445	Leu 450
Lys 370	Val 375	Arg 380	Asp 385	Gly 390	Lys 395	Gly 400	Ile 405	Pro 410	Phe 415	Thr 420	Ala 425	Thr 430	Leu 435	Ala 440	Ser 445
Ser 365	Ser 370	Val 375	Asn 380	Ser 385	Ala 390	Glu 395	Glu 400	His 405	Val 410	Ala 415	Ser 420	Thr 425	Asn 430	Glu 435	Gly 440
Arg 360	Glu 365	Pro 370	Thr 375	Asp 380	Ser 385	Val 390	Pro 395	Lys 400	Glu 405	Gln 410	Met 415	Ser 420	Leu 425	Val 430	His 435
Val 355	Gly 360	Thr 365	Glu 370	Ala 375	Phe 380	Pro 385	Ile 390	Ser 395	Asp 400	Glu 405	Ser 410	Met 415	Ile 420	Lys 425	Asp 430
Arg 350	Lys 355	Asp 360	Arg 365	Leu 370	Pro 375	Leu 380	Glu 385	Ser 390	Ala 395	Val 400	Val 405	Arg 410	His 415	Ser 420	Asp 425
Ala 345	Pro 350	Gly 355	Leu 360	Pro 365	Asn 370	Gly 375	Arg 380	Glu 385	Leu 390	Thr 395	Pro 400	Ala 405	Ser 410	Pro 415	Thr 420
Cys 340	Thr 345	Asn 350	Ser 355	Val 360	Ser 365	Lys 370	Asn 375	Glu 380	Thr 385	His 390	Ala 395	Glu 400	Val 405	Leu 410	Glu 415
Gln 335	Gln 340	Asn 345	Glu 350	Leu 355	Cys 360	Pro 365	Tyr 370	Glu 375	Ser 380	Lys 385	Phe 390	Asp 395	Ser 400	Ser 405	Leu 410
Leu 330	Ser 335	Ser 340	Asp 345	Asp 350	Glu 355	Thr 360	Lys 365	Cys 370	Lys 375	Pro 380	Asn 385	Ser 390	Ala 395	Ser 400	Glu 405
Val 325	Ile 330	Gly 335	Pro 340	Val 345	Ser 350	Leu 355	Gln 360	Glu 365	Thr 370	Ser 375	Ser 380	Ile 385	Val 390	Ser 395	Val 400
Pro 320	Ser 325	Glu 330	Ala 335	Val 340	Asp 345	Asn 350	Val 355	Glu 360	Asn 365	Val 370	Val 375	Ser 380	Phe 385	Asn 390	Ala 395
Lys 315	Glu 320	His 325	Glu 330	Asn 335	Phe 340	Leu 345	Glu 350	Thr 355	Ile 360	Gln 365	Glu 370	Gln 375	Gln 380	Thr 385	Thr 390
Glu 310	Ser 315	Ala 320	Gly 325	Gln 330	Asp 335	Leu 340	Ile 345	Ser 350	Ile 355	Pro 360	Lys 365	Ala 370	Val 375	Glu 380	Pro 385
Met 305	Glu 310	Ile 315	Asp 320	Ser 325	Glu 330	Glu 335	Ser 340	Glu 345	Ser 350	Asp 355	Gly 360	Ser 365	Phe 370	Ile 375	Glu 380
Val 300	Gln 305	Ser 310	Val 315	Ile 320	Ser 325	Asp 330	Glu 335	Glu 340	Leu 345	Gln 350	Ala 355	Glu 360	Phe 365	Pro 370	Glu 375
Thr 295	Ser 300	Lys 305	Pro 310	Pro 315	Ser 320	Glu 325	Gln 330	Gly 335	Glu 340	Glu 345	Glu 350	Leu 355	Val 360	Gly 365	Thr 370
Arg 290	Glu 295	Gly 300	Glu 305	Ala 310	Pro 315	Ala 320	Glu 325	Ser 330	Glu 335	Ser 340	Leu 345	Leu 350	Arg 355	Asp 360	Asn 365
Ser 285	Glu 290	Arg 295	Asp 300	Asp 305	Val 310	Asp 315	Gly 320	Glu 325	Pro 330	Gln 335	Glu 340	Ala 345	Glu 350	Lys 355	Asp 360
Ala 280	Glu 285	Asp 290	Ser 295	Leu 300	His 305	Glu 310	Trp 315	Gln 320	Asp 325	Ile 330	Asn 335	Leu 340	Glu 345	Glu 350	Leu 355
Glu 275	Thr 280	Leu 285	Glu 290	Ser 295	Asn 300	Leu 305	Leu 310	Ala 315	Gln 320	Gln 325	Asn 330	Ser 335	Leu 340	Lys 345	Ala 350
Gln 270	Lys 275	Gln 280	Gln 285	Gln 290	G										

<213> Homo sapiens

<220>

<223> >FAM175A|ENSG00000163322|ENST00000321945|1230

<400> 144

Met	Glu	Gly	Glu	Ser	Thr	Ser	Ala	Val	Leu	Ser	Gly	Phe	Val	Leu	Gly
1				5					10					15	
Ala	Leu	Ala	Phe	Gln	His	Leu	Asn	Thr	Asp	Ser	Asp	Thr	Glu	Gly	Phe
			20					25					30		
Leu	Leu	Gly	Glu	Val	Lys	Gly	Glu	Ala	Lys	Asn	Ser	Ile	Thr	Asp	Ser
		35					40					45			
Gln	Met	Asp	Asp	Val	Glu	Val	Val	Tyr	Thr	Ile	Asp	Ile	Gln	Lys	Tyr
	50					55					60				
Ile	Pro	Cys	Tyr	Gln	Leu	Phe	Ser	Phe	Tyr	Asn	Ser	Ser	Gly	Glu	Val
65					70					75					80
Asn	Glu	Gln	Ala	Leu	Lys	Lys	Ile	Leu	Ser	Asn	Val	Lys	Lys	Asn	Val
				85					90					95	
Val	Gly	Trp	Tyr	Lys	Phe	Arg	Arg	His	Ser	Asp	Gln	Ile	Met	Thr	Phe
			100					105					110		
Arg	Glu	Arg	Leu	Leu	His	Lys	Asn	Leu	Gln	Glu	His	Phe	Ser	Asn	Gln
		115					120					125			
Asp	Leu	Val	Phe	Leu	Leu	Leu	Thr	Pro	Ser	Ile	Ile	Thr	Glu	Ser	Cys
	130					135					140				
Ser	Thr	His	Arg	Leu	Glu	His	Ser	Leu	Tyr	Lys	Pro	Gln	Lys	Gly	Leu
145					150					155					160
Phe	His	Arg	Val	Pro	Leu	Val	Val	Ala	Asn	Leu	Gly	Met	Ser	Glu	Gln
				165				170						175	
Leu	Gly	Tyr	Lys	Thr	Val	Ser	Gly	Ser	Cys	Met	Ser	Thr	Gly	Phe	Ser
			180					185					190		
Arg	Ala	Val	Gln	Thr	His	Ser	Ser	Lys	Phe	Phe	Glu	Glu	Asp	Gly	Ser
		195					200					205			
Leu	Lys	Glu	Val	His	Lys	Ile	Asn	Glu	Met	Tyr	Ala	Ser	Leu	Gln	Glu
	210					215					220				
Glu	Leu	Lys	Ser	Ile	Cys	Lys	Lys	Val	Glu	Asp	Ser	Glu	Gln	Ala	Val
225					230					235					240
Asp	Lys	Leu	Val	Lys	Asp	Val	Asn	Arg	Leu	Lys	Arg	Glu	Ile	Glu	Lys
				245				250						255	
Arg	Arg	Gly	Ala	Gln	Ile	Gln	Ala	Ala	Arg	Glu	Lys	Asn	Ile	Gln	Lys
		260					265					270			
Asp	Pro	Gln	Glu	Asn	Ile	Phe	Leu	Cys	Gln	Ala	Leu	Arg	Thr	Phe	Phe
		275					280					285			
Pro	Asn	Ser	Glu	Phe	Leu	His	Ser	Cys	Val	Met	Ser	Leu	Lys	Asn	Arg
	290					295					300				
His	Val	Ser	Lys	Ser	Ser	Cys	Asn	Tyr	Asn	His	His	Leu	Asp	Val	Val
305					310					315					320
Asp	Asn	Leu	Thr	Leu	Met	Val	Glu	His	Thr	Asp	Ile	Pro	Glu	Ala	Ser
				325					330					335	
Pro	Ala	Ser	Thr	Pro	Gln	Ile	Ile	Lys	His	Lys	Ala	Leu	Asp	Leu	Asp
			340					345					350		
Asp	Arg	Trp	Gln	Phe	Lys	Arg	Ser	Arg	Leu	Leu	Asp	Thr	Gln	Asp	Lys
		355					360					365			
Arg	Ser	Lys	Ala	Asp	Thr	Gly	Ser	Ser	Asn	Gln	Asp	Lys	Ala	Ser	Lys

370	375	380
Met Ser Ser Pro Glu Thr Asp Glu Glu Ile Glu Lys Met Lys Gly Phe		
385	390	395
Gly Glu Tyr Ser Arg Ser Pro Thr Phe		400
405		

<210> 145
 <211> 1455
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >FANCA|ENSG00000187741|ENST00000389301|4368

<400> 145
Met Ser Asp Ser Trp Val Pro Asn Ser Ala Ser Gly Gln Asp Pro Gly
1 5 10 15
Gly Arg Arg Arg Ala Trp Ala Glu Leu Leu Ala Gly Arg Val Lys Arg
20 25 30
Glu Lys Tyr Asn Pro Glu Arg Ala Gln Lys Leu Lys Glu Ser Ala Val
35 40 45
Arg Leu Leu Arg Ser His Gln Asp Leu Asn Ala Leu Leu Glu Val
50 55 60
Glu Gly Pro Leu Cys Lys Lys Leu Ser Leu Ser Lys Val Ile Asp Cys
65 70 75 80
Asp Ser Ser Glu Ala Tyr Ala Asn His Ser Ser Ser Phe Ile Gly Ser
85 90 95
Ala Leu Gln Asp Gln Ala Ser Arg Leu Gly Val Pro Val Gly Ile Leu
100 105 110
Ser Ala Gly Met Val Ala Ser Ser Val Gly Gln Ile Cys Thr Ala Pro
115 120 125
Ala Glu Thr Ser His Pro Val Leu Leu Thr Val Glu Gln Arg Lys Lys
130 135 140
Leu Ser Ser Leu Leu Glu Phe Ala Gln Tyr Leu Leu Ala His Ser Met
145 150 155 160
Phe Ser Arg Leu Ser Phe Cys Gln Glu Leu Trp Lys Ile Gln Ser Ser
165 170 175
Leu Leu Leu Glu Ala Val Trp His Leu His Val Gln Gly Ile Val Ser
180 185 190
Leu Gln Glu Leu Leu Glu Ser His Pro Asp Met His Ala Val Gly Ser
195 200 205
Trp Leu Phe Arg Asn Leu Cys Cys Leu Cys Glu Gln Met Glu Ala Ser
210 215 220
Cys Gln His Ala Asp Val Ala Arg Ala Met Leu Ser Asp Phe Val Gln
225 230 235 240
Met Phe Val Leu Arg Gly Phe Gln Lys Asn Ser Asp Leu Arg Arg Thr
245 250 255
Val Glu Pro Glu Lys Met Pro Gln Val Thr Val Asp Val Leu Gln Arg
260 265 270
Met Leu Ile Phe Ala Leu Asp Ala Leu Ala Ala Gly Val Gln Glu Glu
275 280 285
Ser Ser Thr His Lys Ile Val Arg Cys Trp Phe Gly Val Phe Ser Gly
290 295 300

His	Thr	Leu	Gly	Ser	Val	Ile	Ser	Thr	Asp	Pro	Leu	Lys	Arg	Phe	Phe	305	310	315	320
Ser	His	Thr	Leu	Thr	Gln	Ile	Leu	Thr	His	Ser	Pro	Val	Leu	Lys	Ala	325	330	335	
Ser	Asp	Ala	Val	Gln	Met	Gln	Arg	Glu	Trp	Ser	Phe	Ala	Arg	Thr	His	340	345	350	
Pro	Leu	Leu	Thr	Ser	Leu	Tyr	Arg	Arg	Leu	Phe	Val	Met	Leu	Ser	Ala	355	360	365	
Glu	Glu	Leu	Val	Gly	His	Leu	Gln	Glu	Val	Leu	Glu	Thr	Gln	Glu	Val	370	375	380	
His	Trp	Gln	Arg	Val	Leu	Ser	Phe	Val	Ser	Ala	Leu	Val	Val	Cys	Phe	385	390	395	400
Pro	Glu	Ala	Gln	Gln	Leu	Leu	Glu	Asp	Trp	Val	Ala	Arg	Leu	Met	Ala	405	410	415	
Gln	Ala	Phe	Glu	Ser	Cys	Gln	Leu	Asp	Ser	Met	Val	Thr	Ala	Phe	Leu	420	425	430	
Val	Val	Arg	Gln	Ala	Ala	Leu	Glu	Gly	Pro	Ser	Ala	Phe	Leu	Ser	Tyr	435	440	445	
Ala	Asp	Trp	Phe	Lys	Ala	Ser	Phe	Gly	Ser	Thr	Arg	Gly	Tyr	His	Gly	450	455	460	
Cys	Ser	Lys	Lys	Ala	Leu	Val	Phe	Leu	Phe	Thr	Phe	Leu	Ser	Glu	Leu	465	470	475	480
Val	Pro	Phe	Glu	Ser	Pro	Arg	Tyr	Leu	Gln	Val	His	Ile	Leu	His	Pro	485	490	495	
Pro	Leu	Val	Pro	Gly	Lys	Tyr	Arg	Ser	Leu	Leu	Thr	Asp	Tyr	Ile	Ser	500	505	510	
Leu	Ala	Lys	Thr	Arg	Leu	Ala	Asp	Leu	Lys	Val	Ser	Ile	Glu	Asn	Met	515	520	525	
Gly	Leu	Tyr	Glu	Asp	Leu	Ser	Ser	Ala	Gly	Asp	Ile	Thr	Glu	Pro	His	530	535	540	
Ser	Gln	Ala	Leu	Gln	Asp	Val	Glu	Lys	Ala	Ile	Met	Val	Phe	Glu	His	545	550	555	560
Thr	Gly	Asn	Ile	Pro	Val	Thr	Val	Met	Glu	Ala	Ser	Ile	Phe	Arg	Arg	565	570	575	
Pro	Tyr	Tyr	Val	Ser	His	Phe	Leu	Pro	Ala	Leu	Leu	Thr	Pro	Arg	Val	580	585	590	
Leu	Pro	Lys	Val	Pro	Asp	Ser	Arg	Val	Ala	Phe	Ile	Glu	Ser	Leu	Lys	595	600	605	
Arg	Ala	Asp	Lys	Ile	Pro	Pro	Ser	Leu	Tyr	Ser	Thr	Tyr	Cys	Gln	Ala	610	615	620	
Cys	Ser	Ala	Ala	Glu	Glu	Lys	Pro	Glu	Asp	Ala	Ala	Leu	Gly	Val	Arg	625	630	635	640
Ala	Glu	Pro	Asn	Ser	Ala	Glu	Glu	Pro	Leu	Gly	Gln	Leu	Thr	Ala	Ala	645	650	655	
Leu	Gly	Glu	Leu	Arg	Ala	Ser	Met	Thr	Asp	Pro	Ser	Gln	Arg	Asp	Val	660	665	670	
Ile	Ser	Ala	Gln	Val	Ala	Val	Ile	Ser	Glu	Arg	Leu	Arg	Ala	Val	Leu	675	680	685	
Gly	His	Asn	Glu	Asp	Asp	Ser	Ser	Val	Glu	Ile	Ser	Lys	Ile	Gln	Leu	690	695	700	
Ser	Ile	Asn	Thr	Pro	Arg	Leu	Glu	Pro	Arg	Glu	His	Met	Ala	Val	Asp	705	710	715	720
Leu	Leu	Leu	Thr	Ser	Phe	Cys	Gln	Asn	Leu	Met	Ala	Ala	Ser	Ser	Val	725	730	735	

Ala	Pro	Pro	Glu	Arg	Gln	Gly	Pro	Trp	Ala	Ala	Leu	Phe	Val	Arg	Thr		
			740					745					750				
Met	Cys	Gly	Arg	Val	Leu	Pro	Ala	Val	Leu	Thr	Arg	Leu	Cys	Gln	Leu		
		755					760					765					
Leu	Arg	His	Gln	Gly	Pro	Ser	Leu	Ser	Ala	Pro	His	Val	Leu	Gly	Leu		
	770					775					780						
Ala	Ala	Leu	Ala	Val	His	Leu	Gly	Glu	Ser	Arg	Ser	Ala	Leu	Pro	Glu		
785					790					795					800		
Val	Asp	Val	Gly	Pro	Pro	Ala	Pro	Gly	Ala	Gly	Leu	Pro	Val	Pro	Ala		
			805					810						815			
Leu	Phe	Asp	Ser	Leu	Leu	Thr	Cys	Arg	Thr	Arg	Asp	Ser	Leu	Phe	Phe		
		820						825					830				
Cys	Leu	Lys	Phe	Cys	Thr	Ala	Ala	Ile	Ser	Tyr	Ser	Leu	Cys	Lys	Phe		
		835				840						845					
Ser	Ser	Gln	Ser	Arg	Asp	Thr	Leu	Cys	Ser	Cys	Leu	Ser	Pro	Gly	Leu		
	850				855						860						
Ile	Lys	Lys	Phe	Gln	Phe	Leu	Met	Phe	Arg	Leu	Phe	Ser	Glu	Ala	Arg		
865				870					875						880		
Gln	Pro	Leu	Ser	Glu	Glu	Asp	Val	Ala	Ser	Leu	Ser	Trp	Arg	Pro	Leu		
			885					890						895			
His	Leu	Pro	Ser	Ala	Asp	Trp	Gln	Arg	Ala	Ala	Leu	Ser	Leu	Trp	Thr		
		900					905						910				
His	Arg	Thr	Phe	Arg	Glu	Val	Leu	Lys	Glu	Glu	Asp	Val	His	Leu	Thr		
	915					920					925						
Tyr	Gln	Asp	Trp	Leu	His	Leu	Glu	Leu	Glu	Ile	Gln	Pro	Glu	Ala	Asp		
	930				935					940							
Ala	Leu	Ser	Asp	Thr	Glu	Arg	Gln	Asp	Phe	His	Gln	Trp	Ala	Ile	His		
945				950					955						960		
Glu	His	Phe	Leu	Pro	Glu	Ser	Ser	Ala	Ser	Gly	Gly	Cys	Asp	Gly	Asp		
			965					970						975			
Leu	Gln	Ala	Ala	Cys	Thr	Ile	Leu	Val	Asn	Ala	Leu	Met	Asp	Phe	His		
		980					985						990				
Gln	Ser	Ser	Arg	Ser	Tyr	Asp	His	Ser	Glu	Asn	Ser	Asp	Leu	Val	Phe		
	995					1000						1005					
Gly	Gly	Arg	Thr	Gly	Asn	Glu	Asp	Ile	Ile	Ser	Arg	Leu	Gln	Glu	Met		
	1010				1015						1020						
Val	Ala	Asp	Leu	Glu	Leu	Gln	Gln	Asp	Leu	Ile	Val	Pro	Leu	Gly	His		
1025				1030					1035						1040		
Thr	Pro	Ser	Gln	Glu	His	Phe	Leu	Phe	Glu	Ile	Phe	Arg	Arg	Arg	Leu		
			1045					1050						1055			
Gln	Ala	Leu	Thr	Ser	Gly	Trp	Ser	Val	Ala	Ala	Ser	Leu	Gln	Arg	Gln		
		1060					1065						1070				
Arg	Glu	Leu	Leu	Met	Tyr	Lys	Arg	Ile	Leu	Leu	Arg	Leu	Pro	Ser	Ser		
	1075					1080						1085					
Val	Leu	Cys	Gly	Ser	Ser	Phe	Gln	Ala	Glu	Gln	Pro	Ile	Thr	Ala	Arg		
	1090					1095					1100						
Cys	Glu	Gln	Phe	Phe	His	Leu	Val	Asn	Ser	Glu	Met	Arg	Asn	Phe	Cys		
1105				1110					1115						1120		
Ser	His	Gly	Gly	Ala	Leu	Thr	Gln	Asp	Ile	Thr	Ala	His	Phe	Phe	Arg		
			1125					1130						1135			
Gly	Leu	Leu	Asn	Ala	Cys	Leu	Arg	Ser	Arg	Asp	Pro	Ser	Leu	Met	Val		
		1140					1145						1150				
Asp	Phe	Ile	Leu	Ala	Lys	Cys	Gln	Thr	Lys	Cys	Pro	Leu	Ile	Leu	Thr		
	1155						1160					1165					

Ser Ala Leu Val Trp Trp Pro Ser Leu Glu Pro Val Leu Leu Cys Arg
 1170 1175 1180
 Trp Arg Arg His Cys Gln Ser Pro Leu Pro Arg Glu Leu Gln Lys Leu
 1185 1190 1195 1200
 Gln Glu Gly Arg Gln Phe Ala Ser Asp Phe Leu Ser Pro Glu Ala Ala
 1205 1210 1215
 Ser Pro Ala Pro Asn Pro Asp Trp Leu Ser Ala Ala Ala Leu His Phe
 1220 1225 1230
 Ala Ile Gln Gln Val Arg Glu Glu Asn Ile Arg Lys Gln Leu Lys Lys
 1235 1240 1245
 Leu Asp Cys Glu Arg Glu Glu Leu Leu Val Phe Leu Phe Phe Phe Ser
 1250 1255 1260
 Leu Met Gly Leu Leu Ser Ser His Leu Thr Ser Asn Ser Thr Thr Asp
 1265 1270 1275 1280
 Leu Pro Lys Ala Phe His Val Cys Ala Ala Ile Leu Glu Cys Leu Glu
 1285 1290 1295
 Lys Arg Lys Ile Ser Trp Leu Ala Leu Phe Gln Leu Thr Glu Ser Asp
 1300 1305 1310
 Leu Arg Leu Gly Arg Leu Leu Leu Arg Val Ala Pro Asp Gln His Thr
 1315 1320 1325
 Arg Leu Leu Pro Phe Ala Phe Tyr Ser Leu Leu Ser Tyr Phe His Glu
 1330 1335 1340
 Asp Ala Ala Ile Arg Glu Glu Ala Phe Leu His Val Ala Val Asp Met
 1345 1350 1355 1360
 Tyr Leu Lys Leu Val Gln Leu Phe Val Ala Gly Asp Thr Ser Thr Val
 1365 1370 1375
 Ser Pro Pro Ala Gly Arg Ser Leu Glu Leu Lys Gly Gln Gly Asn Pro
 1380 1385 1390
 Val Glu Leu Ile Thr Lys Ala Arg Leu Phe Leu Leu Gln Leu Ile Pro
 1395 1400 1405
 Arg Cys Pro Lys Lys Ser Phe Ser His Val Ala Glu Leu Leu Ala Asp
 1410 1415 1420
 Arg Gly Asp Cys Asp Pro Glu Val Ser Ala Ala Leu Gln Ser Arg Gln
 1425 1430 1435 1440
 Gln Ala Ala Pro Asp Ala Asp Leu Ser Gln Glu Pro His Leu Phe
 1445 1450 1455

<210> 146

<211> 859

<212> PRT

<213> Homo sapiens

<220>

<223> >FANCB|ENSG00000181544|ENST00000398334|2580

<400> 146

Met Thr Ser Lys Gln Ala Met Ser Ser Asn Glu Gln Glu Arg Leu Leu
 1 5 10 15
 Cys Tyr Asn Gly Glu Val Leu Val Phe Gln Leu Ser Lys Gly Asn Phe
 20 25 30
 Ala Asp Lys Glu Pro Thr Lys Thr Pro Ile Leu His Val Arg Arg Met
 35 40 45
 Val Phe Asp Arg Gly Thr Lys Val Phe Val Gln Lys Ser Thr Gly Phe

50		55		60											
Phe	Thr	Ile	Lys	Glu	Glu	Asn	Ser	His	Leu	Lys	Ile	Met	Cys	Cys	Asn
65				70						75					80
Cys	Val	Ser	Asp	Phe	Arg	Thr	Gly	Ile	Asn	Leu	Pro	Tyr	Ile	Val	Ile
			85						90					95	
Glu	Lys	Asn	Lys	Lys	Asn	Asn	Val	Phe	Glu	Tyr	Phe	Leu	Leu	Ile	Leu
			100					105					110		
His	Ser	Thr	Asn	Lys	Phe	Glu	Met	Arg	Leu	Ser	Phe	Lys	Leu	Gly	Tyr
		115				120						125			
Glu	Met	Lys	Asp	Gly	Leu	Arg	Val	Leu	Asn	Gly	Pro	Leu	Ile	Leu	Trp
130						135					140				
Arg	His	Val	Lys	Ala	Phe	Phe	Phe	Ile	Ser	Ser	Gln	Thr	Gly	Lys	Val
145					150					155					160
Val	Ser	Val	Ser	Gly	Asn	Phe	Ser	Ser	Ile	Gln	Trp	Ala	Gly	Glu	Ile
				165					170					175	
Glu	Asn	Leu	Gly	Met	Val	Leu	Leu	Gly	Leu	Lys	Glu	Cys	Cys	Leu	Ser
			180					185					190		
Glu	Glu	Glu	Cys	Thr	Gln	Glu	Pro	Ser	Lys	Ser	Asp	Tyr	Ala	Ile	Trp
		195					200					205			
Asn	Thr	Lys	Phe	Cys	Val	Tyr	Ser	Leu	Glu	Ser	Gln	Glu	Val	Leu	Ser
210						215					220				
Asp	Ile	Tyr	Ile	Ile	Pro	Pro	Ala	Tyr	Ser	Ser	Val	Val	Thr	Tyr	Val
225					230					235					240
His	Ile	Cys	Ala	Thr	Glu	Ile	Ile	Lys	Asn	Gln	Leu	Arg	Ile	Ser	Leu
			245						250					255	
Ile	Ala	Leu	Thr	Arg	Lys	Asn	Gln	Leu	Ile	Ser	Phe	Gln	Asn	Gly	Thr
			260					265					270		
Pro	Lys	Asn	Val	Cys	Gln	Leu	Pro	Phe	Gly	Asp	Pro	Cys	Ala	Val	Gln
		275					280					285			
Leu	Met	Asp	Ser	Gly	Gly	Gly	Asn	Leu	Phe	Phe	Val	Val	Ser	Phe	Ile
290						295					300				
Ser	Asn	Asn	Ala	Cys	Ala	Val	Trp	Lys	Glu	Ser	Phe	Gln	Val	Ala	Ala
305					310					315					320
Lys	Trp	Glu	Lys	Leu	Ser	Leu	Val	Leu	Ile	Asp	Asp	Phe	Ile	Gly	Ser
			325						330					335	
Gly	Thr	Glu	Gln	Val	Leu	Leu	Leu	Phe	Lys	Asp	Ser	Leu	Asn	Ser	Asp
		340						345					350		
Cys	Leu	Thr	Ser	Phe	Lys	Ile	Thr	Asp	Leu	Gly	Lys	Ile	Asn	Tyr	Ser
		355					360					365			
Ser	Glu	Pro	Ser	Asp	Cys	Asn	Glu	Asp	Asp	Leu	Phe	Glu	Asp	Lys	Gln
370						375					380				
Glu	Asn	Arg	Tyr	Leu	Val	Val	Pro	Pro	Leu	Glu	Thr	Gly	Leu	Lys	Val
385					390					395					400
Cys	Phe	Ser	Ser	Phe	Arg	Glu	Leu	Arg	Gln	His	Leu	Leu	Leu	Lys	Glu
				405					410					415	
Lys	Ile	Ile	Ser	Lys	Ser	Tyr	Lys	Ala	Leu	Ile	Asn	Leu	Val	Gln	Gly
			420					425					430		
Lys	Asp	Asp	Asn	Thr	Ser	Ser	Ala	Glu	Glu	Lys	Glu	Cys	Leu	Val	Pro
		435					440					445			
Leu	Cys	Gly	Glu	Glu	Glu	Asn	Ser	Val	His	Ile	Leu	Asp	Glu	Lys	Leu
450						455					460				
Ser	Asp	Asn	Phe	Gln	Asp	Ser	Glu	Gln	Leu	Val	Glu	Lys	Ile	Trp	Tyr
465					470					475					480
Arg	Val	Ile	Asp	Asp	Ser	Leu	Val	Val	Gly	Val	Lys	Thr	Thr	Ser	Ser

<220>

<223> >FANCC|ENSG00000158169|ENST00000289081|1677

<400> 147

Met	Ala	Gln	Asp	Ser	Val	Asp	Leu	Ser	Cys	Asp	Tyr	Gln	Phe	Trp	Met
1				5					10					15	
Gln	Lys	Leu	Ser	Val	Trp	Asp	Gln	Ala	Ser	Thr	Leu	Glu	Thr	Gln	Gln
			20					25					30		
Asp	Thr	Cys	Leu	His	Val	Ala	Gln	Phe	Gln	Glu	Phe	Leu	Arg	Lys	Met
		35					40					45			
Tyr	Glu	Ala	Leu	Lys	Glu	Met	Asp	Ser	Asn	Thr	Val	Ile	Glu	Arg	Phe
	50					55					60				
Pro	Thr	Ile	Gly	Gln	Leu	Leu	Ala	Lys	Ala	Cys	Trp	Asn	Pro	Phe	Ile
65					70					75					80
Leu	Ala	Tyr	Asp	Glu	Ser	Gln	Lys	Ile	Leu	Ile	Trp	Cys	Leu	Cys	Cys
				85					90				95		
Leu	Ile	Asn	Lys	Glu	Pro	Gln	Asn	Ser	Gly	Gln	Ser	Lys	Leu	Asn	Ser
			100					105					110		
Trp	Ile	Gln	Gly	Val	Leu	Ser	His	Ile	Leu	Ser	Ala	Leu	Arg	Phe	Asp
		115					120						125		
Lys	Glu	Val	Ala	Leu	Phe	Thr	Gln	Gly	Leu	Gly	Tyr	Ala	Pro	Ile	Asp
	130					135					140				
Tyr	Tyr	Pro	Gly	Leu	Leu	Lys	Asn	Met	Val	Leu	Ser	Leu	Ala	Ser	Glu
145					150					155					160
Leu	Arg	Glu	Asn	His	Leu	Asn	Gly	Phe	Asn	Thr	Gln	Arg	Arg	Met	Ala
				165					170					175	
Pro	Glu	Arg	Val	Ala	Ser	Leu	Ser	Arg	Val	Cys	Val	Pro	Leu	Ile	Thr
			180					185					190		
Leu	Thr	Asp	Val	Asp	Pro	Leu	Val	Glu	Ala	Leu	Leu	Ile	Cys	His	Gly
		195					200					205			
Arg	Glu	Pro	Gln	Glu	Ile	Leu	Gln	Pro	Glu	Phe	Phe	Glu	Ala	Val	Asn
	210					215					220				
Glu	Ala	Ile	Leu	Leu	Lys	Lys	Ile	Ser	Leu	Pro	Met	Ser	Ala	Val	Val
225					230					235					240
Cys	Leu	Trp	Leu	Arg	His	Leu	Pro	Ser	Leu	Glu	Lys	Ala	Met	Leu	His
				245					250					255	
Leu	Phe	Glu	Lys	Leu	Ile	Ser	Ser	Glu	Arg	Asn	Cys	Leu	Arg	Arg	Ile
			260					265					270		
Glu	Cys	Phe	Ile	Lys	Asp	Ser	Ser	Leu	Pro	Gln	Ala	Ala	Cys	His	Pro
		275					280					285			
Ala	Ile	Phe	Arg	Val	Val	Asp	Glu	Met	Phe	Arg	Cys	Ala	Leu	Leu	Glu
	290					295					300				
Thr	Asp	Gly	Ala	Leu	Glu	Ile	Ile	Ala	Thr	Ile	Gln	Val	Phe	Thr	Gln
305					310					315					320
Cys	Phe	Val	Glu	Ala	Leu	Glu	Lys	Ala	Ser	Lys	Gln	Leu	Arg	Phe	Ala
				325					330					335	
Leu	Lys	Thr	Tyr	Phe	Pro	Tyr	Thr	Ser	Pro	Ser	Leu	Ala	Met	Val	Leu
			340					345					350		
Leu	Gln	Asp	Pro	Gln	Asp	Ile	Pro	Arg	Gly	His	Trp	Leu	Gln	Thr	Leu
		355					360					365			
Lys	His	Ile	Ser	Glu	Leu	Leu	Arg	Glu	Ala	Val	Glu	Asp	Gln	Thr	His
	370					375					380				
Gly	Ser	Cys	Gly	Gly	Pro	Phe	Glu	Ser	Trp	Phe	Leu	Phe	Ile	His	Phe
385					390					395					400

Gly	Gly	Trp	Ala	Glu	Met	Val	Ala	Glu	Gln	Leu	Leu	Met	Ser	Ala	Ala	
				405					410					415		
Glu	Pro	Pro	Thr	Ala	Leu	Leu	Trp	Leu	Leu	Ala	Phe	Tyr	Tyr	Gly	Pro	
			420					425					430			
Arg	Asp	Gly	Arg	Gln	Gln	Arg	Ala	Gln	Thr	Met	Val	Gln	Val	Lys	Ala	
		435					440					445				
Val	Leu	Gly	His	Leu	Leu	Ala	Met	Ser	Arg	Ser	Ser	Ser	Leu	Ser	Ala	
	450					455					460					
Gln	Asp	Leu	Gln	Thr	Val	Ala	Gly	Gln	Gly	Thr	Asp	Thr	Asp	Leu	Arg	
465					470					475					480	
Ala	Pro	Ala	Gln	Gln	Leu	Ile	Arg	His	Leu	Leu	Leu	Asn	Phe	Leu	Leu	
			485						490					495		
Trp	Ala	Pro	Gly	Gly	His	Thr	Ile	Ala	Trp	Asp	Val	Ile	Thr	Leu	Met	
			500					505					510			
Ala	His	Thr	Ala	Glu	Ile	Thr	His	Glu	Ile	Ile	Gly	Phe	Leu	Asp	Gln	
		515					520				525					
Thr	Leu	Tyr	Arg	Trp	Asn	Arg	Leu	Gly	Ile	Glu	Ser	Pro	Arg	Ser	Glu	
	530					535					540					
Lys	Leu	Ala	Arg	Glu	Leu	Leu	Lys	Glu	Leu	Arg	Thr	Gln	Val			
545					550					555						

<210> 148

<211> 1471

<212> PRT

<213> Homo sapiens

<220>

<223> >FANCD2|ENSG00000144554|ENST00000287647|4416

<400> 148

Met	Val	Ser	Lys	Arg	Arg	Leu	Ser	Lys	Ser	Glu	Asp	Lys	Glu	Ser	Leu	
1				5					10					15		
Thr	Glu	Asp	Ala	Ser	Lys	Thr	Arg	Lys	Gln	Pro	Leu	Ser	Lys	Lys	Thr	
			20					25					30			
Lys	Lys	Ser	His	Ile	Ala	Asn	Glu	Val	Glu	Glu	Asn	Asp	Ser	Ile	Phe	
		35				40					45					
Val	Lys	Leu	Leu	Lys	Ile	Ser	Gly	Ile	Ile	Leu	Lys	Thr	Gly	Glu	Ser	
	50					55					60					
Gln	Asn	Gln	Leu	Ala	Val	Asp	Gln	Ile	Ala	Phe	Gln	Lys	Lys	Leu	Phe	
65					70				75						80	
Gln	Thr	Leu	Arg	Arg	His	Pro	Ser	Tyr	Pro	Lys	Ile	Ile	Glu	Glu	Phe	
			85						90					95		
Val	Ser	Gly	Leu	Glu	Ser	Tyr	Ile	Glu	Asp	Glu	Asp	Ser	Phe	Arg	Asn	
		100						105					110			
Cys	Leu	Leu	Ser	Cys	Glu	Arg	Leu	Gln	Asp	Glu	Glu	Ala	Ser	Met	Gly	
		115					120					125				
Ala	Ser	Tyr	Ser	Lys	Ser	Leu	Ile	Lys	Leu	Leu	Leu	Gly	Ile	Asp	Ile	
		130				135					140					
Leu	Gln	Pro	Ala	Ile	Ile	Lys	Thr	Leu	Phe	Glu	Lys	Leu	Pro	Glu	Tyr	
145					150					155					160	
Phe	Phe	Glu	Asn	Lys	Asn	Ser	Asp	Glu	Ile	Asn	Ile	Pro	Arg	Leu	Ile	
			165						170					175		
Val	Ser	Gln	Leu	Lys	Trp	Leu	Asp	Arg	Val	Val	Asp	Gly	Lys	Asp	Leu	

Page 455

<210> 149
 <211> 536
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >FANCE|ENSG00000112039|ENST00000229769|1611

<400> 149
 Met Ala Thr Pro Asp Ala Gly Leu Pro Gly Ala Glu Gly Val Glu Pro
 1 5 10 15
 Ala Pro Trp Ala Gln Leu Glu Ala Pro Ala Arg Leu Leu Leu Gln Ala
 20 25 30
 Leu Gln Ala Gly Pro Glu Gly Ala Arg Arg Gly Leu Gly Val Leu Arg
 35 40 45
 Ala Leu Gly Ser Arg Gly Trp Glu Pro Phe Asp Trp Gly Arg Leu Leu
 50 55 60
 Glu Ala Leu Cys Arg Glu Glu Pro Val Val Gln Gly Pro Asp Gly Arg
 65 70 75 80
 Leu Glu Leu Lys Pro Leu Leu Leu Arg Leu Pro Arg Ile Cys Gln Arg
 85 90 95
 Asn Leu Met Ser Leu Leu Met Ala Val Arg Pro Ser Leu Pro Glu Ser
 100 105 110
 Gly Leu Leu Ser Val Leu Gln Ile Ala Gln Gln Asp Leu Ala Pro Asp
 115 120 125
 Pro Asp Ala Trp Leu Arg Ala Leu Gly Glu Leu Leu Arg Arg Asp Leu
 130 135 140
 Gly Val Gly Thr Ser Met Glu Gly Ala Ser Pro Leu Ser Glu Arg Cys
 145 150 155 160
 Gln Arg Gln Leu Gln Ser Leu Cys Arg Gly Leu Gly Leu Gly Gly Arg
 165 170 175
 Arg Leu Lys Ser Pro Gln Ala Pro Asp Pro Glu Glu Glu Glu Asn Arg
 180 185 190
 Asp Ser Gln Gln Pro Gly Lys Arg Arg Lys Asp Ser Glu Glu Glu Ala
 195 200 205
 Ala Ser Pro Glu Gly Lys Arg Val Pro Lys Arg Leu Arg Cys Trp Glu
 210 215 220
 Glu Glu Glu Asp His Glu Lys Glu Arg Pro Glu His Lys Ser Leu Glu
 225 230 235 240
 Ser Leu Ala Asp Gly Gly Ser Ala Ser Pro Ile Lys Asp Gln Pro Val
 245 250 255
 Met Ala Val Lys Thr Gly Glu Asp Gly Ser Asn Leu Asp Asp Ala Lys
 260 265 270
 Gly Leu Ala Glu Ser Leu Glu Leu Pro Lys Ala Ile Gln Asp Gln Leu
 275 280 285
 Pro Arg Leu Gln Gln Leu Leu Lys Thr Leu Glu Glu Gly Leu Glu Gly
 290 295 300
 Leu Glu Asp Ala Pro Pro Val Glu Leu Gln Leu Leu His Glu Cys Ser
 305 310 315 320
 Pro Ser Gln Met Asp Leu Leu Cys Ala Gln Leu Gln Leu Pro Gln Leu
 325 330 335
 Ser Asp Leu Gly Leu Leu Arg Leu Cys Thr Trp Leu Leu Ala Leu Ser
 340 345 350

Pro Asp Leu Ser Leu Ser Asn Ala Thr Val Leu Thr Arg Ser Leu Phe
 355 360 365
 Leu Gly Arg Ile Leu Ser Leu Thr Ser Ser Ala Ser Arg Leu Leu Thr
 370 375 380
 Thr Ala Leu Thr Ser Phe Cys Ala Lys Tyr Thr Tyr Pro Val Cys Ser
 385 390 395 400
 Ala Leu Leu Asp Pro Val Leu Gln Ala Pro Gly Thr Gly Pro Ala Gln
 405 410 415
 Thr Glu Leu Leu Cys Cys Leu Val Lys Met Glu Ser Leu Glu Pro Asp
 420 425 430
 Ala Gln Val Leu Met Leu Gly Gln Ile Leu Glu Leu Pro Trp Lys Glu
 435 440 445
 Glu Thr Phe Leu Val Leu Gln Ser Leu Leu Glu Arg Gln Val Glu Met
 450 455 460
 Thr Pro Glu Lys Phe Ser Val Leu Met Glu Lys Leu Cys Lys Lys Gly
 465 470 475 480
 Leu Ala Ala Thr Thr Ser Met Ala Tyr Ala Lys Leu Met Leu Thr Val
 485 490 495
 Met Thr Lys Tyr Gln Ala Asn Ile Thr Glu Thr Gln Arg Leu Gly Leu
 500 505 510
 Ala Met Ala Leu Glu Pro Asn Thr Thr Phe Leu Arg Lys Ser Leu Lys
 515 520 525
 Ala Ala Leu Lys His Leu Gly Pro
 530 535

<210> 150

<211> 374

<212> PRT

<213> Homo sapiens

<220>

<223> >FANCF|ENSG00000183161|ENST00000327470|1125

<400> 150

Met Glu Ser Leu Leu Gln His Leu Asp Arg Phe Ser Glu Leu Leu Ala
 1 5 10 15
 Val Ser Ser Thr Thr Tyr Val Ser Thr Trp Asp Pro Ala Thr Val Arg
 20 25 30
 Arg Ala Leu Gln Trp Ala Arg Tyr Leu Arg His Ile His Arg Arg Phe
 35 40 45
 Gly Arg His Gly Pro Ile Arg Thr Ala Leu Glu Arg Arg Leu His Asn
 50 55 60
 Gln Trp Arg Gln Glu Gly Gly Phe Gly Arg Gly Pro Val Pro Gly Leu
 65 70 75 80
 Ala Asn Phe Gln Ala Leu Gly His Cys Asp Val Leu Leu Ser Leu Arg
 85 90 95
 Leu Leu Glu Asn Arg Ala Leu Gly Asp Ala Ala Arg Tyr His Leu Val
 100 105 110
 Gln Gln Leu Phe Pro Gly Pro Gly Val Arg Asp Ala Asp Glu Glu Thr
 115 120 125
 Leu Gln Glu Ser Leu Ala Arg Leu Ala Arg Arg Arg Ser Ala Val His
 130 135 140
 Met Leu Arg Phe Asn Gly Tyr Arg Glu Asn Pro Asn Leu Gln Glu Asp

145					150				155				160	
Ser	Leu	Met	Lys	Thr	Gln	Ala	Glu	Leu	Leu	Glu	Arg	Leu	Gln	Glu
				165					170				175	
Val	Gly	Lys	Ala	Glu	Ala	Glu	Arg	Pro	Ala	Arg	Phe	Leu	Ser	Ser
			180					185					190	
Trp	Glu	Arg	Leu	Pro	Gln	Asn	Asn	Phe	Leu	Lys	Val	Ile	Ala	Val
		195					200					205		
Leu	Leu	Gln	Pro	Pro	Leu	Ser	Arg	Arg	Pro	Gln	Glu	Glu	Leu	Glu
	210					215					220			
Gly	Ile	His	Lys	Ser	Pro	Gly	Glu	Gly	Ser	Gln	Val	Leu	Val	His
225					230					235				240
Leu	Leu	Gly	Asn	Ser	Glu	Val	Phe	Ala	Ala	Phe	Cys	Arg	Ala	Leu
			245					250						255
Ala	Gly	Leu	Leu	Thr	Leu	Val	Thr	Ser	Arg	His	Pro	Ala	Leu	Ser
			260					265					270	
Val	Tyr	Leu	Gly	Leu	Leu	Thr	Asp	Trp	Gly	Gln	Arg	Leu	His	Tyr
	275						280					285		
Leu	Gln	Lys	Gly	Ile	Trp	Val	Gly	Thr	Glu	Ser	Gln	Asp	Val	Pro
	290					295					300			
Glu	Glu	Leu	His	Asn	Arg	Phe	Gln	Ser	Leu	Cys	Gln	Ala	Pro	Pro
305					310					315				320
Leu	Lys	Asp	Lys	Val	Leu	Thr	Ala	Leu	Glu	Thr	Cys	Lys	Ala	Gln
			325						330					335
Gly	Asp	Phe	Glu	Val	Pro	Gly	Leu	Ser	Ile	Trp	Thr	Asp	Leu	Leu
		340					345					350		
Ala	Leu	Arg	Ser	Gly	Ala	Phe	Arg	Lys	Arg	Gln	Val	Leu	Gly	Leu
		355					360					365		
Ala	Gly	Leu	Ser	Ser	Val									
	370													

<210> 151
 <211> 622
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >FANCG|ENSG00000221829|ENST00000378643|1869

<400> 151
Met Ser Arg Gln Thr Thr Ser Val Gly Ser Ser Cys Leu Asp Leu Trp
1 5 10 15
Arg Glu Lys Asn Asp Arg Leu Val Arg Gln Ala Lys Val Ala Gln Asn
20 25 30
Ser Gly Leu Thr Leu Arg Arg Gln Gln Leu Ala Gln Asp Ala Leu Glu
35 40 45
Gly Leu Arg Gly Leu Leu His Ser Leu Gln Gly Leu Pro Ala Ala Val
50 55 60
Pro Val Leu Pro Leu Glu Leu Thr Val Thr Cys Asn Phe Ile Ile Leu
65 70 75 80
Arg Ala Ser Leu Ala Gln Gly Phe Thr Glu Asp Gln Ala Gln Asp Ile
85 90 95
Gln Arg Ser Leu Glu Arg Val Leu Glu Thr Gln Glu Gln Gln Gly Pro
100 105 110

Arg	Leu	Glu	Gln	Gly	Leu	Arg	Glu	Leu	Trp	Asp	Ser	Val	Leu	Arg	Ala	115	120	125
Ser	Cys	Leu	Leu	Pro	Glu	Leu	Leu	Ser	Ala	Leu	His	Arg	Leu	Val	Gly	130	135	140
Leu	Gln	Ala	Ala	Leu	Trp	Leu	Ser	Ala	Asp	Arg	Leu	Gly	Asp	Leu	Ala	145	150	155
Leu	Leu	Leu	Glu	Thr	Leu	Asn	Gly	Ser	Gln	Ser	Gly	Ala	Ser	Lys	Asp	165	170	175
Leu	Leu	Leu	Leu	Lys	Thr	Trp	Ser	Pro	Pro	Ala	Glu	Glu	Leu	Asp		180	185	190
Ala	Pro	Leu	Thr	Leu	Gln	Asp	Ala	Gln	Gly	Leu	Lys	Asp	Val	Leu	Leu	195	200	205
Thr	Ala	Phe	Ala	Tyr	Arg	Gln	Gly	Leu	Gln	Glu	Leu	Ile	Thr	Gly	Asn	210	215	220
Pro	Asp	Lys	Ala	Leu	Ser	Ser	Leu	His	Glu	Ala	Ala	Ser	Gly	Leu	Cys	225	230	235
Pro	Arg	Pro	Val	Leu	Val	Gln	Val	Tyr	Thr	Ala	Leu	Gly	Ser	Cys	His	245	250	255
Arg	Lys	Met	Gly	Asn	Pro	Gln	Arg	Ala	Leu	Leu	Tyr	Leu	Val	Ala	Ala	260	265	270
Leu	Lys	Glu	Gly	Ser	Ala	Trp	Gly	Pro	Pro	Leu	Leu	Glu	Ala	Ser	Arg	275	280	285
Leu	Tyr	Gln	Gln	Leu	Gly	Asp	Thr	Thr	Ala	Glu	Leu	Glu	Ser	Leu	Glu	290	295	300
Leu	Leu	Val	Glu	Ala	Leu	Asn	Val	Pro	Cys	Ser	Ser	Lys	Ala	Pro	Gln	305	310	315
Phe	Leu	Ile	Glu	Val	Glu	Leu	Leu	Leu	Pro	Pro	Pro	Asp	Leu	Ala	Ser	325	330	335
Pro	Leu	His	Cys	Gly	Thr	Gln	Ser	Gln	Thr	Lys	His	Ile	Leu	Ala	Ser	340	345	350
Arg	Cys	Leu	Gln	Thr	Gly	Arg	Ala	Gly	Asp	Ala	Ala	Glu	His	Tyr	Leu	355	360	365
Asp	Leu	Leu	Ala	Leu	Leu	Leu	Asp	Ser	Ser	Glu	Pro	Arg	Phe	Ser	Pro	370	375	380
Pro	Pro	Ser	Pro	Pro	Gly	Pro	Cys	Met	Pro	Glu	Val	Phe	Leu	Glu	Ala	385	390	395
Ala	Val	Ala	Leu	Ile	Gln	Ala	Gly	Arg	Ala	Gln	Asp	Ala	Leu	Thr	Leu	405	410	415
Cys	Glu	Glu	Leu	Leu	Ser	Arg	Thr	Ser	Ser	Leu	Leu	Pro	Lys	Met	Ser	420	425	430
Arg	Leu	Trp	Glu	Asp	Ala	Arg	Lys	Gly	Thr	Lys	Glu	Leu	Pro	Tyr	Cys	435	440	445
Pro	Leu	Trp	Val	Ser	Ala	Thr	His	Leu	Leu	Gln	Gly	Gln	Ala	Trp	Val	450	455	460
Gln	Leu	Gly	Ala	Gln	Lys	Val	Ala	Ile	Ser	Glu	Phe	Ser	Arg	Cys	Leu	465	470	475
Glu	Leu	Leu	Phe	Arg	Ala	Thr	Pro	Glu	Glu	Lys	Glu	Gln	Gly	Ala	Ala	485	490	495
Phe	Asn	Cys	Glu	Gln	Gly	Cys	Lys	Ser	Asp	Ala	Ala	Leu	Gln	Gln	Leu	500	505	510
Arg	Ala	Ala	Ala	Leu	Ile	Ser	Arg	Gly	Leu	Glu	Trp	Val	Ala	Ser	Gly	515	520	525
Gln	Asp	Thr	Lys	Ala	Leu	Gln	Asp	Phe	Leu	Leu	Ser	Val	Gln	Met	Cys	530	535	540

Pro	Gly	Asn	Arg	Asp	Thr	Tyr	Phe	His	Leu	Leu	Gln	Thr	Leu	Lys	Arg
545					550					555					560
Leu	Asp	Arg	Arg	Asp	Glu	Ala	Thr	Ala	Leu	Trp	Trp	Arg	Leu	Glu	Ala
				565					570					575	
Gln	Thr	Lys	Gly	Ser	His	Glu	Asp	Ala	Leu	Trp	Ser	Leu	Pro	Leu	Tyr
			580					585					590		
Leu	Glu	Ser	Tyr	Leu	Ser	Trp	Ile	Arg	Pro	Ser	Asp	Arg	Asp	Ala	Phe
		595					600					605			
Leu	Glu	Glu	Phe	Arg	Thr	Ser	Leu	Pro	Lys	Ser	Cys	Asp	Leu		
	610					615					620				

<210> 152
 <211> 1328
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >FANCI|ENSG00000140525|ENST00000310775|3987

<400> 152

Met	Asp	Gln	Lys	Ile	Leu	Ser	Leu	Ala	Ala	Glu	Lys	Thr	Ala	Asp	Lys
1				5					10					15	
Leu	Gln	Glu	Phe	Leu	Gln	Thr	Leu	Arg	Glu	Gly	Asp	Leu	Thr	Asn	Leu
			20					25					30		
Leu	Gln	Asn	Gln	Ala	Val	Lys	Gly	Lys	Val	Ala	Gly	Ala	Leu	Leu	Arg
		35					40					45			
Ala	Ile	Phe	Lys	Gly	Ser	Pro	Cys	Ser	Glu	Glu	Ala	Gly	Thr	Leu	Arg
	50					55					60				
Arg	Arg	Lys	Ile	Tyr	Thr	Cys	Cys	Ile	Gln	Leu	Val	Glu	Ser	Gly	Asp
65				70					75					80	
Leu	Gln	Lys	Glu	Ile	Ala	Ser	Glu	Ile	Ile	Gly	Leu	Leu	Met	Leu	Glu
			85					90					95		
Ala	His	His	Phe	Pro	Gly	Pro	Leu	Leu	Val	Glu	Leu	Ala	Asn	Glu	Phe
			100					105					110		
Ile	Ser	Ala	Val	Arg	Glu	Gly	Ser	Leu	Val	Asn	Gly	Lys	Ser	Leu	Glu
	115					120						125			
Leu	Leu	Pro	Ile	Ile	Leu	Thr	Ala	Leu	Ala	Thr	Lys	Lys	Glu	Asn	Leu
	130					135					140				
Ala	Tyr	Gly	Lys	Gly	Val	Leu	Ser	Gly	Glu	Glu	Cys	Lys	Lys	Gln	Leu
145					150				155					160	
Ile	Asn	Thr	Leu	Cys	Ser	Gly	Arg	Trp	Asp	Gln	Gln	Tyr	Val	Ile	Gln
			165						170					175	
Leu	Thr	Ser	Met	Phe	Lys	Asp	Val	Pro	Leu	Thr	Ala	Glu	Glu	Val	Glu
			180					185					190		
Phe	Val	Val	Glu	Lys	Ala	Leu	Ser	Met	Phe	Ser	Lys	Met	Asn	Leu	Gln
	195						200					205			
Glu	Ile	Pro	Pro	Leu	Val	Tyr	Gln	Leu	Leu	Val	Leu	Ser	Ser	Lys	Gly
	210					215					220				
Ser	Arg	Lys	Ser	Val	Leu	Glu	Gly	Ile	Ile	Ala	Phe	Phe	Ser	Ala	Leu
225					230					235				240	
Asp	Lys	Gln	His	Asn	Glu	Glu	Gln	Ser	Gly	Asp	Glu	Leu	Leu	Asp	Val
				245					250					255	
Val	Thr	Val	Pro	Ser	Gly	Glu	Leu	Arg	His	Val	Glu	Gly	Thr	Ile	Ile

Page 461

690	Ile	Leu	Glu	Ser	Ile	Thr	Asn	Arg	Met	Ile	Lys	Ser	Glu	Leu	Glu	Asp
705	Phe	Glu	Leu	Asp	Lys	Ser	Ala	Asp	Phe	Ser	Gln	Ser	Thr	Ser	Ile	Gly
					725					730					735	
	Ile	Lys	Asn	Asn	Ile	Cys	Ala	Phe	Leu	Val	Met	Gly	Val	Cys	Glu	Val
					740					745					750	
	Leu	Ile	Glu	Tyr	Asn	Phe	Ser	Ile	Ser	Ser	Phe	Ser	Lys	Asn	Arg	Phe
					755					760					765	
	Glu	Asp	Ile	Leu	Ser	Leu	Phe	Met	Cys	Tyr	Lys	Lys	Leu	Ser	Asp	Ile
					770										780	
	Leu	Asn	Glu	Lys	Ala	Gly	Lys	Ala	Lys	Thr	Lys	Met	Ala	Asn	Lys	Thr
					785										790	
	Ser	Asp	Ser	Leu	Leu	Ser	Met	Lys	Phe	Val	Ser	Ser	Leu	Leu	Thr	Ala
					805										810	
	Leu	Phe	Arg	Asp	Ser	Ile	Gln	Ser	His	Gln	Glu	Ser	Leu	Ser	Val	Leu
					820										830	
	Arg	Ser	Ser	Asn	Glu	Phe	Met	Arg	Tyr	Ala	Val	Asn	Val	Ala	Leu	Gln
					835										840	
	Lys	Val	Gln	Gln	Leu	Lys	Glu	Thr	Gly	His	Val	Ser	Gly	Pro	Asp	Gly
					850										855	
	Gln	Asn	Pro	Glu	Lys	Ile	Phe	Gln	Asn	Leu	Cys	Asp	Ile	Thr	Arg	Val
					865										870	
	Leu	Leu	Trp	Arg	Tyr	Thr	Ser	Ile	Pro	Thr	Ser	Val	Glu	Glu	Ser	Gly
					885										890	
	Lys	Lys	Glu	Lys	Gly	Lys	Ser	Ile	Ser	Leu	Leu	Cys	Leu	Glu	Gly	Leu
					900										905	
	Gln	Lys	Ile	Phe	Ser	Ala	Val	Gln	Gln	Phe	Tyr	Gln	Pro	Lys	Ile	Gln
					915										920	
	Gln	Phe	Leu	Arg	Ala	Leu	Asp	Val	Thr	Asp	Lys	Glu	Gly	Glu	Glu	Arg
					930										935	
	Glu	Asp	Ala	Asp	Val	Ser	Val	Thr	Gln	Arg	Thr	Ala	Phe	Gln	Ile	Arg
					945										950	
	Gln	Phe	Gln	Arg	Ser	Leu	Leu	Asn	Leu	Leu	Ser	Ser	Gln	Glu	Glu	Asp
					965										970	
	Phe	Asn	Ser	Lys	Glu	Ala	Leu	Leu	Leu	Val	Thr	Val	Leu	Thr	Ser	Leu
					980										985	
	Ser	Lys	Leu	Glu	Pro	Ser	Ser	Pro	Gln	Phe	Val	Gln	Met	Leu	Ser	
					995										1000	
	Trp	Thr	Ser	Lys	Ile	Cys	Lys	Glu	Asn	Ser	Arg	Glu	Asp	Ala	Leu	Phe
					1010										1015	
	Cys	Lys	Ser	Leu	Met	Asn	Leu	Leu	Phe	Ser	Leu	His	Val	Ser	Tyr	Lys
					1025										1030	
	Ser	Pro	Val	Ile	Leu	Leu	Arg	Asp	Leu	Ser	Gln	Asp	Ile	His	Gly	His
					1045										1050	
	Leu	Gly	Asp	Ile	Asp	Gln	Asp	Val	Glu	Val	Glu	Lys	Thr	Asn	His	Phe
					1060										1065	
	Ala	Ile	Val	Asn	Leu	Arg	Thr	Ala	Ala	Pro	Thr	Val	Cys	Leu	Leu	Val
					1075										1080	
	Leu	Ser	Gln	Ala	Glu	Lys	Val	Leu	Glu	Glu	Val	Asp	Trp	Leu	Ile	Thr
					1090										1095	
	Lys	Leu	Lys	Gly	Gln	Val	Ser	Gln	Glu	Thr	Leu	Ser	Glu	Glu	Ala	Ser
					1105										1110	
	Ser	Gln	Ala	Thr	Leu	Pro	Asn	Gln	Pro	Val	Glu	Lys	Ala	Ile	Ile	Met
															1115	
															1120	

Ala	Asp	Thr	Cys	Phe	Ser	Thr	Ile	Lys	Leu	Lys	Ala	Glu	Asp	Ala	Ser
130						135					140				
Gly	Arg	Glu	His	Leu	Ile	Thr	Leu	Lys	Leu	Lys	Ala	Lys	Tyr	Pro	Ala
145					150					155					160
Glu	Ser	Pro	Asp	Tyr	Phe	Val	Asp	Phe	Pro	Val	Pro	Phe	Cys	Ala	Ser
				165					170					175	
Trp	Thr	Pro	Gln	Val	Asn	Ser	Pro	Gln	Ser	Ser	Leu	Ile	Ser	Ile	Tyr
			180					185					190		
Ser	Gln	Phe	Leu	Ala	Ala	Ile	Glu	Ser	Leu	Lys	Ala	Phe	Trp	Asp	Val
		195					200				205				
Met	Asp	Glu	Ile	Asp	Glu	Lys	Thr	Trp	Val	Leu	Glu	Pro	Glu	Lys	Pro
		210				215					220				
Pro	Arg	Ser	Ala	Thr	Ala	Arg	Arg	Ile	Ala	Leu	Gly	Asn	Asn	Val	Ser
225					230					235					240
Ile	Asn	Ile	Glu	Val	Asp	Pro	Arg	His	Pro	Thr	Met	Leu	Pro	Glu	Cys
				245					250					255	
Phe	Phe	Leu	Gly	Ala	Asp	His	Val	Val	Lys	Pro	Leu	Gly	Ile	Lys	Leu
			260					265					270		
Ser	Arg	Asn	Ile	His	Leu	Trp	Asp	Pro	Glu	Asn	Ser	Val	Leu	Gln	Asn
		275					280					285			
Leu	Lys	Asp	Val	Leu	Glu	Ile	Asp	Phe	Pro	Ala	Arg	Ala	Ile	Leu	Glu
		290				295				300					
Lys	Ser	Asp	Phe	Thr	Met	Asp	Cys	Gly	Ile	Cys	Tyr	Ala	Tyr	Gln	Leu
305					310					315					320
Asp	Gly	Thr	Ile	Pro	Asp	Gln	Val	Cys	Asp	Asn	Ser	Gln	Cys	Gly	Gln
				325					330					335	
Pro	Phe	His	Gln	Ile	Cys	Leu	Tyr	Glu	Trp	Leu	Arg	Gly	Leu	Leu	Thr
			340					345					350		
Ser	Arg	Gln	Ser	Phe	Asn	Ile	Ile	Phe	Gly	Glu	Cys	Pro	Tyr	Cys	Ser
		355				360						365			
Lys	Pro	Ile	Thr	Leu	Lys	Met	Ser	Gly	Arg	Lys	His				
		370				375					380				

<210> 154

<211> 2048

<212> PRT

<213> Homo sapiens

<220>

<223> >FANCM|ENSG00000187790|ENST00000267430|6147

<400> 154

Met	Ser	Gly	Arg	Gln	Arg	Thr	Leu	Phe	Gln	Thr	Trp	Gly	Ser	Ser	Ile
1				5					10					15	
Ser	Arg	Ser	Ser	Gly	Thr	Pro	Gly	Cys	Ser	Ser	Gly	Thr	Glu	Arg	Pro
			20					25					30		
Gln	Ser	Pro	Gly	Ser	Ser	Lys	Ala	Pro	Leu	Pro	Ala	Ala	Ala	Glu	Ala
			35				40					45			
Gln	Leu	Glu	Ser	Asp	Asp	Asp	Val	Leu	Leu	Val	Ala	Ala	Tyr	Glu	Ala
		50				55					60				
Glu	Arg	Gln	Leu	Cys	Leu	Glu	Asn	Gly	Gly	Phe	Cys	Thr	Ser	Ala	Gly
65					70					75					80
Ala	Leu	Trp	Ile	Tyr	Pro	Thr	Asn	Cys	Pro	Val	Arg	Asp	Tyr	Gln	Leu

945					950					955				960	
Leu	Pro	Phe	Glu	Glu	Glu	Leu	Tyr	Ile	Val	Arg	Thr	Asp	Asp	Gln	Phe
				965					970					975	
Tyr	Asn	Cys	His	Ser	Leu	Thr	Lys	Glu	Val	Leu	Ala	Asn	Val	Glu	Arg
			980					985					990		
Phe	Leu	Ser	Tyr	Ser	Pro	Pro	Pro	Leu	Ser	Gly	Leu	Ser	Asp	Leu	Glu
		995					1000					1005			
Tyr	Glu	Ile	Ala	Lys	Gly	Thr	Ala	Leu	Glu	Asn	Leu	Leu	Phe	Leu	Pro
	1010					1015					1020				
Cys	Ala	Glu	His	Leu	Arg	Ser	Asp	Lys	Cys	Thr	Cys	Leu	Leu	Ser	His
1025					1030					1035					1040
Ser	Ala	Val	Asn	Ser	Gln	Gln	Asn	Leu	Glu	Leu	Asn	Ser	Leu	Lys	Cys
				1045					1050					1055	
Ile	Asn	Tyr	Pro	Ser	Glu	Lys	Ser	Cys	Leu	Tyr	Asp	Ile	Pro	Asn	Asp
			1060					1065					1070		
Asn	Ile	Ser	Asp	Glu	Pro	Ser	Leu	Cys	Asp	Cys	Asp	Val	His	Lys	His
	1075						1080					1085			
Asn	Gln	Asn	Glu	Asn	Leu	Val	Pro	Asn	Asn	Arg	Val	Gln	Ile	His	Arg
	1090					1095					1100				
Ser	Pro	Ala	Gln	Asn	Leu	Val	Gly	Glu	Asn	Asn	His	Asp	Val	Asp	Asn
1105					1110					1115					1120
Ser	Asp	Leu	Pro	Val	Leu	Ser	Thr	Asp	Gln	Asp	Glu	Ser	Leu	Leu	Leu
				1125					1130					1135	
Phe	Glu	Asp	Val	Asn	Thr	Glu	Phe	Asp	Asp	Val	Ser	Leu	Ser	Pro	Leu
		1140					1145					1150			
Asn	Ser	Lys	Ser	Glu	Ser	Leu	Pro	Val	Ser	Asp	Lys	Thr	Ala	Ile	Ser
	1155						1160					1165			
Glu	Thr	Pro	Leu	Val	Ser	Gln	Phe	Leu	Ile	Ser	Asp	Glu	Leu	Leu	Leu
	1170					1175					1180				
Asp	Asn	Asn	Ser	Glu	Leu	Gln	Asp	Gln	Ile	Thr	Arg	Asp	Ala	Asn	Ser
1185					1190					1195					1200
Phe	Lys	Ser	Arg	Asp	Gln	Arg	Gly	Val	Gln	Glu	Glu	Lys	Val	Lys	Asn
			1205						1210					1215	
His	Glu	Asp	Ile	Phe	Asp	Cys	Ser	Arg	Asp	Leu	Phe	Ser	Val	Thr	Phe
		1220						1225					1230		
Asp	Leu	Gly	Phe	Cys	Ser	Pro	Asp	Ser	Asp	Asp	Glu	Ile	Leu	Glu	His
	1235						1240					1245			
Thr	Ser	Asp	Ser	Asn	Arg	Pro	Leu	Asp	Asp	Leu	Tyr	Gly	Arg	Tyr	Leu
	1250					1255					1260				
Glu	Ile	Lys	Glu	Ile	Ser	Asp	Ala	Asn	Tyr	Val	Ser	Asn	Gln	Ala	Leu
1265					1270					1275					1280
Ile	Pro	Arg	Asp	His	Ser	Lys	Asn	Phe	Thr	Ser	Gly	Thr	Val	Ile	Ile
				1285					1290					1295	
Pro	Ser	Asn	Glu	Asp	Met	Gln	Asn	Pro	Asn	Tyr	Val	His	Leu	Pro	Leu
		1300					1305						1310		
Ser	Ala	Ala	Lys	Asn	Glu	Glu	Leu	Leu	Ser	Pro	Gly	Tyr	Ser	Gln	Phe
		1315					1320					1325			
Ser	Leu	Pro	Val	Gln	Lys	Lys	Val	Met	Ser	Thr	Pro	Leu	Ser	Lys	Ser
	1330					1335					1340				
Asn	Thr	Leu	Asn	Ser	Phe	Ser	Lys	Ile	Arg	Lys	Glu	Ile	Leu	Lys	Thr
1345					1350					1355					1360
Pro	Asp	Ser	Ser	Lys	Glu	Lys	Val	Asn	Leu	Gln	Arg	Phe	Lys	Glu	Ala
				1365					1370					1375	
Leu	Asn	Ser	Thr	Phe	Asp	Tyr	Ser	Glu	Phe	Ser	Leu	Glu	Lys	Ser	Lys

Page 468

Arg	Thr	Lys	Arg	Gly	Ser	Arg	Gly	Gln	Gly	Ser	Gln	Arg	Cys	Ile	Pro
			100					105					110		
Glu	Phe	Phe	Leu	Ala	Gly	Lys	Gln	Pro	Cys	Thr	Asn	Asp	Met	Ala	Lys
		115					120					125			
Ser	Asn	Ser	Val	Gly	Gln	Asp	Ser	Cys	Gln	Asp	Ser	Glu	Gly	Asp	Met
		130				135					140				
Ile	Phe	Pro	Ala	Glu	Ser	Ser	Cys	Ala	Leu	Pro	Gln	Glu	Gly	Ser	Ala
145					150					155					160
Gly	Pro	Gly	Ser	Pro	Gly	Ser	Ala	Pro	Pro	Ser	Arg	Lys	Arg	Ser	Trp
				165					170					175	
Ser	Ser	Glu	Glu	Glu	Ser	Asn	Gln	Ala	Thr	Gly	Thr	Ser	Arg	Trp	Asp
			180					185					190		
Gly	Val	Ser	Lys	Lys	Ala	Pro	Arg	His	His	Leu	Ser	Val	Pro	Cys	Thr
		195					200					205			
Arg	Pro	Arg	Glu	Ala	Arg	Gln	Glu	Ala	Glu	Asp	Ser	Thr	Ser	Arg	Leu
		210				215					220				
Ser	Ala	Glu	Ser	Gly	Glu	Thr	Asp	Gln	Asp	Ala	Gly	Asp	Val	Gly	Pro
225				230					235					240	
Asp	Pro	Ile	Pro	Asp	Ser	Tyr	Tyr	Gly	Leu	Leu	Gly	Thr	Leu	Pro	Cys
				245				250						255	
Gln	Glu	Ala	Leu	Ser	His	Ile	Cys	Ser	Leu	Pro	Ser	Glu	Val	Leu	Arg
			260					265					270		
His	Val	Phe	Ala	Phe	Leu	Pro	Val	Glu	Asp	Leu	Tyr	Trp	Asn	Leu	Ser
		275					280					285			
Leu	Val	Cys	His	Leu	Trp	Arg	Glu	Ile	Ile	Ser	Asp	Pro	Leu	Phe	Ile
		290				295					300				
Pro	Trp	Lys	Lys	Leu	Tyr	His	Arg	Tyr	Leu	Met	Asn	Glu	Glu	Gln	Ala
305					310					315					320
Val	Ser	Lys	Val	Asp	Gly	Ile	Leu	Ser	Asn	Cys	Gly	Ile	Glu	Lys	Glu
				325					330					335	
Ser	Asp	Leu	Cys	Val	Leu	Asn	Leu	Ile	Arg	Tyr	Thr	Ala	Thr	Thr	Lys
			340					345					350		
Cys	Ser	Pro	Ser	Val	Asp	Pro	Glu	Arg	Val	Leu	Trp	Ser	Leu	Arg	Asp
		355					360					365			
His	Pro	Leu	Leu	Pro	Glu	Ala	Glu	Ala	Cys	Val	Arg	Gln	His	Leu	Pro
					375						380				
Asp	Leu	Tyr	Ala	Ala	Ala	Gly	Gly	Val	Asn	Ile	Trp	Ala	Leu	Val	Ala
385					390					395					400
Ala	Val	Val	Leu	Leu	Ser	Ser	Ser	Val	Asn	Asp	Ile	Gln	Arg	Leu	Leu
				405					410					415	
Phe	Cys	Leu	Arg	Arg	Pro	Ser	Ser	Thr	Val	Thr	Met	Pro	Asp	Val	Thr
			420					425					430		
Glu	Thr	Leu	Tyr	Cys	Ile	Ala	Val	Leu	Leu	Tyr	Ala	Met	Arg	Glu	Lys
			435				440					445			
Gly	Ile	Asn	I												

Glu	Lys	Trp	Ser	Gln	Ser	Arg	Phe	Leu	Tyr	Val	Thr	Phe	Asn	Lys	Ser	530	535	540
Ile	Ala	Lys	Gln	Ala	Glu	Arg	Val	Phe	Pro	Ser	Asn	Val	Ile	Cys	Lys	545	550	555
Thr	Phe	His	Ser	Met	Ala	Tyr	Gly	His	Ile	Gly	Arg	Lys	Tyr	Gln	Ser	565	570	575
Lys	Lys	Lys	Leu	Asn	Leu	Phe	Lys	Leu	Thr	Pro	Phe	Met	Val	Asn	Ser	580	585	590
Val	Leu	Ala	Glu	Gly	Lys	Gly	Gly	Phe	Ile	Arg	Ala	Lys	Leu	Val	Cys	595	600	605
Lys	Thr	Leu	Glu	Asn	Phe	Phe	Ala	Ser	Ala	Asp	Glu	Glu	Leu	Thr	Ile	610	615	620
Asp	His	Val	Pro	Ile	Trp	Cys	Lys	Asn	Ser	Gln	Gly	Gln	Arg	Val	Met	625	630	635
Val	Glu	Gln	Ser	Glu	Lys	Leu	Asn	Gly	Val	Leu	Glu	Ala	Ser	Arg	Leu	645	650	655
Trp	Asp	Asn	Met	Arg	Lys	Leu	Gly	Glu	Cys	Thr	Glu	Glu	Ala	His	Gln	660	665	670
Met	Thr	His	Asp	Gly	Tyr	Leu	Lys	Leu	Trp	Gln	Leu	Ser	Lys	Pro	Ser	675	680	685
Leu	Ala	Ser	Phe	Asp	Ala	Ile	Phe	Val	Asp	Glu	Ala	Gln	Asp	Cys	Thr	690	695	700
Pro	Ala	Ile	Met	Asn	Ile	Val	Leu	Ser	Gln	Pro	Cys	Gly	Lys	Ile	Phe	705	710	715
Val	Gly	Asp	Pro	His	Gln	Gln	Ile	Tyr	Thr	Phe	Arg	Gly	Ala	Val	Asn	725	730	735
Ala	Leu	Phe	Thr	Val	Pro	His	Thr	His	Val	Phe	Tyr	Leu	Thr	Gln	Ser	740	745	750
Phe	Arg	Phe	Gly	Val	Glu	Ile	Ala	Tyr	Val	Gly	Ala	Thr	Ile	Leu	Asp	755	760	765
Val	Cys	Lys	Arg	Val	Arg	Lys	Lys	Thr	Leu	Val	Gly	Gly	Asn	His	Gln	770	775	780
Ser	Gly	Ile	Arg	Gly	Asp	Ala	Lys	Gly	Gln	Val	Ala	Leu	Leu	Ser	Arg	785	790	795
Thr	Asn	Ala	Asn	Val	Phe	Asp	Glu	Ala	Val	Arg	Val	Thr	Glu	Gly	Glu	805	810	815
Phe	Pro	Ser	Arg	Ile	His	Leu	Ile	Gly	Gly	Ile	Lys	Ser	Phe	Gly	Leu	820	825	830
Asp	Arg	Ile	Ile	Asp	Ile	Trp	Ile	Leu	Leu	Gln	Pro	Glu	Glu	Glu	Arg	835	840	845
Arg	Lys	Gln	Asn	Leu	Val	Ile	Lys	Asp	Lys	Phe	Ile	Arg	Arg	Trp	Val	850	855	860
His	Lys	Glu	Gly	Phe	Ser	Gly	Phe	Lys	Arg	Tyr	Val	Thr	Ala	Ala	Glu	865	870	875
Asp	Lys	Glu	Leu	Glu	Ala	Lys	Ile	Ala	Val	Val	Glu	Lys	Tyr	Asn	Ile	885	890	895
Arg	Ile	Pro	Glu	Leu	Val	Gln	Arg	Ile	Glu	Lys	Cys	His	Ile	Glu	Asp	900	905	910
Leu	Asp	Phe	Ala	Glu	Tyr	Ile	Leu	Gly	Thr	Val	His	Lys	Ala	Lys	Gly	915	920	925
Leu	Glu	Phe	Asp	Thr	Val	His	Val	Leu	Asp	Asp	Phe	Val	Lys	Val	Pro	930	935	940
Cys	Ala	Arg	His	Asn	Leu	Pro	Gln	Leu	Pro	His	Phe	Arg	Val	Glu	Ser	945	950	955

Phe	Ser	Glu	Asp	Glu	Trp	Asn	Leu	Leu	Tyr	Val	Ala	Val	Thr	Arg	Ala
				965					970					975	
Lys	Lys	Arg	Leu	Ile	Met	Thr	Lys	Ser	Leu	Glu	Asn	Ile	Leu	Thr	Leu
			980					985					990		
Ala	Gly	Glu	Tyr	Phe	Leu	Gln	Ala	Glu	Leu	Thr	Ser	Asn	Val	Leu	Lys
		995					1000					1005			
Thr	Gly	Val	Val	Arg	Cys	Cys	Val	Gly	Gln	Cys	Asn	Asn	Ala	Ile	Pro
	1010				1015						1020				
Val	Asp	Thr	Val	Leu	Thr	Met	Lys	Lys	Leu	Pro	Ile	Thr	Tyr	Ser	Asn
1025					1030					1035					1040
Arg	Lys	Glu	Asn	Lys	Gly	Gly	Tyr	Leu	Cys	His	Ser	Cys	Ala	Glu	Gln
			1045						1050					1055	
Arg	Ile	Gly	Pro	Leu	Ala	Phe	Leu	Thr	Ala	Ser	Pro	Glu	Gln	Val	Arg
			1060					1065					1070		
Ala	Met	Glu	Arg	Thr	Val	Glu	Asn	Ile	Val	Leu	Pro	Arg	His	Glu	Ala
		1075					1080					1085			
Leu	Leu	Phe	Leu	Val	Phe										
			1090												

<210> 156

<211> 707

<212> PRT

<213> Homo sapiens

<220>

<223> >FBXW7|ENSG00000109670|ENST00000281708|2124

<400> 156

Met	Asn	Gln	Glu	Leu	Leu	Ser	Val	Gly	Ser	Lys	Arg	Arg	Arg	Thr	Gly
1				5					10					15	
Gly	Ser	Leu	Arg	Gly	Asn	Pro	Ser	Ser	Ser	Gln	Val	Asp	Glu	Glu	Gln
			20					25					30		
Met	Asn	Arg	Val	Val	Glu	Glu	Glu	Gln	Gln	Gln	Gln	Leu	Arg	Gln	Gln
		35					40					45			
Glu	Glu	Glu	His	Thr	Ala	Arg	Asn	Gly	Glu	Val	Val	Gly	Val	Glu	Pro
	50					55					60				
Arg	Pro	Gly	Gly	Gln	Asn	Asp	Ser	Gln	Gln	Gly	Gln	Leu	Glu	Glu	Asn
65					70					75					80
Asn	Asn	Arg	Phe	Ile	Ser	Val	Asp	Glu	Asp	Ser	Ser	Gly	Asn	Gln	Glu
			85						90				95		
Glu	Gln	Glu	Glu	Asp	Glu	Glu	His	Ala	Gly	Glu	Gln	Asp	Glu	Glu	Asp
			100					105					110		
Glu	Glu	Glu	Glu	Glu	Met	Asp	Gln	Glu	Ser	Asp	Asp	Phe	Asp	Gln	Ser
		115					120					125			
Asp	Asp	Ser	Ser	Arg	Glu	Asp	Glu	His	Thr	His	Thr	Asn	Ser	Val	Thr
	130					135					140				
Asn	Ser	Ser	Ser	Ile	Val	Asp	Leu	Pro	Val	His	Gln	Leu	Ser	Ser	Pro
145					150					155					160
Phe	Tyr	Thr	Lys	Thr	Thr	Lys	Met	Lys	Arg	Lys	Leu	Asp	His	Gly	Ser
			165						170					175	
Glu	Val	Arg	Ser	Phe	Ser	Leu	Gly	Lys	Lys	Pro	Cys	Lys	Val	Ser	Glu
		180						185					190		
Tyr	Thr	Ser	Thr	Thr	Gly	Leu	Val	Pro	Cys	Ser	Ala	Thr	Pro	Thr	Thr

625 630 635 640
 Ser Asp Asp Gly Thr Val Lys Leu Trp Asp Leu Lys Thr Gly Glu Phe
 645 650 655
 Ile Arg Asn Leu Val Thr Leu Glu Ser Gly Gly Ser Gly Gly Val Val
 660 665 670
 Trp Arg Ile Arg Ala Ser Asn Thr Lys Leu Val Cys Ala Val Gly Ser
 675 680 685
 Arg Asn Gly Thr Glu Glu Thr Lys Leu Leu Val Leu Asp Phe Asp Val
 690 700
 Asp Met Lys
 705

<210> 157
 <211> 380
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >FEN1|ENSG00000168496|ENST00000305885|1143

<400> 157
 Met Gly Ile Gln Gly Leu Ala Lys Leu Ile Ala Asp Val Ala Pro Ser
 1 5 10 15
 Ala Ile Arg Glu Asn Asp Ile Lys Ser Tyr Phe Gly Arg Lys Val Ala
 20 25 30
 Ile Asp Ala Ser Met Ser Ile Tyr Gln Phe Leu Ile Ala Val Arg Gln
 35 40 45
 Gly Gly Asp Val Leu Gln Asn Glu Glu Gly Glu Thr Thr Ser His Leu
 50 55 60
 Met Gly Met Phe Tyr Arg Thr Ile Arg Met Met Glu Asn Gly Ile Lys
 65 70 75 80
 Pro Val Tyr Val Phe Asp Gly Lys Pro Pro Gln Leu Lys Ser Gly Glu
 85 90 95
 Leu Ala Lys Arg Ser Glu Arg Arg Ala Glu Ala Glu Lys Gln Leu Gln
 100 105 110
 Gln Ala Gln Ala Ala Gly Ala Glu Gln Glu Val Glu Lys Phe Thr Lys
 115 120 125
 Arg Leu Val Lys Val Thr Lys Gln His Asn Asp Glu Cys Lys His Leu
 130 135 140
 Leu Ser Leu Met Gly Ile Pro Tyr Leu Asp Ala Pro Ser Glu Ala Glu
 145 150 155 160
 Ala Ser Cys Ala Ala Leu Val Lys Ala Gly Lys Val Tyr Ala Ala Ala
 165 170 175
 Thr Glu Asp Met Asp Cys Leu Thr Phe Gly Ser Pro Val Leu Met Arg
 180 185 190
 His Leu Thr Ala Ser Glu Ala Lys Lys Leu Pro Ile Gln Glu Phe His
 195 200 205
 Leu Ser Arg Ile Leu Gln Glu Leu Gly Leu Asn Gln Glu Gln Phe Val
 210 215 220
 Asp Leu Cys Ile Leu Leu Gly Ser Asp Tyr Cys Glu Ser Ile Arg Gly
 225 230 235 240
 Ile Gly Pro Lys Arg Ala Val Asp Leu Ile Gln Lys His Lys Ser Ile
 245 250 255

Glu	Glu	Ile	Val	Arg	Arg	Leu	Asp	Pro	Asn	Lys	Tyr	Pro	Val	Pro	Glu
			260					265					270		
Asn	Trp	Leu	His	Lys	Glu	Ala	His	Gln	Leu	Phe	Leu	Glu	Pro	Glu	Val
		275					280					285			
Leu	Asp	Pro	Glu	Ser	Val	Glu	Leu	Lys	Trp	Ser	Glu	Pro	Asn	Glu	Glu
	290					295					300				
Glu	Leu	Ile	Lys	Phe	Met	Cys	Gly	Glu	Lys	Gln	Phe	Ser	Glu	Glu	Arg
305					310					315					320
Ile	Arg	Ser	Gly	Val	Lys	Arg	Leu	Ser	Lys	Ser	Arg	Gln	Gly	Ser	Thr
			325						330					335	
Gln	Gly	Arg	Leu	Asp	Asp	Phe	Phe	Lys	Val	Thr	Gly	Ser	Leu	Ser	Ser
			340					345					350		
Ala	Lys	Arg	Lys	Glu	Pro	Glu	Pro	Lys	Gly	Ser	Thr	Lys	Lys	Lys	Ala
		355					360					365			
Lys	Thr	Gly	Ala	Ala	Gly	Lys	Phe	Lys	Arg	Gly	Lys				
	370					375					380				

<210> 158

<211> 908

<212> PRT

<213> Homo sapiens

<220>

<223> >GEN1|ENSG00000178295|ENST00000381254|2727

<400> 158

Met	Gly	Val	Asn	Asp	Leu	Trp	Gln	Ile	Leu	Glu	Pro	Val	Lys	Gln	His
1			5					10					15		
Ile	Pro	Leu	Arg	Asn	Leu	Gly	Gly	Lys	Thr	Ile	Ala	Val	Asp	Leu	Ser
		20					25					30			
Leu	Trp	Val	Cys	Glu	Ala	Gln	Thr	Val	Lys	Lys	Met	Met	Gly	Ser	Val
	35					40					45				
Met	Lys	Pro	His	Leu	Arg	Asn	Leu	Phe	Phe	Arg	Ile	Ser	Tyr	Leu	Thr
	50				55					60					
Gln	Met	Asp	Val	Lys	Leu	Val	Phe	Val	Met	Glu	Gly	Glu	Pro	Pro	Lys
65			70					75							80
Leu	Lys	Ala	Asp	Val	Ile	Ser	Lys	Arg	Asn	Gln	Ser	Arg	Tyr	Gly	Ser
			85				90						95		
Ser	Gly	Lys	Ser	Trp	Ser	Gln	Lys	Thr	Gly	Arg	Ser	His	Phe	Lys	Ser
		100				105						110			
Val	Leu	Arg	Glu	Cys	Leu	His	Met	Leu	Glu	Cys	Leu	Gly	Ile	Pro	Trp
	115				120							125			
Val	Gln	Ala	Ala	Gly	Glu	Ala	Glu	Ala	Met	Cys	Ala	Tyr	Leu	Asn	Ala
	130				135						140				
Gly	Gly	His	Val	Asp	Gly	Cys	Leu	Thr	Asn	Asp	Gly	Asp	Thr	Phe	Leu
145				150					155						160
Tyr	Gly	Ala	Gln	Thr	Val	Tyr	Arg	Asn	Phe	Thr	Met	Asn	Thr	Lys	Asp
			165					170						175	
Pro	His	Val	Asp	Cys	Tyr	Thr	Met	Ser	Ser	Ile	Lys	Ser	Lys	Leu	Gly
		180					185						190		
Leu	Asp	Arg	Asp	Ala	Leu	Val	Gly	Leu	Ala	Ile	Leu	Leu	Gly	Cys	Asp
	195					200					205				
Tyr	Leu	Pro	Lys	Gly	Val	Pro	Gly	Val	Gly	Lys	Glu	Gln	Ala	Leu	Lys

210	215	220
Leu Ile Gln Ile Leu Lys Gly Gln Ser Leu Leu Gln Arg Phe Asn Arg		
225	230	235
Trp Asn Glu Thr Ser Cys Asn Ser Ser Pro Gln Leu Leu Val Thr Lys		240
	245	250
Lys Leu Ala His Cys Ser Val Cys Ser His Pro Gly Ser Pro Lys Asp		255
	260	265
His Glu Arg Asn Gly Cys Arg Leu Cys Lys Ser Asp Lys Tyr Cys Glu		270
	275	280
Pro His Asp Tyr Glu Tyr Cys Cys Pro Cys Glu Trp His Arg Thr Glu		285
	290	295
His Asp Arg Gln Leu Ser Glu Val Glu Asn Asn Ile Lys Lys Lys Ala		300
305	310	315
Cys Cys Cys Glu Gly Phe Pro Phe His Glu Val Ile Gln Glu Phe Leu		320
	325	330
Leu Asn Lys Asp Lys Leu Val Lys Val Ile Arg Tyr Gln Arg Pro Asp		335
	340	345
Leu Leu Leu Phe Gln Arg Phe Thr Leu Glu Lys Met Glu Trp Pro Asn		350
	355	360
His Tyr Ala Cys Glu Lys Leu Leu Val Leu Leu Thr His Tyr Asp Met		365
	370	375
Ile Glu Arg Lys Leu Gly Ser Arg Asn Ser Asn Gln Leu Gln Pro Ile		380
385	390	395
Arg Ile Val Lys Thr Arg Ile Arg Asn Gly Val His Cys Phe Glu Ile		400
	405	410
Glu Trp Glu Lys Pro Glu His Tyr Ala Met Glu Asp Lys Gln His Gly		415
	420	425
Glu Phe Ala Leu Leu Thr Ile Glu Glu Glu Ser Leu Phe Glu Ala Ala		430
	435	440
Tyr Pro Glu Ile Val Ala Val Tyr Gln Lys Gln Lys Leu Glu Ile Lys		445
	450	455
Gly Lys Lys Gln Lys Arg Ile Lys Pro Lys Glu Asn Asn Leu Pro Glu		460
465	470	475
Pro Asp Glu Val Met Ser Phe Gln Ser His Met Thr Leu Lys Pro Thr		480
	485	490
Cys Glu Ile Phe His Lys Gln Asn Ser Lys Leu Asn Ser Gly Ile Ser		495
	500	505
Pro Asp Pro Thr Leu Pro Gln Glu Ser Ile Ser Ala Ser Leu Asn Ser		510
	515	520
Leu Leu Leu Pro Lys Asn Thr Pro Cys Leu Asn Ala Gln Glu Gln Phe		525
	530	535
Met Ser Ser Leu Arg Pro Leu Ala Ile Gln Gln Ile Lys Ala Val Ser		540
545	550	555
Lys Ser Leu Ile Ser Glu Ser Ser Gln Pro Asn Thr Ser Ser His Asn		560
	565	570
Ile Ser Val Ile Ala Asp Leu His Leu Ser Thr Ile Asp Trp Glu Gly		575
	580	585
Thr Ser Phe Ser Asn Ser Pro Ala Ile Gln Arg Asn Thr Phe Ser His		590
	595	600
Asp Leu Lys Ser Glu Val Glu Ser Glu Leu Ser Ala Ile Pro Asp Gly		605
	610	615
Phe Glu Asn Ile Pro Glu Gln Leu Ser Cys Glu Ser Glu Arg Tyr Thr		620
625	630	635
Ala Asn Ile Lys Lys Val Leu Asp Glu Asp Ser Asp Gly Ile Ser Pro		640

Pro	Arg	His	Leu	Gln	Leu	Ala	Ile	Arg	Asn	Asp	Glu	Glu	Leu	Asn	Lys
				85					90					95	
Leu	Leu	Gly	Gly	Val	Thr	Ile	Ala	Gln	Gly	Gly	Val	Leu	Pro	Asn	Ile
			100					105					110		
Gln	Ala	Val	Leu	Leu	Pro	Lys	Lys	Thr	Ser	Ala	Thr	Val	Gly	Pro	Lys
		115					120					125			
Ala	Pro	Ser	Gly	Gly	Lys	Lys	Ala	Thr	Gln	Ala	Ser	Gln	Glu	Tyr	
	130					135					140				

<210> 160
 <211> 488
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >HDAC2|ENSG00000196591|ENST00000519065|1467

<400> 160

Met	Ala	Tyr	Ser	Gln	Gly	Gly	Gly	Lys	Lys	Lys	Val	Cys	Tyr	Tyr	Tyr
1				5				10						15	
Asp	Gly	Asp	Ile	Gly	Asn	Tyr	Tyr	Tyr	Gly	Gln	Gly	His	Pro	Met	Lys
			20				25						30		
Pro	His	Arg	Ile	Arg	Met	Thr	His	Asn	Leu	Leu	Leu	Asn	Tyr	Gly	Leu
		35				40						45			
Tyr	Arg	Lys	Met	Glu	Ile	Tyr	Arg	Pro	His	Lys	Ala	Thr	Ala	Glu	Glu
	50					55				60					
Met	Thr	Lys	Tyr	His	Ser	Asp	Glu	Tyr	Ile	Lys	Phe	Leu	Arg	Ser	Ile
65				70					75					80	
Arg	Pro	Asp	Asn	Met	Ser	Glu	Tyr	Ser	Lys	Gln	Met	Gln	Arg	Phe	Asn
			85					90						95	
Val	Gly	Glu	Asp	Cys	Pro	Val	Phe	Asp	Gly	Leu	Phe	Glu	Phe	Cys	Gln
			100					105					110		
Leu	Ser	Thr	Gly	Gly	Ser	Val	Ala	Gly	Ala	Val	Lys	Leu	Asn	Arg	Gln
		115				120						125			
Gln	Thr	Asp	Met	Ala	Val	Asn	Trp	Ala	Gly	Gly	Leu	His	His	Ala	Lys
	130					135					140				
Lys	Ser	Glu	Ala	Ser	Gly	Phe	Cys	Tyr	Val	Asn	Asp	Ile	Val	Leu	Ala
145				150					155					160	
Ile	Leu	Glu	Leu	Leu	Lys	Tyr	His	Gln	Arg	Val	Leu	Tyr	Ile	Asp	Ile
			165					170					175		
Asp	Ile	His	His	Gly	Asp	Gly	Val	Glu	Glu	Ala	Phe	Tyr	Thr	Thr	Asp
		180						185					190		
Arg	Val	Met	Thr	Val	Ser	Phe	His	Lys	Tyr	Gly	Glu	Tyr	Phe	Pro	Gly
		195				200						205			
Thr	Gly	Asp	Leu	Arg	Asp	Ile	Gly	Ala	Gly	Lys	Gly	Lys	Tyr	Tyr	Ala
	210				215						220				
Val	Asn	Phe	Pro	Met	Arg	Asp	Gly	Ile	Asp	Asp	Glu	Ser	Tyr	Gly	Gln
225				230					235					240	
Ile	Phe	Lys	Pro	Ile	Ile	Ser	Lys	Val	Met	Glu	Met	Tyr	Gln	Pro	Ser
			245					250					255		
Ala	Val	Val	Leu	Gln	Cys	Gly	Ala	Asp	Ser	Leu	Ser	Gly	Asp	Arg	Leu
		260					265					270			
Gly	Cys	Phe	Asn	Leu	Thr	Val	Lys	Gly	His	Ala	Lys	Cys	Val	Glu	Val

Gln Ala Gln Asp Leu Ala Arg Ser Tyr Gly Ile Pro Tyr Ile Glu Thr
130 135 140
Ser Ala Lys Thr Arg Gln Gly Val Glu Asp Ala Phe Tyr Thr Leu Val
145 150 155 160
Arg Glu Ile Arg Gln His Lys Leu Arg Lys Leu Asn Pro Pro Asp Glu
165 170 175
Ser Gly Pro Gly Cys Met Ser Cys Lys Cys Val Leu Ser
180 185

<210> 162
<211> 189
<212> PRT
<213> Homo sapiens

<220>
<223> >KRAS|ENSG00000133703|ENST00000256078|570

<400> 162
Met Thr Glu Tyr Lys Leu Val Val Val Gly Ala Gly Gly Val Gly Lys
1 5 10 15
Ser Ala Leu Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr
20 25 30
Asp Pro Thr Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly
35 40 45
Glu Thr Cys Leu Leu Asp Ile Leu Asp Thr Ala Gly Gln Glu Glu Tyr
50 55 60
Ser Ala Met Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys
65 70 75 80
Val Phe Ala Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr
85 90 95
Arg Glu Gln Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val
100 105 110
Leu Val Gly Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys
115 120 125
Gln Ala Gln Asp Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr
130 135 140
Ser Ala Lys Thr Arg Gln Arg Val Glu Asp Ala Phe Tyr Thr Leu Val
145 150 155 160
Arg Glu Ile Arg Gln Tyr Arg Leu Lys Lys Ile Ser Lys Glu Glu Lys
165 170 175
Thr Pro Gly Cys Val Lys Ile Lys Lys Cys Ile Ile Met
180 185

<210> 163
<211> 911
<212> PRT
<213> Homo sapiens

<220>
<223> >LIG4|ENSG00000174405|ENST00000356922|2736

<400> 163

Met	Ala	Ala	Ser	Gln	Thr	Ser	Gln	Thr	Val	Ala	Ser	His	Val	Pro	Phe	1	5	10	15
Ala	Asp	Leu	Cys	Ser	Thr	Leu	Glu	Arg	Ile	Gln	Lys	Ser	Lys	Gly	Arg	20	25	30	
Ala	Glu	Lys	Ile	Arg	His	Phe	Arg	Glu	Phe	Leu	Asp	Ser	Trp	Arg	Lys	35	40	45	
Phe	His	Asp	Ala	Leu	His	Lys	Asn	His	Lys	Asp	Val	Thr	Asp	Ser	Phe	50	55	60	
Tyr	Pro	Ala	Met	Arg	Leu	Ile	Leu	Pro	Gln	Leu	Glu	Arg	Glu	Arg	Met	65	70	75	80
Ala	Tyr	Gly	Ile	Lys	Glu	Thr	Met	Leu	Ala	Lys	Leu	Tyr	Ile	Glu	Leu	85	90	95	
Leu	Asn	Leu	Pro	Arg	Asp	Gly	Lys	Asp	Ala	Leu	Lys	Leu	Leu	Asn	Tyr	100	105	110	
Arg	Thr	Pro	Thr	Gly	Thr	His	Gly	Asp	Ala	Gly	Asp	Phe	Ala	Met	Ile	115	120	125	
Ala	Tyr	Phe	Val	Leu	Lys	Pro	Arg	Cys	Leu	Gln	Lys	Gly	Ser	Leu	Thr	130	135	140	
Ile	Gln	Gln	Val	Asn	Asp	Leu	Leu	Asp	Ser	Ile	Ala	Ser	Asn	Asn	Ser	145	150	155	160
Ala	Lys	Arg	Lys	Asp	Leu	Ile	Lys	Lys	Ser	Leu	Leu	Gln	Leu	Ile	Thr	165	170	175	
Gln	Ser	Ser	Ala	Leu	Glu	Gln	Lys	Trp	Leu	Ile	Arg	Met	Ile	Ile	Lys	180	185	190	
Asp	Leu	Lys	Leu	Gly	Val	Ser	Gln	Gln	Thr	Ile	Phe	Ser	Val	Phe	His	195	200	205	
Asn	Asp	Ala	Ala	Glu	Leu	His	Asn	Val	Thr	Thr	Asp	Leu	Glu	Lys	Val	210	215	220	
Cys	Arg	Gln	Leu	His	Asp	Pro	Ser	Val	Gly	Leu	Ser	Asp	Ile	Ser	Ile	225	230	235	240
Thr	Leu	Phe	Ser	Ala	Phe	Lys	Pro	Met	Leu	Ala	Ala	Ile	Ala	Asp	Ile	245	250	255	
Glu	His	Ile	Glu	Lys	Asp	Met	Lys	His	Gln	Ser	Phe	Tyr	Ile	Glu	Thr	260	265	270	
Lys	Leu	Asp	Gly	Glu	Arg	Met	Gln	Met	His	Lys	Asp	Gly	Asp	Val	Tyr	275	280	285	
Lys	Tyr	Phe	Ser	Arg	Asn	Gly	Tyr	Asn	Tyr	Thr	Asp	Gln	Phe	Gly	Ala	290	295	300	
Ser	Pro	Thr	Glu	Gly	Ser	Leu	Thr	Pro	Phe	Ile	His	Asn	Ala	Phe	Lys	305	310	315	320
Ala	Asp	Ile	Gln	Ile	Cys	Ile	Leu	Asp	Gly	Glu	Met	Met	Ala	Tyr	Asn	325	330	335	
Pro	Asn	Thr	Gln	Thr	Phe	Met	Gln	Lys	Gly	Thr	Lys	Phe	Asp	Ile	Lys	340	345	350	
Arg	Met	Val	Glu	Asp	Ser	Asp	Leu	Gln	Thr	Cys	Tyr	Cys	Val	Phe	Asp	355	360	365	
Val	Leu	Met	Val	Asn	Asn	Lys	Lys	Leu	Gly	His	Glu	Thr	Leu	Arg	Lys	370	375	380	
Arg	Tyr	Glu	Ile	Leu	Ser	Ser	Ile	Phe	Thr	Pro	Ile	Pro	Gly	Arg	Ile	385	390	395	400
Glu	Ile	Val	Gln	Lys	Thr	Gln	Ala	His	Thr	Lys	Asn	Glu	Val	Ile	Asp	405	410	415	
Ala	Leu	Asn	Glu	Ala	Ile	Asp	Lys	Arg	Glu	Glu	Gly	Ile	Met	Val	Lys	420	425	430	

Gln	Pro	Leu	Ser	Ile	Tyr	Lys	Pro	Asp	Lys	Arg	Gly	Glu	Gly	Trp	Leu	435	440	445
Lys	Ile	Lys	Pro	Glu	Tyr	Val	Ser	Gly	Leu	Met	Asp	Glu	Leu	Asp	Ile	450	455	460
Leu	Ile	Val	Gly	Gly	Tyr	Trp	Gly	Lys	Gly	Ser	Arg	Gly	Gly	Met	Met	465	470	475
Ser	His	Phe	Leu	Cys	Ala	Val	Ala	Glu	Lys	Pro	Pro	Pro	Gly	Glu	Lys	485	490	495
Pro	Ser	Val	Phe	His	Thr	Leu	Ser	Arg	Val	Gly	Ser	Gly	Cys	Thr	Met	500	505	510
Lys	Glu	Leu	Tyr	Asp	Leu	Gly	Leu	Lys	Leu	Ala	Lys	Tyr	Trp	Lys	Pro	515	520	525
Phe	His	Arg	Lys	Ala	Pro	Pro	Ser	Ser	Ile	Leu	Cys	Gly	Thr	Glu	Lys	530	535	540
Pro	Glu	Val	Tyr	Ile	Glu	Pro	Cys	Asn	Ser	Val	Ile	Val	Gln	Ile	Lys	545	550	555
Ala	Ala	Glu	Ile	Val	Pro	Ser	Asp	Met	Tyr	Lys	Thr	Gly	Cys	Thr	Leu	565	570	575
Arg	Phe	Pro	Arg	Ile	Glu	Lys	Ile	Arg	Asp	Asp	Lys	Glu	Trp	His	Glu	580	585	590
Cys	Met	Thr	Leu	Asp	Asp	Leu	Glu	Gln	Leu	Arg	Gly	Lys	Ala	Ser	Gly	595	600	605
Lys	Leu	Ala	Ser	Lys	His	Leu	Tyr	Ile	Gly	Gly	Asp	Asp	Glu	Pro	Gln	610	615	620
Glu	Lys	Lys	Arg	Lys	Ala	Ala	Pro	Lys	Met	Lys	Lys	Val	Ile	Gly	Ile	625	630	635
Ile	Glu	His	Leu	Lys	Ala	Pro	Asn	Leu	Thr	Asn	Val	Asn	Lys	Ile	Ser	645	650	655
Asn	Ile	Phe	Glu	Asp	Val	Glu	Phe	Cys	Val	Met	Ser	Gly	Thr	Asp	Ser	660	665	670
Gln	Pro	Lys	Pro	Asp	Leu	Glu	Asn	Arg	Ile	Ala	Glu	Phe	Gly	Gly	Tyr	675	680	685
Ile	Val	Gln	Asn	Pro	Gly	Pro	Asp	Thr	Tyr	Cys	Val	Ile	Ala	Gly	Ser	690	695	700
Glu	Asn	Ile	Arg	Val	Lys	Asn	Ile	Ile	Leu	Ser	Asn	Lys	His	Asp	Val	705	710	715
Val	Lys	Pro	Ala	Trp	Leu	Leu	Glu	Cys	Phe	Lys	Thr	Lys	Ser	Phe	Val	725	730	735
Pro	Trp	Gln	Pro	Arg	Phe	Met	Ile	His	Met	Cys	Pro	Ser	Thr	Lys	Glu	740	745	750
His	Phe	Ala	Arg	Glu	Tyr	Asp	Cys	Tyr	Gly	Asp	Ser	Tyr	Phe	Ile	Asp	755	760	765
Thr	Asp	Leu	Asn	Gln	Leu	Lys	Glu	Val	Phe	Ser	Gly	Ile	Lys	Asn	Ser	770	775	780
Asn	Glu	Gln	Thr	Pro	Glu	Glu	Met	Ala	Ser	Leu	Ile	Ala	Asp	Leu	Glu	785	790	795
Tyr	Arg	Tyr	Ser	Trp	Asp	Cys	Ser	Pro	Leu	Ser	Met	Phe	Arg	Arg	His	805	810	815
Thr	Val	Tyr	Leu	Asp	Ser	Tyr	Ala	Val	Ile	Asn	Asp	Leu	Ser	Thr	Lys	820	825	830
Asn	Glu	Gly	Thr	Arg	Leu	Ala	Ile	Lys	Ala	Leu	Glu	Leu	Arg	Phe	His	835	840	845
Gly	Ala	Lys	Val	Val	Ser	Cys	Leu	Ala	Glu	Gly	Val	Ser	His	Val	Ile	850	855	860

Ile	Gly	Glu	Asp	His	Ser	Arg	Val	Ala	Asp	Phe	Lys	Ala	Phe	Arg	Arg
865					870				875					880	
Thr	Phe	Lys	Arg	Lys	Phe	Lys	Ile	Leu	Lys	Glu	Ser	Trp	Val	Thr	Asp
				885					890					895	
Ser	Ile	Asp	Lys	Cys	Glu	Leu	Gln	Glu	Glu	Asn	Gln	Tyr	Leu	Ile	
			900					905					910		

<210> 164
 <211> 2089
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >MDC1|ENSG00000137337|ENST00000376406|6270

<400> 164

Met	Glu	Asp	Thr	Gln	Ala	Ile	Asp	Trp	Asp	Val	Glu	Glu	Glu	Glu	Glu
1				5				10					15		
Thr	Glu	Gln	Ser	Ser	Glu	Ser	Leu	Arg	Cys	Asn	Val	Glu	Pro	Val	Gly
			20					25					30		
Arg	Leu	His	Ile	Phe	Ser	Gly	Ala	His	Gly	Pro	Glu	Lys	Asp	Phe	Pro
		35					40					45			
Leu	His	Leu	Gly	Lys	Asn	Val	Val	Gly	Arg	Met	Pro	Asp	Cys	Ser	Val
	50				55					60					
Ala	Leu	Pro	Phe	Pro	Ser	Ile	Ser	Lys	Gln	His	Ala	Glu	Ile	Glu	Ile
65					70				75					80	
Leu	Ala	Trp	Asp	Lys	Ala	Pro	Ile	Leu	Arg	Asp	Cys	Gly	Ser	Leu	Asn
				85					90					95	
Gly	Thr	Gln	Ile	Leu	Arg	Pro	Pro	Lys	Val	Leu	Ser	Pro	Gly	Val	Ser
			100					105					110		
His	Arg	Leu	Arg	Asp	Gln	Glu	Leu	Ile	Leu	Phe	Ala	Asp	Leu	Leu	Cys
		115					120					125			
Gln	Tyr	His	Arg	Leu	Asp	Val	Ser	Leu	Pro	Phe	Val	Ser	Arg	Gly	Pro
	130					135					140				
Leu	Thr	Val	Glu	Glu	Thr	Pro	Arg	Val	Gln	Gly	Glu	Thr	Gln	Pro	Gln
145					150					155					160
Arg	Leu	Leu	Leu	Ala	Glu	Asp	Ser	Glu	Glu	Val	Asp	Phe	Leu	Ser	
				165					170					175	
Glu	Arg	Arg	Met	Val	Lys	Lys	Ser	Arg	Thr	Thr	Ser	Ser	Ser	Val	Ile
			180					185					190		
Val	Pro	Glu	Ser	Asp	Glu	Glu	Gly	His	Ser	Pro	Val	Leu	Gly	Gly	Leu
		195					200					205			
Gly	Pro	Pro	Phe	Ala	Phe	Asn	Leu	Asn	Ser	Asp	Thr	Asp	Val	Glu	Glu
	210					215					220				
Gly	Gln	Gln	Pro	Ala	Thr	Glu	Glu	Ala	Ser	Ser	Ala	Ala	Arg	Arg	Gly
225					230					235					240
Ala	Thr	Val	Glu	Ala	Lys	Gln	Ser	Glu	Ala	Glu	Val	Val	Thr	Glu	Ile
				245					250					255	
Gln	Leu	Glu	Lys	Asp	Gln	Pro	Leu	Val	Lys	Glu	Arg	Asp	Asn	Asp	Thr
			260					265					270		
Lys	Val	Lys	Arg	Gly	Ala	Gly	Asn	Gly	Val	Val	Pro	Ala	Gly	Val	Ile
		275					280					285			
Leu	Glu	Arg	Ser	Gln	Pro	Pro	Gly	Glu	Asp	Ser	Asp	Thr	Asp	Val	Asp

290	295	300
Asp Asp Ser Arg Pro Pro Gly Arg Pro Ala Glu Val His Leu Glu Arg		
305	310	315
Ala Gln Pro Phe Gly Phe Ile Asp Ser Asp Thr Asp Ala Glu Glu Glu		
	325	330
Arg Ile Pro Ala Thr Pro Val Val Ile Pro Met Lys Lys Arg Lys Ile		
	340	345
Phe His Gly Val Gly Thr Arg Gly Pro Gly Ala Pro Gly Leu Ala His		
	355	360
Leu Gln Glu Ser Gln Ala Gly Ser Asp Thr Asp Val Glu Glu Gly Lys		
	370	375
Ala Pro Gln Ala Val Pro Leu Glu Lys Ser Gln Ala Ser Met Val Ile		
385	390	395
Asn Ser Asp Thr Asp Asp Glu Glu Glu Val Ser Ala Ala Leu Thr Leu		
	405	410
Ala His Leu Lys Glu Ser Gln Pro Ala Ile Trp Asn Arg Asp Ala Glu		
	420	425
Glu Asp Met Pro Gln Arg Val Val Leu Leu Gln Arg Ser Gln Thr Thr		
	435	440
Thr Glu Arg Asp Ser Asp Thr Asp Val Glu Glu Glu Glu Leu Pro Val		
	450	455
Glu Asn Arg Glu Ala Val Leu Lys Asp His Thr Lys Ile Arg Ala Leu		
465	470	475
Val Arg Ala His Ser Glu Lys Asp Gln Pro Pro Phe Gly Asp Ser Asp		
	485	490
Asp Ser Val Glu Ala Asp Lys Ser Ser Pro Gly Ile His Leu Glu Arg		
	500	505
Ser Gln Ala Ser Thr Thr Val Asp Ile Asn Thr Gln Val Glu Lys Glu		
	515	520
Val Pro Pro Gly Ser Ala Ile Ile His Ile Lys Lys His Gln Val Ser		
	530	535
Val Glu Gly Thr Asn Gln Thr Asp Val Lys Ala Val Gly Gly Pro Ala		
545	550	555
Lys Leu Leu Val Val Ser Leu Glu Glu Ala Trp Pro Leu His Gly Asp		
	565	570
Cys Glu Thr Asp Ala Glu Glu Gly Thr Ser Leu Thr Ala Ser Val Val		
	580	585
Ala Asp Val Arg Lys Ser Gln Leu Pro Ala Glu Gly Asp Ala Gly Ala		
	595	600
Glu Trp Ala Ala Ala Val Leu Lys Gln Glu Arg Ala His Glu Val Gly		
	610	615
Ala Gln Gly Gly Pro Pro Val Ala Gln Val Glu Gln Asp Leu Pro Ile		
625	630	635
Ser Arg Glu Asn Leu Thr Asp Leu Val Val Asp Thr Asp Thr Leu Gly		
	645	650
Glu Ser Thr Gln Pro Gln Arg Glu Gly Ala Gln Val Pro Thr Gly Arg		
	660	665
Glu Arg Glu Gln His Val Gly Gly Thr Lys Asp Ser Glu Asp Asn Tyr		
	675	680
Gly Asp Ser Glu Asp Leu Asp Leu Gln Ala Thr Gln Cys Phe Leu Glu		
	690	695
Asn Gln Gly Leu Glu Ala Val Gln Ser Met Glu Asp Glu Pro Thr Gln		
705	710	715
Ala Phe Met Leu Thr Pro Pro Gln Glu Leu Gly Pro Ser His Cys Ser		

1155					1160					1165					
Gln	Pro	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Ser	Glu	Pro	Thr	Ser
1170					1175					1180					
Gln	Val	Thr	Arg	Gly	Arg	Lys	Ser	Arg	Ser	Ser	Val	Lys	Thr	Pro	Glu
1185					1190					1195					1200
Thr	Val	Val	Pro	Thr	Ala	Leu	Glu	Leu	Gln	Pro	Ser	Thr	Ser	Thr	Asp
1205					1210					1215					
Arg	Pro	Val	Thr	Ser	Glu	Pro	Thr	Ser	Gln	Ala	Thr	Arg	Gly	Arg	Lys
1220					1225					1230					
Asn	Arg	Ser	Ser	Val	Lys	Thr	Pro	Glu	Pro	Val	Val	Pro	Thr	Ala	Pro
1235					1240					1245					
Glu	Leu	Gln	Pro	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Ser	Glu	Pro
1250					1255					1260					
Thr	Tyr	Gln	Ala	Thr	Arg	Gly	Arg	Lys	Asn	Arg	Ser	Ser	Val	Lys	Thr
1265					1270					1275					1280
Pro	Glu	Pro	Val	Val	Pro	Thr	Ala	Pro	Glu	Leu	Arg	Pro	Ser	Thr	Ser
1285					1290					1295					
Thr	Asp	Arg	Pro	Val	Thr	Pro	Lys	Pro	Thr	Ser	Arg	Thr	Thr	Arg	Ser
1300					1305					1310					
Arg	Thr	Asn	Met	Ser	Ser	Val	Lys	Thr	Pro	Glu	Thr	Val	Val	Pro	Thr
1315					1320					1325					
Ala	Pro	Glu	Leu	Gln	Ile	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Pro
1330					1335					1340					
Lys	Pro	Thr	Ser	Arg	Thr	Thr	Arg	Ser	Arg	Thr	Asn	Met	Ser	Ser	Val
1345					1350					1355					1360
Lys	Asn	Pro	Glu	Ser	Thr	Val	Pro	Ile	Ala	Pro	Glu	Leu	Pro	Pro	Ser
1365					1370					1375					
Thr	Ser	Thr	Glu	Gln	Pro	Val	Thr	Pro	Glu	Pro	Thr	Ser	Arg	Ala	Thr
1380					1385					1390					
Arg	Gly	Arg	Lys	Asn	Arg	Ser	Ser	Gly	Lys	Thr	Pro	Glu	Thr	Leu	Val
1395					1400					1405					
Pro	Thr	Ala	Pro	Lys	Leu	Glu	Pro	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val
1410					1415					1420					
Thr	Pro	Glu	Pro	Thr	Ser	Gln	Ala	Thr	Arg	Gly	Arg	Thr	Asn	Arg	Ser
1425					1430					1435					1440
Ser	Val	Lys	Thr	Pro	Glu	Thr	Val	Val	Pro	Thr	Ala	Pro	Glu	Leu	Gln
1445					1450					1455					
Pro	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Pro	Glu	Pro	Thr	Ser	Gln
1460					1465					1470					
Ala	Thr	Arg	Gly	Arg	Thr	Asp	Arg	Ser	Ser	Val	Lys	Thr	Pro	Glu	Thr
1475					1480					1485					
Val	Val	Pro	Thr	Ala	Pro	Glu	Leu	Gln	Ala	Ser	Ala	Ser	Thr	Asp	Gln
1490					1495					1500					
Pro	Val	Thr	Ser	Glu	Pro	Thr	Ser	Arg	Thr	Thr	Arg	Gly	Arg	Lys	Asn
1505					1510					1515					1520
Arg	Ser	Ser	Val	Lys	Thr	Pro	Glu	Thr	Val	Val	Pro	Ala	Ala	Pro	Glu
1525					1530					1535					
Leu	Gln	Pro	Ser	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Pro	Glu	Pro	Thr
1540					1545					1550					
Ser	Arg	Ala	Thr	Arg	Gly	Arg	Thr	Asn	Arg	Ser	Ser	Val	Lys	Thr	Pro
1555					1560					1565					
Glu	Ser	Ile	Val	Pro	Ile	Ala	Pro	Glu	Leu	Gln	Pro	Ser	Thr	Ser	Arg
1570					1575					1580					
Asn	Gln	Leu	Val	Thr	Pro	Glu	Pro	Thr	Ser	Arg	Ala	Thr	Arg	Cys	Arg

1585					1590					1595				1600
Thr	Asn	Arg	Ser	Ser	Val	Lys	Thr	Pro	Glu	Pro	Val	Val	Pro	Thr
					1605					1610				1615
Pro	Glu	Pro	His	Pro	Thr	Thr	Ser	Thr	Asp	Gln	Pro	Val	Thr	Pro
					1620					1625				1630
Leu	Thr	Ser	Arg	Ala	Thr	Arg	Arg	Lys	Thr	Asn	Arg	Ser	Ser	Val
					1635					1640				1645
Thr	Pro	Lys	Pro	Val	Glu	Pro	Ala	Ala	Ser	Asp	Leu	Glu	Pro	Phe
					1650					1655				1660
Pro	Thr	Asp	Gln	Ser	Val	Thr	Pro	Glu	Ala	Ile	Ala	Gln	Gly	Gly
					1665					1670				1675
Ser	Lys	Thr	Leu	Arg	Ser	Ser	Thr	Val	Arg	Ala	Met	Pro	Val	Pro
					1685					1690				1695
Thr	Pro	Glu	Phe	Gln	Ser	Pro	Val	Thr	Thr	Asp	Gln	Pro	Ile	Ser
					1700					1705				1710
Glu	Pro	Ile	Thr	Gln	Pro	Ser	Cys	Ile	Lys	Arg	Gln	Arg	Ala	Ala
					1715					1720				1725
Asn	Pro	Gly	Ser	Leu	Ala	Ala	Pro	Ile	Asp	His	Lys	Pro	Cys	Ser
					1730					1735				1740
Pro	Leu	Glu	Pro	Lys	Ser	Gln	Ala	Ser	Arg	Asn	Gln	Arg	Trp	Gly
					1745					1750				1755
Val	Arg	Ala	Ala	Glu	Ser	Leu	Thr	Ala	Ile	Pro	Glu	Pro	Ala	Ser
					1765					1770				1775
Gln	Leu	Leu	Glu	Thr	Pro	Ile	His	Ala	Ser	Gln	Ile	Gln	Lys	Val
					1780					1785				1790
Pro	Ala	Gly	Arg	Ser	Arg	Phe	Thr	Pro	Glu	Leu	Gln	Pro	Lys	Ala
					1795					1800				1805
Gln	Ser	Arg	Lys	Arg	Ser	Leu	Ala	Thr	Met	Asp	Ser	Pro	Pro	His
					1810					1815				1820
Lys	Gln	Pro	Gln	Arg	Gly	Glu	Val	Ser	Gln	Lys	Thr	Val	Ile	Ile
					1825					1830				1835
Glu	Glu	Glu	Glu	Asp	Thr	Ala	Glu	Lys	Pro	Gly	Lys	Glu	Glu	Asp
					1845					1850				1855
Val	Thr	Pro	Lys	Pro	Gly	Lys	Arg	Lys	Arg	Asp	Gln	Ala	Glu	Glu
					1860					1865				1870
Pro	Asn	Arg	Ile	Pro	Ser	Arg	Ser	Leu	Arg	Arg	Thr	Lys	Leu	Asn
					1875					1880				1885
Glu	Ser	Thr	Ala	Pro	Lys	Val	Leu	Phe	Thr	Gly	Val	Val	Asp	Ala
					1890					1895				1900
Gly	Glu	Arg	Ala	Val	Leu	Ala	Leu	Gly	Gly	Ser	Leu	Ala	Gly	Ser
					1905					1910				1915
Ala	Glu	Ala	Ser	His	Leu	Val	Thr	Asp	Arg	Ile	Arg	Arg	Thr	Val
					1925					1930				1935
Phe	Leu	Cys	Ala	Leu	Gly	Arg	Gly	Ile	Pro	Ile	Leu	Ser	Leu	Asp
					1940					1945				1950
Leu	His	Gln	Ser	Arg	Lys	Ala	Gly	Phe	Phe	Leu	Pro	Pro	Asp	Glu
					1955					1960				1965
Val	Val	Thr	Asp	Pro	Glu	Gln	Glu	Lys	Asn	Phe	Gly	Phe	Ser	Leu
					1970					1975				1980
Asp	Ala	Leu	Ser	Arg	Ala	Arg	Glu	Arg	Arg	Leu	Leu	Glu	Gly	Tyr
					1985					1990				1995
Ile	Tyr	Val	Thr	Pro	Gly	Val	Gln	Pro	Pro	Pro	Pro	Gln	Met	Gly
					2005					2010				2015
Ile	Ile	Ser	Cys	Cys	Gly	Gly	Thr	Tyr	Leu	Pro	Ser	Met	Pro	Arg

Arg	Lys	Ala	Ile	Glu	Thr	Val	Tyr	Ala	Ala	Tyr	Leu	Pro	Lys	Asn	Thr	275	280	285
His	Pro	Phe	Leu	Tyr	Leu	Ser	Leu	Glu	Ile	Ser	Pro	Gln	Asn	Val	Asp	290	295	300
Val	Asn	Val	His	Pro	Thr	Lys	His	Glu	Val	His	Phe	Leu	His	Glu	Glu	305	310	315
Ser	Ile	Leu	Glu	Arg	Val	Gln	Gln	His	Ile	Glu	Ser	Lys	Leu	Leu	Gly	325	330	335
Ser	Asn	Ser	Ser	Arg	Met	Tyr	Phe	Thr	Gln	Thr	Leu	Leu	Pro	Gly	Leu	340	345	350
Ala	Gly	Pro	Ser	Gly	Glu	Met	Val	Lys	Ser	Thr	Thr	Ser	Leu	Thr	Ser	355	360	365
Ser	Ser	Thr	Ser	Gly	Ser	Ser	Asp	Lys	Val	Tyr	Ala	His	Gln	Met	Val	370	375	380
Arg	Thr	Asp	Ser	Arg	Glu	Gln	Lys	Leu	Asp	Ala	Phe	Leu	Gln	Pro	Leu	385	390	395
Ser	Lys	Pro	Leu	Ser	Ser	Gln	Pro	Gln	Ala	Ile	Val	Thr	Glu	Asp	Lys	405	410	415
Thr	Asp	Ile	Ser	Ser	Gly	Arg	Ala	Arg	Gln	Gln	Asp	Glu	Glu	Met	Leu	420	425	430
Glu	Leu	Pro	Ala	Pro	Ala	Glu	Val	Ala	Ala	Lys	Asn	Gln	Ser	Leu	Glu	435	440	445
Gly	Asp	Thr	Thr	Lys	Gly	Thr	Ser	Glu	Met	Ser	Glu	Lys	Arg	Gly	Pro	450	455	460
Thr	Ser	Ser	Asn	Pro	Arg	Lys	Arg	His	Arg	Glu	Asp	Ser	Asp	Val	Glu	465	470	475
Met	Val	Glu	Asp	Asp	Ser	Arg	Lys	Glu	Met	Thr	Ala	Ala	Cys	Thr	Pro	485	490	495
Arg	Arg	Arg	Ile	Ile	Asn	Leu	Thr	Ser	Val	Leu	Ser	Leu	Gln	Glu	Glu	500	505	510
Ile	Asn	Glu	Gln	Gly	His	Glu	Val	Leu	Arg	Glu	Met	Leu	His	Asn	His	515	520	525
Ser	Phe	Val	Gly	Cys	Val	Asn	Pro	Gln	Trp	Ala	Leu	Ala	Gln	His	Gln	530	535	540
Thr	Lys	Leu	Tyr	Leu	Leu	Asn	Thr	Thr	Lys	Leu	Ser	Glu	Glu	Leu	Phe	545	550	555
Tyr	Gln	Ile	Leu	Ile	Tyr	Asp	Phe	Ala	Asn	Phe	Gly	Val	Leu	Arg	Leu	565	570	575
Ser	Glu	Pro	Ala	Pro	Leu	Phe	Asp	Leu	Ala	Met	Leu	Ala	Leu	Asp	Ser	580	585	590
Pro	Glu	Ser	Gly	Trp	Thr	Glu	Glu	Asp	Gly	Pro	Lys	Glu	Gly	Leu	Ala	595	600	605
Glu	Tyr	Ile	Val	Glu	Phe	Leu	Lys	Lys	Lys	Ala	Glu	Met	Leu	Ala	Asp	610	615	620
Tyr	Phe	Ser	Leu	Glu	Ile	Asp	Glu	Glu	Gly	Asn	Leu	Ile	Gly	Leu	Pro	625	630	635
Leu	Leu	Ile	Asp	Asn	Tyr	Val	Pro	Pro	Leu	Glu	Gly	Leu	Pro	Ile	Phe	645	650	655
Ile	Leu	Arg	Leu	Ala	Thr	Glu	Val	Asn	Trp	Asp	Glu	Glu	Lys	Glu	Cys	660	665	670
Phe	Glu	Ser	Leu	Ser	Lys	Glu	Cys	Ala	Met	Phe	Tyr	Ser	Ile	Arg	Lys	675	680	685
Gln	Tyr	Ile	Ser	Glu	Glu	Ser	Thr	Leu	Ser	Gly	Gln	Gln	Ser	Glu	Val	690	695	700

Pro	Gly	Ser	Ile	Pro	Asn	Ser	Trp	Lys	Trp	Thr	Val	Glu	His	Ile	Val
705					710					715					720
Tyr	Lys	Ala	Leu	Arg	Ser	His	Ile	Leu	Pro	Pro	Lys	His	Phe	Thr	Glu
				725					730					735	
Asp	Gly	Asn	Ile	Leu	Gln	Leu	Ala	Asn	Leu	Pro	Asp	Leu	Tyr	Lys	Val
			740					745					750		
Phe	Glu	Arg	Cys												
		755													

<210> 166
 <211> 1453
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >MLH3|ENSG00000119684|ENST00000355774|4362

<400> 166

Met	Ile	Lys	Cys	Leu	Ser	Val	Glu	Val	Gln	Ala	Lys	Leu	Arg	Ser	Gly
1				5					10					15	
Leu	Ala	Ile	Ser	Ser	Leu	Gly	Gln	Cys	Val	Glu	Glu	Leu	Ala	Leu	Asn
			20					25					30		
Ser	Ile	Asp	Ala	Glu	Ala	Lys	Cys	Val	Ala	Val	Arg	Val	Asn	Met	Glu
		35					40				45				
Thr	Phe	Gln	Val	Gln	Val	Ile	Asp	Asn	Gly	Phe	Gly	Met	Gly	Ser	Asp
	50					55					60				
Asp	Val	Glu	Lys	Val	Gly	Asn	Arg	Tyr	Phe	Thr	Ser	Lys	Cys	His	Ser
65					70					75				80	
Val	Gln	Asp	Leu	Glu	Asn	Pro	Arg	Phe	Tyr	Gly	Phe	Arg	Gly	Glu	Ala
			85						90					95	
Leu	Ala	Asn	Ile	Ala	Asp	Met	Ala	Ser	Ala	Val	Glu	Ile	Ser	Ser	Lys
			100					105					110		
Lys	Asn	Arg	Thr	Met	Lys	Thr	Phe	Val	Lys	Leu	Phe	Gln	Ser	Gly	Lys
		115					120					125			
Ala	Leu	Lys	Ala	Cys	Glu	Ala	Asp	Val	Thr	Arg	Ala	Ser	Ala	Gly	Thr
	130					135					140				
Thr	Val	Thr	Val	Tyr	Asn	Leu	Phe	Tyr	Gln	Leu	Pro	Val	Arg	Arg	Lys
145					150					155					160
Cys	Met	Asp	Pro	Arg	Leu	Glu	Phe	Glu	Lys	Val	Arg	Gln	Arg	Ile	Glu
				165					170					175	
Ala	Leu	Ser	Leu	Met	His	Pro	Ser	Ile	Ser	Phe	Ser	Leu	Arg	Asn	Asp
			180					185					190		
Val	Ser	Gly	Ser	Met	Val	Leu	Gln	Leu	Pro	Lys	Thr	Lys	Asp	Val	Cys
		195					200					205			
Ser	Arg	Phe	Cys	Gln	Ile	Tyr	Gly	Leu	Gly	Lys	Ser	Gln	Lys	Leu	Arg
	210					215					220				
Glu	Ile	Ser	Phe	Lys	Tyr	Lys	Glu	Phe	Glu	Leu	Ser	Gly	Tyr	Ile	Ser
225					230					235					240
Ser	Glu	Ala	His	Tyr	Asn	Lys	Asn	Met	Gln	Phe	Leu	Phe	Val	Asn	Lys
				245					250					255	
Arg	Leu	Val	Leu	Arg	Thr	Lys	Leu	His	Lys	Leu	Ile	Asp	Phe	Leu	Leu
		260						265					270		
Arg	Lys	Glu	Ser	Ile	Ile	Cys	Lys	Pro	Lys	Asn	Gly	Pro	Thr	Ser	Arg

705					710					715					720
Arg	His	Val	Ser	Asn	Asp	Ser	Arg	Lys	Thr	Asp	Lys	Leu	Ile	Gly	Phe
				725					730					735	
Ser	Lys	Pro	Ile	Val	Arg	Lys	Lys	Leu	Ser	Leu	Ser	Ser	Gln	Leu	Gly
			740					745					750		
Ser	Leu	Glu	Lys	Phe	Lys	Arg	Gln	Tyr	Gly	Lys	Val	Glu	Asn	Pro	Leu
		755					760					765			
Asp	Thr	Glu	Val	Glu	Glu	Ser	Asn	Gly	Val	Thr	Thr	Asn	Leu	Ser	Leu
	770					775					780				
Gln	Val	Glu	Pro	Asp	Ile	Leu	Leu	Lys	Asp	Lys	Asn	Arg	Leu	Glu	Asn
785				790					795						800
Ser	Asp	Val	Cys	Lys	Ile	Thr	Thr	Met	Glu	His	Ser	Asp	Ser	Asp	Ser
			805						810					815	
Ser	Cys	Gln	Pro	Ala	Ser	His	Ile	Leu	Asn	Ser	Glu	Lys	Phe	Pro	Phe
			820					825					830		
Ser	Lys	Asp	Glu	Asp	Cys	Leu	Glu	Gln	Gln	Met	Pro	Ser	Leu	Arg	Glu
		835					840					845			
Ser	Pro	Met	Thr	Leu	Lys	Glu	Leu	Ser	Leu	Phe	Asn	Arg	Lys	Pro	Leu
	850					855					860				
Asp	Leu	Glu	Lys	Ser	Ser	Glu	Ser	Leu	Ala	Ser	Lys	Leu	Ser	Arg	Leu
865					870					875					880
Lys	Gly	Ser	Glu	Arg	Glu	Thr	Gln	Thr	Met	Gly	Met	Met	Ser	Arg	Phe
				885					890						895
Asn	Glu	Leu	Pro	Asn	Ser	Asp	Ser	Ser	Arg	Lys	Asp	Ser	Lys	Leu	Cys
			900					905					910		
Ser	Val	Leu	Thr	Gln	Asp	Phe	Cys	Met	Leu	Phe	Asn	Asn	Lys	His	Glu
	915					920						925			
Lys	Thr	Glu	Asn	Gly	Val	Ile	Pro	Thr	Ser	Asp	Ser	Ala	Thr	Gln	Asp
	930					935					940				
Asn	Ser	Phe	Asn	Lys	Asn	Ser	Lys	Thr	His	Ser	Asn	Ser	Asn	Thr	Thr
945					950					955					960
Glu	Asn	Cys	Val	Ile	Ser	Glu	Thr	Pro	Leu	Val	Leu	Pro	Tyr	Asn	Asn
			965						970					975	
Ser	Lys	Val	Thr	Gly	Lys	Asp	Ser	Asp	Val	Leu	Ile	Arg	Ala	Ser	Glu
			980					985					990		
Gln	Gln	Ile	Gly	Ser	Leu	Asp	Ser	Pro	Ser	Gly	Met	Leu	Met	Asn	Pro
	995						1000					1005			
Val	Glu	Asp	Ala	Thr	Gly	Asp	Gln	Asn	Gly	Ile	Cys	Phe	Gln	Ser	Glu
	1010					1015					1020				
Glu	Ser	Lys	Ala	Arg	Ala	Cys	Ser	Glu	Thr	Glu	Glu	Ser	Asn	Thr	Cys
1025					1030					1035					1040
Cys	Ser	Asp	Trp	Gln	Arg	His	Phe	Asp	Val	Ala	Leu	Gly	Arg	Met	Val
			1045						1050					1055	
Tyr	Val	Asn	Lys	Met	Thr	Gly	Leu	Ser	Thr	Phe	Ile	Ala	Pro	Thr	Glu
		1060						1065					1070		
Asp	Ile	Gln	Ala	Ala	Cys	Thr	Lys	Asp	Leu	Thr	Thr	Val	Ala	Val	Asp
	1075						1080					1085			
Val	Val	Leu	Glu	Asn	Gly	Ser	Gln	Tyr	Arg	Cys	Gln	Pro	Phe	Arg	Ser
	1090					1095					1100				
Asp	Leu	Val	Leu	Pro	Phe	Leu	Pro	Arg	Ala	Arg	Ala	Glu	Arg	Thr	Val
1105					1110					1115					1120
Met	Arg	Gln	Asp	Asn	Arg	Asp	Thr	Val	Asp	Asp	Thr	Val	Ser	Ser	Glu
			1125						1130					1135	
Ser	Leu	Gln	Ser	Leu	Phe	Ser	Glu	Trp	Asp	Asn	Pro	Val	Phe	Ala	Arg

Page 494

Glu Glu Lys Asp Ala Ile Glu Glu Leu Val Lys Tyr Gln Leu Glu Lys
 465 470 475 480
 Thr Gln Arg Phe Leu Lys Glu Arg His Ile Asp Ala Leu Glu Asp Lys
 485 490 495
 Ile Asp Glu Glu Val Arg Arg Phe Arg Glu Thr Arg Gln Lys Asn Thr
 500 505 510
 Asn Glu Glu Asp Asp Glu Val Arg Glu Ala Met Thr Arg Ala Arg Ala
 515 520 525
 Leu Arg Ser Gln Ser Glu Glu Ser Ala Ser Ala Phe Ser Ala Asp Asp
 530 535 540
 Leu Met Ser Ile Asp Leu Ala Glu Gln Met Ala Asn Asp Ser Asp Asp
 545 550 555 560
 Ser Ile Ser Ala Ala Thr Asn Lys Gly Arg Gly Arg Gly Arg Gly Arg
 565 570 575
 Arg Gly Gly Arg Gly Gln Asn Ser Ala Ser Arg Gly Gly Ser Gln Arg
 580 585 590
 Gly Arg Ala Asp Thr Gly Leu Glu Thr Ser Thr Arg Ser Arg Asn Ser
 595 600 605
 Lys Thr Ala Val Ser Ala Ser Arg Asn Met Ser Ile Ile Asp Ala Phe
 610 615 620
 Lys Ser Thr Arg Gln Gln Pro Ser Arg Asn Val Thr Thr Lys Asn Tyr
 625 630 635 640
 Ser Glu Val Ile Glu Val Asp Glu Ser Asp Val Glu Glu Asp Ile Phe
 645 650 655
 Pro Thr Thr Ser Lys Thr Asp Gln Arg Trp Ser Ser Thr Ser Ser Ser
 660 665 670
 Lys Ile Met Ser Gln Ser Gln Val Ser Lys Gly Val Asp Phe Glu Ser
 675 680 685
 Ser Glu Asp Asp Asp Asp Asp Pro Phe Met Asn Thr Ser Ser Leu Arg
 690 695 700
 Arg Asn Arg Arg
 705

<210> 168

<211> 934

<212> PRT

<213> Homo sapiens

<220>

<223> >MSH2|ENSG00000095002|ENST00000233146|2805

<400> 168

Met Ala Val Gln Pro Lys Glu Thr Leu Gln Leu Glu Ser Ala Ala Glu
 1 5 10 15
 Val Gly Phe Val Arg Phe Phe Gln Gly Met Pro Glu Lys Pro Thr Thr
 20 25 30
 Thr Val Arg Leu Phe Asp Arg Gly Asp Phe Tyr Thr Ala His Gly Glu
 35 40 45
 Asp Ala Leu Leu Ala Ala Arg Glu Val Phe Lys Thr Gln Gly Val Ile
 50 55 60
 Lys Tyr Met Gly Pro Ala Gly Ala Lys Asn Leu Gln Ser Val Val Leu
 65 70 75 80
 Ser Lys Met Asn Phe Glu Ser Phe Val Lys Asp Leu Leu Leu Val Arg

				85				90					95				
Gln	Tyr	Arg	Val	Glu	Val	Tyr	Lys	Asn	Arg	Ala	Gly	Asn	Lys	Ala	Ser		
			100					105					110				
Lys	Glu	Asn	Asp	Trp	Tyr	Leu	Ala	Tyr	Lys	Ala	Ser	Pro	Gly	Asn	Leu		
		115					120					125					
Ser	Gln	Phe	Glu	Asp	Ile	Leu	Phe	Gly	Asn	Asn	Asp	Met	Ser	Ala	Ser		
	130					135					140						
Ile	Gly	Val	Val	Gly	Val	Lys	Met	Ser	Ala	Val	Asp	Gly	Gln	Arg	Gln		
145				150						155					160		
Val	Gly	Val	Gly	Tyr	Val	Asp	Ser	Ile	Gln	Arg	Lys	Leu	Gly	Leu	Cys		
			165					170						175			
Glu	Phe	Pro	Asp	Asn	Asp	Gln	Phe	Ser	Asn	Leu	Glu	Ala	Leu	Leu	Ile		
			180					185					190				
Gln	Ile	Gly	Pro	Lys	Glu	Cys	Val	Leu	Pro	Gly	Gly	Glu	Thr	Ala	Gly		
		195				200						205					
Asp	Met	Gly	Lys	Leu	Arg	Gln	Ile	Ile	Gln	Arg	Gly	Gly	Ile	Leu	Ile		
	210					215					220						
Thr	Glu	Arg	Lys	Lys	Ala	Asp	Phe	Ser	Thr	Lys	Asp	Ile	Tyr	Gln	Asp		
225					230					235					240		
Leu	Asn	Arg	Leu	Leu	Lys	Gly	Lys	Lys	Gly	Glu	Gln	Met	Asn	Ser	Ala		
			245						250						255		
Val	Leu	Pro	Glu	Met	Glu	Asn	Gln	Val	Ala	Val	Ser	Ser	Leu	Ser	Ala		
		260						265					270				
Val	Ile	Lys	Phe	Leu	Glu	Leu	Leu	Ser	Asp	Asp	Ser	Asn	Phe	Gly	Gln		
	275					280						285					
Phe	Glu	Leu	Thr	Thr	Phe	Asp	Phe	Ser	Gln	Tyr	Met	Lys	Leu	Asp	Ile		
	290					295					300						
Ala	Ala	Val	Arg	Ala	Leu	Asn	Leu	Phe	Gln	Gly	Ser	Val	Glu	Asp	Thr		
305					310					315					320		
Thr	Gly	Ser	Gln	Ser	Leu	Ala	Ala	Leu	Leu	Asn	Lys	Cys	Lys	Thr	Pro		
			325						330						335		
Gln	Gly	Gln	Arg	Leu	Val	Asn	Gln	Trp	Ile	Lys	Gln	Pro	Leu	Met	Asp		
		340						345					350				
Lys	Asn	Arg	Ile	Glu	Glu	Arg	Leu	Asn	Leu	Val	Glu	Ala	Phe	Val	Glu		
	355						360					365					
Asp	Ala	Glu	Leu	Arg	Gln	Thr	Leu	Gln	Glu	Asp	Leu	Leu	Arg	Arg	Phe		
	370				375						380						
Pro	Asp	Leu	Asn	Arg	Leu	Ala	Lys	Lys	Phe	Gln	Arg	Gln	Ala	Ala	Asn		
385					390					395					400		
Leu	Gln	Asp	Cys	Tyr	Arg	Leu	Tyr	Gln	Gly	Ile	Asn	Gln	Leu	Pro	Asn		
		405						410						415			
Val	Ile	Gln	Ala	Leu	Glu	Lys	His	Glu	Gly	Lys	His	Gln	Lys	Leu	Leu		
		420						425					430				
Leu	Ala	Val	Phe	Val	Thr	Pro	Leu	Thr	Asp	Leu	Arg	Ser	Asp	Phe	Ser		
	435					440						445					
Lys	Phe	Gln	Glu	Met	Ile	Glu	Thr	Thr	Leu	Asp	Met	Asp	Gln	Val	Glu		
	450					455					460						
Asn	His	Glu	Phe	Leu	Val	Lys	Pro	Ser	Phe	Asp	Pro	Asn	Leu	Ser	Glu		
465					470					475					480		
Leu	Arg	Glu	Ile	Met	Asn	Asp	Leu	Glu	Lys	Lys	Met	Gln	Ser	Thr	Leu		
			485					490							495		
Ile	Ser	Ala	Ala	Arg	Asp	Leu	Gly	Leu	Asp	Pro	Gly	Lys	Gln	Ile	Lys		
		500					505					510					
Leu	Asp	Ser	Ser	Ala	Gln	Phe	Gly	Tyr	Tyr	Phe	Arg	Val	Thr	Cys	Lys		

<210> 169
 <211> 1137
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >MSH3|ENSG00000113318|ENST00000265081|3414

<400> 169
 Met Ser Arg Arg Lys Pro Ala Ser Gly Gly Leu Ala Ala Ser Ser Ser
 1 5 10 15
 Ala Pro Ala Arg Gln Ala Val Leu Ser Arg Phe Phe Gln Ser Thr Gly
 20 25 30
 Ser Leu Lys Ser Thr Ser Ser Ser Thr Gly Ala Ala Asp Gln Val Asp
 35 40 45
 Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Pro Pro
 50 55 60
 Ala Pro Pro Ala Pro Ala Phe Pro Pro Gln Leu Pro Pro His Ile Ala
 65 70 75 80
 Thr Glu Ile Asp Arg Arg Lys Lys Arg Pro Leu Glu Asn Asp Gly Pro
 85 90 95
 Val Lys Lys Lys Val Lys Lys Val Gln Lys Glu Gly Gly Ser Asp
 100 105 110
 Leu Gly Met Ser Gly Asn Ser Glu Pro Lys Lys Cys Leu Arg Thr Arg
 115 120 125
 Asn Val Ser Lys Ser Leu Glu Lys Leu Lys Glu Phe Cys Cys Asp Ser
 130 135 140
 Ala Leu Pro Gln Ser Arg Val Gln Thr Glu Ser Leu Gln Glu Arg Phe
 145 150 155 160
 Ala Val Leu Pro Lys Cys Thr Asp Phe Asp Asp Ile Ser Leu Leu His
 165 170 175
 Ala Lys Asn Ala Val Ser Ser Glu Asp Ser Lys Arg Gln Ile Asn Gln
 180 185 190
 Lys Asp Thr Thr Leu Phe Asp Leu Ser Gln Phe Gly Ser Ser Asn Thr
 195 200 205
 Ser His Glu Asn Leu Gln Lys Thr Ala Ser Lys Ser Ala Asn Lys Arg
 210 215 220
 Ser Lys Ser Ile Tyr Thr Pro Leu Glu Leu Gln Tyr Ile Glu Met Lys
 225 230 235 240
 Gln Gln His Lys Asp Ala Val Leu Cys Val Glu Cys Gly Tyr Lys Tyr
 245 250 255
 Arg Phe Phe Gly Glu Asp Ala Glu Ile Ala Ala Arg Glu Leu Asn Ile
 260 265 270
 Tyr Cys His Leu Asp His Asn Phe Met Thr Ala Ser Ile Pro Thr His
 275 280 285
 Arg Leu Phe Val His Val Arg Arg Leu Val Ala Lys Gly Tyr Lys Val
 290 295 300
 Gly Val Val Lys Gln Thr Glu Thr Ala Ala Leu Lys Ala Ile Gly Asp
 305 310 315 320
 Asn Arg Ser Ser Leu Phe Ser Arg Lys Leu Thr Ala Leu Tyr Thr Lys
 325 330 335
 Ser Thr Leu Ile Gly Glu Asp Val Asn Pro Leu Ile Lys Leu Asp Asp
 340 345 350

Ala	Val	Asn	Val	Asp	Glu	Ile	Met	Thr	Asp	Thr	Ser	Thr	Ser	Tyr	Leu
		355					360					365			
Leu	Cys	Ile	Ser	Glu	Asn	Lys	Glu	Asn	Val	Arg	Asp	Lys	Lys	Lys	Gly
	370					375					380				
Asn	Ile	Phe	Ile	Gly	Ile	Val	Gly	Val	Gln	Pro	Ala	Thr	Gly	Glu	Val
385					390					395					400
Val	Phe	Asp	Ser	Phe	Gln	Asp	Ser	Ala	Ser	Arg	Ser	Glu	Leu	Glu	Thr
				405					410					415	
Arg	Met	Ser	Ser	Leu	Gln	Pro	Val	Glu	Leu	Leu	Leu	Pro	Ser	Ala	Leu
			420					425					430		
Ser	Glu	Gln	Thr	Glu	Ala	Leu	Ile	His	Arg	Ala	Thr	Ser	Val	Ser	Val
		435					440					445			
Gln	Asp	Asp	Arg	Ile	Arg	Val	Glu	Arg	Met	Asp	Asn	Ile	Tyr	Phe	Glu
	450					455				460					
Tyr	Ser	His	Ala	Phe	Gln	Ala	Val	Thr	Glu	Phe	Tyr	Ala	Lys	Asp	Thr
465					470					475					480
Val	Asp	Ile	Lys	Gly	Ser	Gln	Ile	Ile	Ser	Gly	Ile	Val	Asn	Leu	Glu
				485					490					495	
Lys	Pro	Val	Ile	Cys	Ser	Leu	Ala	Ala	Ile	Ile	Lys	Tyr	Leu	Lys	Glu
			500					505					510		
Phe	Asn	Leu	Glu	Lys	Met	Leu	Ser	Lys	Pro	Glu	Asn	Phe	Lys	Gln	Leu
		515					520					525			
Ser	Ser	Lys	Met	Glu	Phe	Met	Thr	Ile	Asn	Gly	Thr	Thr	Leu	Arg	Asn
	530					535					540				
Leu	Glu	Ile	Leu	Gln	Asn	Gln	Thr	Asp	Met	Lys	Thr	Lys	Gly	Ser	Leu
545					550					555					560
Leu	Trp	Val	Leu	Asp	His	Thr	Lys	Thr	Ser	Phe	Gly	Arg	Arg	Lys	Leu
				565					570					575	
Lys	Lys	Trp	Val	Thr	Gln	Pro	Leu	Leu	Lys	Leu	Arg	Glu	Ile	Asn	Ala
			580					585					590		
Arg	Leu	Asp	Ala	Val	Ser	Glu	Val	Leu	His	Ser	Glu	Ser	Ser	Val	Phe
		595					600					605			
Gly	Gln	Ile	Glu	Asn	His	Leu	Arg	Lys	Leu	Pro	Asp	Ile	Glu	Arg	Gly
	610					615					620				
Leu	Cys	Ser	Ile	Tyr	His	Lys	Lys	Cys	Ser	Thr	Gln	Glu	Phe	Phe	Leu
625					630					635					640
Ile	Val	Lys	Thr	Leu	Tyr	His	Leu	Lys	Ser	Glu	Phe	Gln	Ala	Ile	Ile
				645					650					655	
Pro	Ala	Val	Asn	Ser	His	Ile	Gln	Ser	Asp	Leu	Leu	Arg	Thr	Val	Ile
			660					665					670		
Leu	Glu	Ile	Pro	Glu	Leu	Leu	Ser	Pro	Val	Glu	His	Tyr	Leu	Lys	Ile
		675					680					685			
Leu	Asn	Glu	Gln	Ala	Ala	Lys	Val	Gly	Asp	Lys	Thr	Glu	Leu	Phe	Lys
	690					695					700				
Asp	Leu	Ser	Asp	Phe	Pro	Leu	Ile	Lys	Lys	Arg	Lys	Asp	Glu	Ile	Gln
705					710					715					720
Gly	Val	Ile	Asp	Glu	Ile	Arg	Met	His	Leu	Gln	Glu	Ile	Arg	Lys	Ile
				725					730					735	
Leu	Lys	Asn	Pro	Ser	Ala	Gln	Tyr	Val	Thr	Val	Ser	Gly	Gln	Glu	Phe
			740					745					750		
Met	Ile	Glu	Ile	Lys	Asn	Ser	Ala	Val	Ser	Cys	Ile	Pro	Thr	Asp	Trp
		755					760					765			
Val	Lys	Val	Gly	Ser	Thr	Lys	Ala	Val	Ser	Arg	Phe	His	Ser	Pro	Phe
	770					775					780				

Ile	Val	Glu	Asn	Tyr	Arg	His	Leu	Asn	Gln	Leu	Arg	Glu	Gln	Leu	Val	785	790	795	800
Leu	Asp	Cys	Ser	Ala	Glu	Trp	Leu	Asp	Phe	Leu	Glu	Lys	Phe	Ser	Glu		805	810	815
His	Tyr	His	Ser	Leu	Cys	Lys	Ala	Val	His	His	Leu	Ala	Thr	Val	Asp		820	825	830
Cys	Ile	Phe	Ser	Leu	Ala	Lys	Val	Ala	Lys	Gln	Gly	Asp	Tyr	Cys	Arg		835	840	845
Pro	Thr	Val	Gln	Glu	Glu	Arg	Lys	Ile	Val	Ile	Lys	Asn	Gly	Arg	His		850	855	860
Pro	Val	Ile	Asp	Val	Leu	Leu	Gly	Glu	Gln	Asp	Gln	Tyr	Val	Pro	Asn		865	870	875
Asn	Thr	Asp	Leu	Ser	Glu	Asp	Ser	Glu	Arg	Val	Met	Ile	Ile	Thr	Gly		885	890	895
Pro	Asn	Met	Gly	Gly	Lys	Ser	Ser	Tyr	Ile	Lys	Gln	Val	Ala	Leu	Ile		900	905	910
Thr	Ile	Met	Ala	Gln	Ile	Gly	Ser	Tyr	Val	Pro	Ala	Glu	Glu	Ala	Thr		915	920	925
Ile	Gly	Ile	Val	Asp	Gly	Ile	Phe	Thr	Arg	Met	Gly	Ala	Ala	Asp	Asn		930	935	940
Ile	Tyr	Lys	Gly	Gln	Ser	Thr	Phe	Met	Glu	Glu	Leu	Thr	Asp	Thr	Ala		945	950	955
Glu	Ile	Ile	Arg	Lys	Ala	Thr	Ser	Gln	Ser	Leu	Val	Ile	Leu	Asp	Glu		965	970	975
Leu	Gly	Arg	Gly	Thr	Ser	Thr	His	Asp	Gly	Ile	Ala	Ile	Ala	Tyr	Ala		980	985	990
Thr	Leu	Glu	Tyr	Phe	Ile	Arg	Asp	Val	Lys	Ser	Leu	Thr	Leu	Phe	Val		995	1000	1005
Thr	His	Tyr	Pro	Pro	Val	Cys	Glu	Leu	Glu	Lys	Asn	Tyr	Ser	His	Gln		1010	1015	1020
Val	Gly	Asn	Tyr	His	Met	Gly	Phe	Leu	Val	Ser	Glu	Asp	Glu	Ser	Lys		1025	1030	1035
Leu	Asp	Pro	Gly	Ala	Ala	Glu	Gln	Val	Pro	Asp	Phe	Val	Thr	Phe	Leu		1045	1050	1055
Tyr	Gln	Ile	Thr	Arg	Gly	Ile	Ala	Ala	Arg	Ser	Tyr	Gly	Leu	Asn	Val		1060	1065	1070
Ala	Lys	Leu	Ala	Asp	Val	Pro	Gly	Glu	Ile	Leu	Lys	Lys	Ala	Ala	His		1075	1080	1085
Lys	Ser	Lys	Glu	Leu	Glu	Gly	Leu	Ile	Asn	Thr	Lys	Arg	Lys	Arg	Leu		1090	1095	1100
Lys	Tyr	Phe	Ala	Lys	Leu	Trp	Thr	Met	His	Asn	Ala	Gln	Asp	Leu	Gln		1105	1110	1115
Lys	Trp	Thr	Glu	Glu	Phe	Asn	Met	Glu	Glu	Thr	Gln	Thr	Ser	Leu	Leu		1125	1130	1135

His

<210> 170

<211> 1360

<212> PRT

<213> Homo sapiens

<220>

<223> >MSH6|ENSG00000116062|ENST00000234420|4083

<400> 170

Met	Ser	Arg	Gln	Ser	Thr	Leu	Tyr	Ser	Phe	Phe	Pro	Lys	Ser	Pro	Ala	1	5	10	15
Leu	Ser	Asp	Ala	Asn	Lys	Ala	Ser	Ala	Arg	Ala	Ser	Arg	Glu	Gly	Gly	20	25	30	
Arg	Ala	Ala	Ala	Ala	Pro	Gly	Ala	Ser	Pro	Ser	Pro	Gly	Gly	Asp	Ala	35	40	45	
Ala	Trp	Ser	Glu	Ala	Gly	Pro	Gly	Pro	Arg	Pro	Leu	Ala	Arg	Ser	Ala	50	55	60	
Ser	Pro	Pro	Lys	Ala	Lys	Asn	Leu	Asn	Gly	Gly	Leu	Arg	Arg	Ser	Val	65	70	75	80
Ala	Pro	Ala	Ala	Pro	Thr	Ser	Cys	Asp	Phe	Ser	Pro	Gly	Asp	Leu	Val	85	90	95	
Trp	Ala	Lys	Met	Glu	Gly	Tyr	Pro	Trp	Trp	Pro	Cys	Leu	Val	Tyr	Asn	100	105	110	
His	Pro	Phe	Asp	Gly	Thr	Phe	Ile	Arg	Glu	Lys	Gly	Lys	Ser	Val	Arg	115	120	125	
Val	His	Val	Gln	Phe	Phe	Asp	Asp	Ser	Pro	Thr	Arg	Gly	Trp	Val	Ser	130	135	140	
Lys	Arg	Leu	Leu	Lys	Pro	Tyr	Thr	Gly	Ser	Lys	Ser	Lys	Glu	Ala	Gln	145	150	155	160
Lys	Gly	Gly	His	Phe	Tyr	Ser	Ala	Lys	Pro	Glu	Ile	Leu	Arg	Ala	Met	165	170	175	
Gln	Arg	Ala	Asp	Glu	Ala	Leu	Asn	Lys	Asp	Lys	Ile	Lys	Arg	Leu	Glu	180	185	190	
Leu	Ala	Val	Cys	Asp	Glu	Pro	Ser	Glu	Pro	Glu	Glu	Glu	Glu	Glu	Met	195	200	205	
Glu	Val	Gly	Thr	Thr	Tyr	Val	Thr	Asp	Lys	Ser	Glu	Glu	Asp	Asn	Glu	210	215	220	
Ile	Glu	Ser	Glu	Glu	Glu	Val	Gln	Pro	Lys	Thr	Gln	Gly	Ser	Arg	Arg	225	230	235	240
Ser	Ser	Arg	Gln	Ile	Lys	Lys	Arg	Arg	Val	Ile	Ser	Asp	Ser	Glu	Ser	245	250	255	
Asp	Ile	Gly	Gly	Ser	Asp	Val	Glu	Phe	Lys	Pro	Asp	Thr	Lys	Glu	Glu	260	265	270	
Gly	Ser	Ser	Asp	Glu	Ile	Ser	Ser	Gly	Val	Gly	Asp	Ser	Glu	Ser	Glu	275	280	285	
Gly	Leu	Asn	Ser	Pro	Val	Lys	Val	Ala	Arg	Lys	Arg	Lys	Arg	Met	Val	290	295	300	
Thr	Gly	Asn	Gly	Ser	Leu	Lys	Arg	Lys	Ser	Ser	Arg	Lys	Glu	Thr	Pro	305	310	315	320
Ser	Ala	Thr	Lys	Gln	Ala	Thr	Ser	Ile	Ser	Ser	Glu	Thr	Lys	Asn	Thr	325	330	335	
Leu	Arg	Ala	Phe	Ser	Ala	Pro	Gln	Asn	Ser	Glu	Ser	Gln	Ala	His	Val	340	345	350	
Ser	Gly	Gly	Gly	Asp	Asp	Ser	Ser	Arg	Pro	Thr	Val	Trp	Tyr	His	Glu	355	360	365	
Thr	Leu	Glu	Trp	Leu	Lys	Glu	Glu	Lys	Arg	Arg	Asp	Glu	His	Arg	Arg	370	375	380	
Arg	Pro	Asp	His	Pro	Asp	Phe	Asp	Ala	Ser	Thr	Leu	Tyr	Val	Pro	Glu	385	390	395	400
Asp	Phe	Leu	Asn	Ser	Cys	Thr	Pro	Gly	Met	Arg	Lys	Trp	Trp	Gln	Ile				

Page 503

1265		1270		1275		1280
Glu Thr Ile Thr Phe Leu Tyr Lys Phe Ile Lys Gly Ala Cys Pro Lys						
	1285		1290		1295	
Ser Tyr Gly Phe Asn Ala Ala Arg Leu Ala Asn Leu Pro Glu Glu Val						
	1300		1305		1310	
Ile Gln Lys Gly His Arg Lys Ala Arg Glu Phe Glu Lys Met Asn Gln						
	1315		1320		1325	
Ser Leu Arg Leu Phe Arg Glu Val Cys Leu Ala Ser Glu Arg Ser Thr						
	1330		1335		1340	
Val Asp Ala Glu Ala Val His Lys Leu Leu Thr Leu Ile Lys Glu Leu						
1345		1350		1355		1360

<210> 171
 <211> 454
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >MYC|ENSG00000136997|ENST00000377970|1365

<400> 171

Leu Asp Phe Phe Arg Val Val Glu Asn Gln Gln Pro Pro Ala Thr Met															
1		5		10		15									
Pro Leu Asn Val Ser Phe Thr Asn Arg Asn Tyr Asp Leu Asp Tyr Asp															
	20		25		30										
Ser Val Gln Pro Tyr Phe Tyr Cys Asp Glu Glu Glu Asn Phe Tyr Gln															
	35		40		45										
Gln Gln Gln Gln Ser Glu Leu Gln Pro Pro Ala Pro Ser Glu Asp Ile															
	50		55		60										
Trp Lys Lys Phe Glu Leu Leu Pro Thr Pro Pro Leu Ser Pro Ser Arg															
65		70		75		80									
Arg Ser Gly Leu Cys Ser Pro Ser Tyr Val Ala Val Thr Pro Phe Ser															
	85		90		95										
Leu Arg Gly Asp Asn Asp Gly Gly Gly Ser Phe Ser Thr Ala Asp															
	100		105		110										
Gln Leu Glu Met Val Thr Glu Leu Leu Gly Gly Asp Met Val Asn Gln															
	115		120		125										
Ser Phe Ile Cys Asp Pro Asp Asp Glu Thr Phe Ile Lys Asn Ile Ile															
	130		135		140										
Ile Gln Asp Cys Met Trp Ser Gly Phe Ser Ala Ala Ala Lys Leu Val															
145		150		155		160									
Ser Glu Lys Leu Ala Ser Tyr Gln Ala Ala Arg Lys Asp Ser Gly Ser															
	165		170		175										
Pro Asn Pro Ala Arg Gly His Ser Val Cys Ser Thr Ser Ser Leu Tyr															
	180		185		190										
Leu Gln Asp Leu Ser Ala Ala Ala Ser Glu Cys Ile Asp Pro Ser Val															
	195		200		205										
Val Phe Pro Tyr Pro Leu Asn Asp Ser Ser Ser Pro Lys Ser Cys Ala															
	210		215		220										
Ser Gln Asp Ser Ser Ala Phe Ser Pro Ser Ser Asp Ser Leu Leu Ser															
225		230		235		240									

Ser	Thr	Glu	Ser	Ser	Pro	Gln	Gly	Ser	Pro	Glu	Pro	Leu	Val	Leu	His
				245					250					255	
Glu	Glu	Thr	Pro	Pro	Thr	Thr	Ser	Ser	Asp	Ser	Glu	Glu	Glu	Gln	Glu
			260					265						270	
Asp	Glu	Glu	Glu	Ile	Asp	Val	Val	Ser	Val	Glu	Lys	Arg	Gln	Ala	Pro
		275					280						285		
Gly	Lys	Arg	Ser	Glu	Ser	Gly	Ser	Pro	Ser	Ala	Gly	Gly	His	Ser	Lys
	290					295					300				
Pro	Pro	His	Ser	Pro	Leu	Val	Leu	Lys	Arg	Cys	His	Val	Ser	Thr	His
					310					315					320
Gln	His	Asn	Tyr	Ala	Ala	Pro	Pro	Ser	Thr	Arg	Lys	Asp	Tyr	Pro	Ala
				325					330					335	
Ala	Lys	Arg	Val	Lys	Leu	Asp	Ser	Val	Arg	Val	Leu	Arg	Gln	Ile	Ser
			340					345						350	
Asn	Asn	Arg	Lys	Cys	Thr	Ser	Pro	Arg	Ser	Ser	Asp	Thr	Glu	Glu	Asn
		355					360					365			
Val	Lys	Arg	Arg	Thr	His	Asn	Val	Leu	Glu	Arg	Gln	Arg	Arg	Asn	Glu
	370					375					380				
Leu	Lys	Arg	Ser	Phe	Phe	Ala	Leu	Arg	Asp	Gln	Ile	Pro	Glu	Leu	Glu
				390						395					400
Asn	Asn	Glu	Lys	Ala	Pro	Lys	Val	Val	Ile	Leu	Lys	Lys	Ala	Thr	Ala
				405					410					415	
Tyr	Ile	Leu	Ser	Val	Gln	Ala	Glu	Glu	Gln	Lys	Leu	Ile	Ser	Glu	Glu
			420					425					430		
Asp	Leu	Leu	Arg	Lys	Arg	Arg	Glu	Gln	Leu	Lys	His	Lys	Leu	Glu	Gln
		435					440					445			
Leu	Arg	Asn	Ser	Cys	Ala										
															450

<210> 172

<211> 754

<212> PRT

<213> Homo sapiens

<220>

<223> >NBN|ENSG00000104320|ENST00000265433|2265

<400> 172

Met	Trp	Lys	Leu	Leu	Pro	Ala	Ala	Gly	Pro	Ala	Gly	Gly	Glu	Pro	Tyr
1				5					10					15	
Arg	Leu	Leu	Thr	Gly	Val	Glu	Tyr	Val	Val	Gly	Arg	Lys	Asn	Cys	Ala
			20					25					30		
Ile	Leu	Ile	Glu	Asn	Asp	Gln	Ser	Ile	Ser	Arg	Asn	His	Ala	Val	Leu
		35					40					45			
Thr	Ala	Asn	Phe	Ser	Val	Thr	Asn	Leu	Ser	Gln	Thr	Asp	Glu	Ile	Pro
	50					55					60				
Val	Leu	Thr	Leu	Lys	Asp	Asn	Ser	Lys	Tyr	Gly	Thr	Phe	Val	Asn	Glu
					70					75					80
Glu	Lys	Met	Gln	Asn	Gly	Phe	Ser	Arg	Thr	Leu	Lys	Ser	Gly	Asp	Gly
				85				90						95	
Ile	Thr	Phe	Gly	Val	Phe	Gly	Ser	Lys	Phe	Arg	Ile	Glu	Tyr	Glu	Pro
			100					105					110		
Leu	Val	Ala	Cys	Ser	Ser	Cys	Leu	Asp	Val	Ser	Gly	Lys	Thr	Ala	Leu

545					550					555				560	
Glu	Val	Leu	Glu	Gln	Leu	Phe	Lys	Asp	Thr	Lys	Pro	Glu	Leu	Glu	Ile
				565					570					575	
Asp	Val	Lys	Val	Gln	Lys	Gln	Glu	Glu	Asp	Val	Asn	Val	Arg	Lys	Arg
			580					585					590		
Pro	Arg	Met	Asp	Ile	Glu	Thr	Asn	Asp	Thr	Phe	Ser	Asp	Glu	Ala	Val
		595					600					605			
Pro	Glu	Ser	Ser	Lys	Ile	Ser	Gln	Glu	Asn	Glu	Ile	Gly	Lys	Lys	Arg
	610				615					620					
Glu	Leu	Lys	Glu	Asp	Ser	Leu	Trp	Ser	Ala	Lys	Glu	Ile	Ser	Asn	Asn
625				630					635					640	
Asp	Lys	Leu	Gln	Asp	Asp	Ser	Glu	Met	Leu	Pro	Lys	Lys	Leu	Leu	Leu
			645					650					655		
Thr	Glu	Phe	Arg	Ser	Leu	Val	Ile	Lys	Asn	Ser	Thr	Ser	Arg	Asn	Pro
		660						665					670		
Ser	Gly	Ile	Asn	Asp	Asp	Tyr	Gly	Gln	Leu	Lys	Asn	Phe	Lys	Lys	Phe
	675					680					685				
Lys	Lys	Val	Thr	Tyr	Pro	Gly	Ala	Gly	Lys	Leu	Pro	His	Ile	Ile	Gly
	690				695						700				
Gly	Ser	Asp	Leu	Ile	Ala	His	His	Ala	Arg	Lys	Asn	Thr	Glu	Leu	Glu
705				710					715					720	
Glu	Trp	Leu	Arg	Gln	Glu	Met	Glu	Val	Gln	Asn	Gln	His	Ala	Lys	Glu
			725					730					735		
Glu	Ser	Leu	Ala	Asp	Asp	Leu	Phe	Arg	Tyr	Asn	Pro	Tyr	Leu	Lys	Arg
		740					745						750		

Arg Arg

<210> 173
 <211> 189
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >NRAS|ENSG00000213281|ENST00000369535|570

<400> 173

Met	Thr	Glu	Tyr	Lys	Leu	Val	Val	Val	Gly	Ala	Gly	Gly	Val	Gly	Lys
1				5					10					15	
Ser	Ala	Leu	Thr	Ile	Gln	Leu	Ile	Gln	Asn	His	Phe	Val	Asp	Glu	Tyr
			20					25					30		
Asp	Pro	Thr	Ile	Glu	Asp	Ser	Tyr	Arg	Lys	Gln	Val	Val	Ile	Asp	Gly
		35					40					45			
Glu	Thr	Cys	Leu	Leu	Asp	Ile	Leu	Asp	Thr	Ala	Gly	Gln	Glu	Glu	Tyr
	50				55						60				
Ser	Ala	Met	Arg	Asp	Gln	Tyr	Met	Arg	Thr	Gly	Glu	Gly	Phe	Leu	Cys
65				70					75					80	
Val	Phe	Ala	Ile	Asn	Asn	Ser	Lys	Ser	Phe	Ala	Asp	Ile	Asn	Leu	Tyr
			85					90					95		
Arg	Glu	Gln	Ile	Lys	Arg	Val	Lys	Asp	Ser	Asp	Asp	Val	Pro	Met	Val
		100					105					110			
Leu	Val	Gly	Asn	Lys	Cys	Asp	Leu	Pro	Thr	Arg	Thr	Val	Asp	Thr	Lys
		115					120					125			

Gln	Ala	His	Glu	Leu	Ala	Lys	Ser	Tyr	Gly	Ile	Pro	Phe	Ile	Glu	Thr
130						135					140				
Ser	Ala	Lys	Thr	Arg	Gln	Gly	Val	Glu	Asp	Ala	Phe	Tyr	Thr	Leu	Val
145					150					155					160
Arg	Glu	Ile	Arg	Gln	Tyr	Arg	Met	Lys	Lys	Leu	Asn	Ser	Ser	Asp	Asp
				165					170					175	
Gly	Thr	Gln	Gly	Cys	Met	Gly	Leu	Pro	Cys	Val	Val	Met			
		180						185							

<210> 174
 <211> 1186
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >PALB2|ENSG00000083093|ENST00000261584|3561

<400> 174

Met	Asp	Glu	Pro	Pro	Gly	Lys	Pro	Leu	Ser	Cys	Glu	Glu	Lys	Glu	Lys
1				5					10					15	
Leu	Lys	Glu	Lys	Leu	Ala	Phe	Leu	Lys	Arg	Glu	Tyr	Ser	Lys	Thr	Leu
		20					25						30		
Ala	Arg	Leu	Gln	Arg	Ala	Gln	Arg	Ala	Glu	Lys	Ile	Lys	His	Ser	Ile
	35					40					45				
Lys	Lys	Thr	Val	Glu	Glu	Gln	Asp	Cys	Leu	Ser	Gln	Gln	Asp	Leu	Ser
	50					55					60				
Pro	Gln	Leu	Lys	His	Ser	Glu	Pro	Lys	Asn	Lys	Ile	Cys	Val	Tyr	Asp
65					70					75					80
Lys	Leu	His	Ile	Lys	Thr	His	Leu	Asp	Glu	Glu	Thr	Gly	Glu	Lys	Thr
			85					90						95	
Ser	Ile	Thr	Leu	Asp	Val	Gly	Pro	Glu	Ser	Phe	Asn	Pro	Gly	Asp	Gly
			100					105					110		
Pro	Gly	Gly	Leu	Pro	Ile	Gln	Arg	Thr	Asp	Asp	Thr	Gln	Glu	His	Phe
	115						120					125			
Pro	His	Arg	Val	Ser	Asp	Pro	Ser	Gly	Glu	Gln	Lys	Gln	Lys	Leu	Pro
	130					135					140				
Ser	Arg	Arg	Lys	Lys	Gln	Lys	Arg	Thr	Phe	Ile	Ser	Gln	Glu	Arg	
145					150					155					160
Asp	Cys	Val	Phe	Gly	Thr	Asp	Ser	Leu	Arg	Leu	Ser	Gly	Lys	Arg	Leu
			165					170						175	
Lys	Glu	Gln	Glu	Glu	Ile	Ser	Ser	Lys	Asn	Pro	Ala	Arg	Ser	Pro	Val
		180						185					190		
Thr	Glu	Ile	Arg	Thr	His	Leu	Leu	Ser	Leu	Lys	Ser	Glu	Leu	Pro	Asp
	195						200					205			
Ser	Pro	Glu	Pro	Val	Thr	Glu	Ile	Asn	Glu	Asp	Ser	Val	Leu	Ile	Pro
	210					215					220				
Pro	Thr	Ala	Gln	Pro	Glu	Lys	Gly	Val	Asp	Thr	Phe	Leu	Arg	Arg	Pro
225					230					235					240
Asn	Phe	Thr	Arg	Ala	Thr	Thr	Val	Pro	Leu	Gln	Thr	Leu	Ser	Asp	Ser
			245						250					255	
Gly	Ser	Ser	Gln	His	Leu	Glu	His	Ile	Pro	Pro	Lys	Gly	Ser	Ser	Glu
		260						265				270			
Leu	Thr	Thr	His	Asp	Leu	Lys	Asn	Ile	Arg	Phe	Thr	Ser	Pro	Val	Ser

705					710				715					720
Thr	Asp	Met	Cys	Ser	Pro	Ala	Phe	Pro	Ile	Leu	Gly	Thr	Thr	Pro
				725					730					735
Phe	Gly	Pro	Gln	Gly	Ser	Tyr	Glu	Lys	Ala	Ser	Thr	Glu	Val	Ala
			740					745					750	Gly
Arg	Thr	Cys	Cys	Thr	Pro	Gln	Leu	Ala	His	Leu	Lys	Asp	Ser	Val
		755					760					765		Cys
Leu	Ala	Ser	Asp	Thr	Lys	Gln	Phe	Asp	Ser	Ser	Gly	Ser	Pro	Ala
	770					775					780			Lys
Pro	His	Thr	Thr	Leu	Gln	Val	Ser	Gly	Arg	Gln	Gly	Gln	Pro	Thr
785					790				795					800
Asp	Cys	Asp	Ser	Val	Pro	Pro	Gly	Thr	Pro	Pro	Pro	Ile	Glu	Ser
				805					810					815
Thr	Phe	Lys	Glu	Asn	Gln	Leu	Cys	Arg	Asn	Thr	Cys	Gln	Glu	Leu
			820					825					830	His
Lys	His	Ser	Val	Glu	Gln	Thr	Glu	Thr	Ala	Glu	Leu	Pro	Ala	Ser
	835						840					845		Asp
Ser	Ile	Asn	Pro	Gly	Asn	Leu	Gln	Leu	Val	Ser	Glu	Leu	Lys	Asn
	850					855					860			Pro
Ser	Gly	Ser	Cys	Ser	Val	Asp	Val	Ser	Ala	Met	Phe	Trp	Glu	Arg
865					870				875					880
Gly	Cys	Lys	Glu	Pro	Cys	Ile	Ile	Thr	Ala	Cys	Glu	Asp	Val	Val
				885					890					895
Leu	Trp	Lys	Ala	Leu	Asp	Ala	Trp	Gln	Trp	Glu	Lys	Leu	Tyr	Thr
			900					905					910	Trp
His	Phe	Ala	Glu	Val	Pro	Val	Leu	Gln	Ile	Val	Pro	Val	Pro	Asp
		915					920				925			Val
Tyr	Asn	Leu	Val	Cys	Val	Ala	Leu	Gly	Asn	Leu	Glu	Ile	Arg	Glu
	930					935					940			Ile
Arg	Ala	Leu	Phe	Cys	Ser	Ser	Asp	Asp	Glu	Ser	Glu	Lys	Gln	Val
945					950					955				960
Leu	Lys	Ser	Gly	Asn	Ile	Lys	Ala	Val	Leu	Gly	Leu	Thr	Lys	Arg
				965					970					975
Leu	Val	Ser	Ser	Ser	Gly	Thr	Leu	Ser	Asp	Gln	Gln	Val	Glu	Val
			980					985					990	Met
Thr	Phe	Ala	Glu	Asp	Gly	Gly	Gly	Lys	Glu	Asn	Gln	Phe	Leu	Met
		995					1000					1005		Pro
Pro	Glu	Glu	Thr	Ile	Leu	Thr	Phe	Ala	Glu	Val	Gln	Gly	Met	Gln
	1010					1015					1020			Glu
Ala	Leu	Leu	Gly	Thr	Thr	Ile	Met	Asn	Asn	Ile	Val	Ile	Trp	Asn
1025					1030				1035					1040
Lys	Thr	Gly	Gln	Leu	Leu	Lys	Lys	Met	His	Ile	Asp	Asp	Ser	Tyr
				1045					1050					1055
Ala	Ser	Val	Cys	His	Lys	Ala	Tyr	Ser	Glu	Met	Gly	Leu	Leu	Phe
			1060					1065					1070	Ile
Val	Leu	Ser	His	Pro	Cys	Ala	Lys	Glu	Ser	Glu	Ser	Leu	Arg	Ser
		1075					1080					1085		Pro
Val	Phe	Gln	Leu	Ile	Val	Ile	Asn	Pro	Lys	Thr	Thr	Leu	Ser	Val
	1090					1095					1100			Gly
Val	Met	Leu	Tyr	Cys	Leu	Pro	Pro	Gly	Gln	Ala	Gly	Arg	Phe	Leu
1105					1110				1115					Glu
Gly	Asp	Val	Lys	Asp	His	Cys	Ala	Ala	Ala	Ile	Leu	Thr	Ser	Gly
				1125					1130					1135
Ile	Ala	Ile	Trp	Asp	Leu	Leu	Leu	Gly	Gln	Cys	Thr	Ala	Leu	Leu
														Pro

			1140					1145					1150						
Pro	Val	Ser	Asp	Gln	His	Trp	Ser	Phe	Val	Lys	Trp	Ser	Gly	Thr	Asp				
			1155					1160					1165						
Ser	His	Leu	Leu	Ala	Gly	Gln	Lys	Asp	Gly	Asn	Ile	Phe	Val	Tyr	His				
			1170				1175					1180							
Tyr	Ser																		
1185																			

<210> 175
 <211> 1014
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >PARP1|ENSG00000143799|ENST00000366794|3045

<400> 175

Met	Ala	Glu	Ser	Ser	Asp	Lys	Leu	Tyr	Arg	Val	Glu	Tyr	Ala	Lys	Ser				
1				5					10					15					
Gly	Arg	Ala	Ser	Cys	Lys	Lys	Cys	Ser	Glu	Ser	Ile	Pro	Lys	Asp	Ser				
			20					25					30						
Leu	Arg	Met	Ala	Ile	Met	Val	Gln	Ser	Pro	Met	Phe	Asp	Gly	Lys	Val				
		35					40					45							
Pro	His	Trp	Tyr	His	Phe	Ser	Cys	Phe	Trp	Lys	Val	Gly	His	Ser	Ile				
	50					55					60								
Arg	His	Pro	Asp	Val	Glu	Val	Asp	Gly	Phe	Ser	Glu	Leu	Arg	Trp	Asp				
65					70				75						80				
Asp	Gln	Gln	Lys	Val	Lys	Lys	Thr	Ala	Glu	Ala	Gly	Gly	Val	Thr	Gly				
				85					90					95					
Lys	Gly	Gln	Asp	Gly	Ile	Gly	Ser	Lys	Ala	Glu	Lys	Thr	Leu	Gly	Asp				
			100					105					110						
Phe	Ala	Ala	Glu	Tyr	Ala	Lys	Ser	Asn	Arg	Ser	Thr	Cys	Lys	Gly	Cys				
		115					120					125							
Met	Glu	Lys	Ile	Glu	Lys	Gly	Gln	Val	Arg	Leu	Ser	Lys	Lys	Met	Val				
	130					135					140								
Asp	Pro	Glu	Lys	Pro	Gln	Leu	Gly	Met	Ile	Asp	Arg	Trp	Tyr	His	Pro				
145					150					155					160				
Gly	Cys	Phe	Val	Lys	Asn	Arg	Glu	Glu	Leu	Gly	Phe	Arg	Pro	Glu	Tyr				
				165				170						175					
Ser	Ala	Ser	Gln	Leu	Lys	Gly	Phe	Ser	Leu	Leu	Ala	Thr	Glu	Asp	Lys				
			180					185					190						
Glu	Ala	Leu	Lys	Lys	Gln	Leu	Pro	Gly	Val	Lys	Ser	Glu	Gly	Lys	Arg				
		195					200					205							
Lys	Gly	Asp	Glu	Val	Asp	Gly	Val	Asp	Glu	Val	Ala	Lys	Lys	Lys	Ser				
	210					215					220								
Lys	Lys	Glu	Lys	Asp	Lys	Asp	Ser	Lys	Leu	Glu	Lys	Ala	Leu	Lys	Ala				
225					230					235					240				
Gln	Asn	Asp	Leu	Ile	Trp	Asn	Ile	Lys	Asp	Glu	Leu	Lys	Lys	Val	Cys				
				245					250					255					
Ser	Thr	Asn	Asp	Leu	Lys	Glu	Leu	Leu	Ile	Phe	Asn	Lys	Gln	Gln	Val				
		260						265					270						
Pro	Ser	Gly	Glu	Ser	Ala	Ile	Leu	Asp	Arg	Val	Ala	Asp	Gly	Met	Val				
		275					280					285							

Phe	Gly	Ala	Leu	Leu	Pro	Cys	Glu	Glu	Cys	Ser	Gly	Gln	Leu	Val	Phe
290						295					300				
Lys	Ser	Asp	Ala	Tyr	Tyr	Cys	Thr	Gly	Asp	Val	Thr	Ala	Trp	Thr	Lys
305					310					315					320
Cys	Met	Val	Lys	Thr	Gln	Thr	Pro	Asn	Arg	Lys	Glu	Trp	Val	Thr	Pro
				325					330						335
Lys	Glu	Phe	Arg	Glu	Ile	Ser	Tyr	Leu	Lys	Lys	Leu	Lys	Val	Lys	Lys
			340					345					350		
Gln	Asp	Arg	Ile	Phe	Pro	Pro	Glu	Thr	Ser	Ala	Ser	Val	Ala	Ala	Thr
		355					360					365			
Pro	Pro	Pro	Ser	Thr	Ala	Ser	Ala	Pro	Ala	Ala	Val	Asn	Ser	Ser	Ala
		370				375					380				
Ser	Ala	Asp	Lys	Pro	Leu	Ser	Asn	Met	Lys	Ile	Leu	Thr	Leu	Gly	Lys
385					390					395					400
Leu	Ser	Arg	Asn	Lys	Asp	Glu	Val	Lys	Ala	Met	Ile	Glu	Lys	Leu	Gly
				405					410						415
Gly	Lys	Leu	Thr	Gly	Thr	Ala	Asn	Lys	Ala	Ser	Leu	Cys	Ile	Ser	Thr
			420					425					430		
Lys	Lys	Glu	Val	Glu	Lys	Met	Asn	Lys	Lys	Met	Glu	Glu	Val	Lys	Glu
		435					440					445			
Ala	Asn	Ile	Arg	Val	Val	Ser	Glu	Asp	Phe	Leu	Gln	Asp	Val	Ser	Ala
		450				455					460				
Ser	Thr	Lys	Ser	Leu	Gln	Glu	Leu	Phe	Leu	Ala	His	Ile	Leu	Ser	Pro
465					470					475					480
Trp	Gly	Ala	Glu	Val	Lys	Ala	Glu	Pro	Val	Glu	Val	Val	Ala	Pro	Arg
				485					490					495	
Gly	Lys	Ser	Gly	Ala	Ala	Leu	Ser	Lys	Lys	Ser	Lys	Gly	Gln	Val	Lys
			500					505					510		
Glu	Glu	Gly	Ile	Asn	Lys	Ser	Glu	Lys	Arg	Met	Lys	Leu	Thr	Leu	Lys
		515					520					525			
Gly	Gly	Ala	Ala	Val	Asp	Pro	Asp	Ser	Gly	Leu	Glu	His	Ser	Ala	His
		530				535					540				
Val	Leu	Glu	Lys	Gly	Gly	Lys	Val	Phe	Ser	Ala	Thr	Leu	Gly	Leu	Val
545					550					555					560
Asp	Ile	Val	Lys	Gly	Thr	Asn	Ser	Tyr	Tyr	Lys	Leu	Gln	Leu	Leu	Glu
				565					570						575
Asp	Asp	Lys	Glu	Asn	Arg	Tyr	Trp	Ile	Phe	Arg	Ser	Trp	Gly	Arg	Val
			580					585					590		
Gly	Thr	Val	Ile	Gly	Ser	Asn	Lys	Leu	Glu	Gln	Met	Pro	Ser	Lys	Glu
		595					600					605			
Asp	Ala	Ile	Glu	His	Phe	Met	Lys	Leu	Tyr	Glu	Glu	Lys	Thr	Gly	Asn
						615					620				
Ala	Trp	His	Ser	Lys	Asn	Phe	Thr	Lys	Tyr	Pro	Lys	Lys	Phe	Tyr	Pro
625					630					635					640
Leu	Glu	Ile	Asp	Tyr	Gly	Gln	Asp	Glu	Glu	Ala	Val	Lys	Lys	Leu	Thr
				645					650						655
Val	Asn	Pro	Gly	Thr	Lys	Ser	Lys	Leu	Pro	Lys	Pro	Val	Gln	Asp	Leu
			660					665					670		
Ile	Lys	Met	Ile	Phe	Asp	Val	Glu	Ser	Met	Lys	Lys	Ala	Met	Val	Glu
		675					680					685			
Tyr	Glu	Ile	Asp	Leu	Gln	Lys	Met	Pro	Leu	Gly	Lys	Leu	Ser	Lys	Arg
	690					695					700				
Gln	Ile	Gln	Ala	Ala	Tyr	Ser	Ile	Leu	Ser	Glu	Val	Gln	Gln	Ala	Val
705					710					715					720

465					470					475					480
Asn	Thr	Gly	Leu	Leu	Leu	Leu	Ser	Glu	Val	Ala	Leu	Gly	Gln	Cys	Asn
				485					490					495	
Glu	Leu	Leu	Glu	Ala	Asn	Pro	Lys	Ala	Glu	Gly	Leu	Leu	Gln	Gly	Lys
			500					505					510		
His	Ser	Thr	Lys	Gly	Leu	Gly	Lys	Met	Ala	Pro	Ser	Ser	Ala	His	Phe
		515					520					525			
Val	Thr	Leu	Asn	Gly	Ser	Thr	Val	Pro	Leu	Gly	Pro	Ala	Ser	Asp	Thr
	530					535					540				
Gly	Ile	Leu	Asn	Pro	Asp	Gly	Tyr	Thr	Leu	Asn	Tyr	Asn	Glu	Tyr	Ile
545					550					555					560
Val	Tyr	Asn	Pro	Asn	Gln	Val	Arg	Met	Arg	Tyr	Leu	Leu	Lys	Val	Gln
				565					570					575	
Phe	Asn	Phe	Leu	Gln	Leu	Trp									
			580												

<210> 177

<211> 540

<212> PRT

<213> Homo sapiens

<220>

<223> >PARP3|ENSG00000041880|ENST00000398755|1623

<400> 177

Met	Ser	Leu	Leu	Phe	Leu	Ala	Met	Ala	Pro	Lys	Pro	Lys	Pro	Trp	Val
1				5				10						15	
Gln	Thr	Glu	Gly	Pro	Glu	Lys	Lys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu
			20					25					30		
Glu	Asp	Pro	Phe	Arg	Ser	Thr	Ala	Glu	Ala	Leu	Lys	Ala	Ile	Pro	Ala
		35					40					45			
Glu	Lys	Arg	Ile	Ile	Arg	Val	Asp	Pro	Thr	Cys	Pro	Leu	Ser	Ser	Asn
	50					55					60				
Pro	Gly	Thr	Gln	Val	Tyr	Glu	Asp	Tyr	Asn	Cys	Thr	Leu	Asn	Gln	Thr
65				70					75						80
Asn	Ile	Glu	Asn	Asn	Asn	Lys	Phe	Tyr	Ile	Ile	Gln	Leu	Leu	Gln	
			85					90						95	
Asp	Ser	Asn	Arg	Phe	Phe	Thr	Cys	Trp	Asn	His	Trp	Gly	Arg	Val	Gly
			100				105						110		
Glu	Val	Gly	Gln	Ser	Lys	Ile	Asn	His	Phe	Thr	Arg	Leu	Glu	Asp	Ala
		115					120					125			
Lys	Lys	Asp	Phe	Glu	Lys	Lys	Phe	Arg	Glu	Lys	Thr	Lys	Asn	Asn	Trp
	130					135					140				
Ala	Glu	Arg	Asp	His	Phe	Val	Ser	His	Pro	Gly	Lys	Tyr	Thr	Leu	Ile
145					150					155					160
Glu	Val	Gln	Ala	Glu	Asp	Glu	Ala	Gln	Glu	Ala	Val	Val	Lys	Val	Asp
				165					170					175	
Arg	Gly	Pro	Val	Arg	Thr	Val	Thr	Lys	Arg	Val	Gln	Pro	Cys	Ser	Leu
			180					185					190		
Asp	Pro	Ala	Thr	Gln	Lys	Leu	Ile	Thr	Asn	Ile	Phe	Ser	Lys	Glu	Met
		195				200						205			
Phe	Lys	Asn	Thr	Met	Ala	Leu	Met	Asp	Leu	Asp	Val	Lys	Lys	Met	Pro
	210					215					220				

Leu Gly Lys Leu Ser Lys Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu
 225 230 235 240
 Glu Ala Leu Glu Glu Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser
 245 250 255
 Leu Glu Glu Leu Ser Ser His Phe Tyr Thr Val Ile Pro His Asn Phe
 260 265 270
 Gly His Ser Gln Pro Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala
 275 280 285
 Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala
 290 295 300
 Leu Gln Ala Val Ser Glu Gln Glu Lys Thr Val Glu Glu Val Pro His
 305 310 315 320
 Pro Leu Asp Arg Asp Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu
 325 330 335
 Asp Ser Gly Ala Pro Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln
 340 345 350
 Thr Gly Ser Asn His Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val
 355 360 365
 Asn Gln Glu Gly Glu Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly
 370 375 380
 Asn Arg Lys Leu Leu Trp His Gly Thr Asn Met Ala Val Val Ala Ala
 385 390 395 400
 Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser Gly Gly Arg Val
 405 410 415
 Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr
 420 425 430
 Val Ile Gly Met Lys Cys Gly Ala His His Val Gly Tyr Met Phe Leu
 435 440 445
 Gly Glu Val Ala Leu Gly Arg Glu His His Ile Asn Thr Asp Asn Pro
 450 455 460
 Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly
 465 470 475 480
 His Thr Glu Pro Asp Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly
 485 490 495
 Gln Gln Val Val Val Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe
 500 505 510
 Ser Ser Ser Thr Phe Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser
 515 520 525
 Gln Cys Arg Leu Arg Tyr Leu Leu Glu Val His Leu
 530 535 540

<210> 178

<211> 1724

<212> PRT

<213> Homo sapiens

<220>

<223> >PARP4|ENSG00000102699|ENST00000381989|5175

<400> 178

Met Val Met Gly Ile Phe Ala Asn Cys Ile Phe Cys Leu Lys Val Lys
 1 5 10 15
 Tyr Leu Pro Gln Gln Gln Lys Lys Lys Leu Gln Thr Asp Ile Lys Glu

450		455		460
Gln Arg Thr Asp Val	Gly Asn Leu Gly Ser	Gly Ile Tyr Phe Ser Asp		
465	470	475	480	
Ser Leu Ser Thr Ser	Ile Lys Tyr Ser His	Pro Gly Glu Thr Asp Gly		
	485	490	495	
Thr Arg Leu Leu Leu	Ile Cys Asp Val Ala	Leu Gly Lys Cys Met Asp		
	500	505	510	
Leu His Glu Lys Asp	Phe Ser Leu Thr Glu	Ala Pro Pro Gly Tyr Asp		
	515	520	525	
Ser Val His Gly Val	Ser Gln Thr Ala Ser	Val Thr Thr Asp Phe Glu		
	530	535	540	
Asp Asp Glu Phe Val	Val Tyr Lys Thr Asn	Gln Val Lys Met Lys Tyr		
545	550	555	560	
Ile Ile Lys Phe Ser	Met Pro Gly Asp Gln	Ile Lys Asp Phe His Pro		
	565	570	575	
Ser Asp His Thr Glu	Leu Glu Glu Tyr Arg	Pro Glu Phe Ser Asn Phe		
	580	585	590	
Ser Lys Val Glu Asp	Tyr Gln Leu Pro Asp	Ala Lys Thr Ser Ser Ser		
	595	600	605	
Thr Lys Ala Gly Leu	Gln Asp Ala Ser Gly	Asn Leu Val Pro Leu Glu		
	610	615	620	
Asp Val His Ile Lys	Gly Arg Ile Ile Asp	Thr Val Ala Gln Val Ile		
625	630	635	640	
Val Phe Gln Thr Tyr	Thr Asn Lys Ser His	Val Pro Ile Glu Ala Lys		
	645	650	655	
Tyr Ile Phe Pro Leu	Asp Asp Lys Ala Ala	Val Cys Gly Phe Glu Ala		
	660	665	670	
Phe Ile Asn Gly Lys	His Ile Val Gly Glu	Ile Lys Glu Lys Glu Glu		
	675	680	685	
Ala Gln Gln Glu Tyr	Leu Glu Ala Val Thr	Gln Gly His Gly Ala Tyr		
	690	695	700	
Leu Met Ser Gln Asp	Ala Pro Asp Val Phe	Thr Val Ser Val Gly Asn		
705	710	715	720	
Leu Pro Pro Lys Ala	Lys Val Leu Ile Lys	Ile Thr Tyr Ile Thr Glu		
	725	730	735	
Leu Ser Ile Leu Gly	Thr Val Gly Val Phe	Phe Met Pro Ala Thr Val		
	740	745	750	
Ala Pro Trp Gln Gln	Asp Lys Ala Leu Asn	Glu Asn Leu Gln Asp Thr		
	755	760	765	
Val Glu Lys Ile Cys	Ile Lys Glu Ile Gly	Thr Lys Gln Ser Phe Ser		
	770	775	780	
Leu Thr Met Ser Ile	Glu Met Pro Tyr Val	Ile Glu Phe Ile Phe Ser		
785	790	795	800	
Asp Thr His Glu Leu	Lys Gln Lys Arg Thr	Asp Cys Lys Ala Val Ile		
	805	810	815	
Ser Thr Met Glu Gly	Ser Ser Leu Asp Ser	Ser Gly Phe Ser Leu His		
	820	825	830	
Ile Gly Leu Ser Ala	Ala Tyr Leu Pro Arg	Met Trp Val Glu Lys His		
	835	840	845	
Pro Glu Lys Glu Ser	Glu Ala Cys Met Leu	Val Phe Gln Pro Asp Leu		
	850	855	860	
Asp Val Asp Leu Pro	Asp Leu Ala Ser Glu	Ser Glu Val Ile Ile Cys		
865	870	875	880	
Leu Asp Cys Ser Ser	Ser Met Glu Gly Val	Thr Phe Leu Gln Ala Lys		

1315	1320	1325
Pro Thr Ala Arg Ala His Ser Pro Ala Ser Leu Ser Phe Ala Ser Tyr		
1330	1335	1340
Arg Gln Val Ala Ser Phe Gly Ser Ala Ala Pro Pro Arg Gln Phe Asp		
1345	1350	1355
Ala Ser Gln Phe Ser Gln Gly Pro Val Pro Gly Thr Cys Ala Asp Trp		
1365	1370	1375
Ile Pro Gln Ser Ala Ser Cys Pro Thr Gly Pro Pro Gln Asn Pro Pro		
1380	1385	1390
Ser Ser Pro Tyr Cys Gly Ile Val Phe Ser Gly Ser Ser Leu Ser Ser		
1395	1400	1405
Ala Gln Ser Ala Pro Leu Gln His Pro Gly Gly Phe Thr Thr Arg Pro		
1410	1415	1420
Ser Ala Gly Thr Phe Pro Glu Leu Asp Ser Pro Gln Leu His Phe Ser		
1425	1430	1435
Leu Pro Thr Asp Pro Asp Pro Ile Arg Gly Phe Gly Ser Tyr His Pro		
1445	1450	1455
Ser Ala Ser Ser Pro Phe His Phe Gln Pro Ser Ala Ala Ser Leu Thr		
1460	1465	1470
Ala Asn Leu Arg Leu Pro Met Ala Ser Ala Leu Pro Glu Ala Leu Cys		
1475	1480	1485
Ser Gln Ser Arg Thr Thr Pro Val Asp Leu Cys Leu Leu Glu Glu Ser		
1490	1495	1500
Val Gly Ser Leu Glu Gly Ser Arg Cys Pro Val Phe Ala Phe Gln Ser		
1505	1510	1515
Ser Asp Thr Glu Ser Asp Glu Leu Ser Glu Val Leu Gln Asp Ser Cys		
1525	1530	1535
Phe Leu Gln Ile Lys Cys Asp Thr Lys Asp Asp Ser Ile Leu Cys Phe		
1540	1545	1550
Leu Glu Val Lys Glu Glu Asp Glu Ile Val Cys Ile Gln His Trp Gln		
1555	1560	1565
Asp Ala Val Pro Trp Thr Glu Leu Leu Ser Leu Gln Thr Glu Asp Gly		
1570	1575	1580
Phe Trp Lys Leu Thr Pro Glu Leu Gly Leu Ile Leu Asn Leu Asn Thr		
1585	1590	1595
Asn Gly Leu His Ser Phe Leu Lys Gln Lys Gly Ile Gln Ser Leu Gly		
1605	1610	1615
Val Lys Gly Arg Glu Cys Leu Leu Asp Leu Ile Ala Thr Met Leu Val		
1620	1625	1630
Leu Gln Phe Ile Arg Thr Arg Leu Glu Lys Glu Gly Ile Val Phe Lys		
1635	1640	1645
Ser Leu Met Lys Met Asp Asp Ala Ser Ile Ser Arg Asn Ile Pro Trp		
1650	1655	1660
Ala Phe Glu Ala Ile Lys Gln Ala Ser Glu Trp Val Arg Arg Thr Glu		
1665	1670	1675
Gly Gln Tyr Pro Ser Ile Cys Pro Arg Leu Glu Leu Gly Asn Asp Trp		
1685	1690	1695
Asp Ser Ala Thr Lys Gln Leu Leu Gly Leu Gln Pro Ile Ser Thr Val		
1700	1705	1710
Ser Pro Leu His Arg Val Leu His Tyr Ser Gln Gly		
1715	1720	

<210> 179

<211> 261

<212> PRT
 <213> Homo sapiens

<220>
 <223> >PCNA|ENSG00000132646|ENST00000379160|786

<400> 179
 Met Phe Glu Ala Arg Leu Val Gln Gly Ser Ile Leu Lys Lys Val Leu
 1 5 10 15
 Glu Ala Leu Lys Asp Leu Ile Asn Glu Ala Cys Trp Asp Ile Ser Ser
 20 25 30
 Ser Gly Val Asn Leu Gln Ser Met Asp Ser Ser His Val Ser Leu Val
 35 40 45
 Gln Leu Thr Leu Arg Ser Glu Gly Phe Asp Thr Tyr Arg Cys Asp Arg
 50 55 60
 Asn Leu Ala Met Gly Val Asn Leu Thr Ser Met Ser Lys Ile Leu Lys
 65 70 75 80
 Cys Ala Gly Asn Glu Asp Ile Ile Thr Leu Arg Ala Glu Asp Asn Ala
 85 90 95
 Asp Thr Leu Ala Leu Val Phe Glu Ala Pro Asn Gln Glu Lys Val Ser
 100 105 110
 Asp Tyr Glu Met Lys Leu Met Asp Leu Asp Val Glu Gln Leu Gly Ile
 115 120 125
 Pro Glu Gln Glu Tyr Ser Cys Val Val Lys Met Pro Ser Gly Glu Phe
 130 135 140
 Ala Arg Ile Cys Arg Asp Leu Ser His Ile Gly Asp Ala Val Val Ile
 145 150 155 160
 Ser Cys Ala Lys Asp Gly Val Lys Phe Ser Ala Ser Gly Glu Leu Gly
 165 170 175
 Asn Gly Asn Ile Lys Leu Ser Gln Thr Ser Asn Val Asp Lys Glu Glu
 180 185 190
 Glu Ala Val Thr Ile Glu Met Asn Glu Pro Val Gln Leu Thr Phe Ala
 195 200 205
 Leu Arg Tyr Leu Asn Phe Phe Thr Lys Ala Thr Pro Leu Ser Ser Thr
 210 215 220
 Val Thr Leu Ser Met Ser Ala Asp Val Pro Leu Val Val Glu Tyr Lys
 225 230 235 240
 Ile Ala Asp Met Gly His Leu Lys Tyr Tyr Leu Ala Pro Lys Ile Glu
 245 250 255
 Asp Glu Glu Gly Ser
 260

<210> 180
 <211> 1068
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >PIK3CA|ENSG00000121879|ENST00000263967|3207

<400> 180
 Met Pro Pro Arg Pro Ser Ser Gly Glu Leu Trp Gly Ile His Leu Met

1				5				10					15				
Pro	Pro	Arg	Ile	Leu	Val	Glu	Cys	Leu	Leu	Pro	Asn	Gly	Met	Ile	Val		
			20					25					30				
Thr	Leu	Glu	Cys	Leu	Arg	Glu	Ala	Thr	Leu	Ile	Thr	Ile	Lys	His	Glu		
		35					40					45					
Leu	Phe	Lys	Glu	Ala	Arg	Lys	Tyr	Pro	Leu	His	Gln	Leu	Leu	Gln	Asp		
	50					55					60						
Glu	Ser	Ser	Tyr	Ile	Phe	Val	Ser	Val	Thr	Gln	Glu	Ala	Glu	Arg	Glu		
65					70					75					80		
Glu	Phe	Phe	Asp	Glu	Thr	Arg	Arg	Leu	Cys	Asp	Leu	Arg	Leu	Phe	Gln		
			85					90						95			
Pro	Phe	Leu	Lys	Val	Ile	Glu	Pro	Val	Gly	Asn	Arg	Glu	Glu	Lys	Ile		
			100					105					110				
Leu	Asn	Arg	Glu	Ile	Gly	Phe	Ala	Ile	Gly	Met	Pro	Val	Cys	Glu	Phe		
		115					120					125					
Asp	Met	Val	Lys	Asp	Pro	Glu	Val	Gln	Asp	Phe	Arg	Arg	Asn	Ile	Leu		
	130					135					140						
Asn	Val	Cys	Lys	Glu	Ala	Val	Asp	Leu	Arg	Asp	Leu	Asn	Ser	Pro	His		
145					150					155					160		
Ser	Arg	Ala	Met	Tyr	Val	Tyr	Pro	Pro	Asn	Val	Glu	Ser	Ser	Pro	Glu		
			165						170					175			
Leu	Pro	Lys	His	Ile	Tyr	Asn	Lys	Leu	Asp	Lys	Gly	Gln	Ile	Ile	Val		
			180					185					190				
Val	Ile	Trp	Val	Ile	Val	Ser	Pro	Asn	Asn	Asp	Lys	Gln	Lys	Tyr	Thr		
		195					200					205					
Leu	Lys	Ile	Asn	His	Asp	Cys	Val	Pro	Glu	Gln	Val	Ile	Ala	Glu	Ala		
	210					215					220						
Ile	Arg	Lys	Lys	Thr	Arg	Ser	Met	Leu	Leu	Ser	Ser	Glu	Gln	Leu	Lys		
225					230					235					240		
Leu	Cys	Val	Leu	Glu	Tyr	Gln	Gly	Lys	Tyr	Ile	Leu	Lys	Val	Cys	Gly		
			245					250						255			
Cys	Asp	Glu	Tyr	Phe	Leu	Glu	Lys	Tyr	Pro	Leu	Ser	Gln	Tyr	Lys	Tyr		
		260						265					270				
Ile	Arg	Ser	Cys	Ile	Met	Leu	Gly	Arg	Met	Pro	Asn	Leu	Met	Leu	Met		
		275					280					285					
Ala	Lys	Glu	Ser	Leu	Tyr	Ser	Gln	Leu	Pro	Met	Asp	Cys	Phe	Thr	Met		
	290					295					300						
Pro	Ser	Tyr	Ser	Arg	Arg	Ile	Ser	Thr	Ala	Thr	Pro	Tyr	Met	Asn	Gly		
305					310					315					320		
Glu	Thr	Ser	Thr	Lys	Ser	Leu	Trp	Val	Ile	Asn	Ser	Ala	Leu	Arg	Ile		
			325					330						335			
Lys	Ile	Leu	Cys	Ala	Thr	Tyr	Val	Asn	Val	Asn	Ile	Arg	Asp	Ile	Asp		
			340					345					350				
Lys	Ile	Tyr	Val	Arg	Thr	Gly	Ile	Tyr	His	Gly	Gly	Glu	Pro	Leu	Cys		
		355				360						365					
Asp	Asn	Val	Asn	Thr	Gln	Arg	Val	Pro	Cys	Ser	Asn	Pro	Arg	Trp	Asn		
	370					375					380						
Glu	Trp	Leu	Asn	Tyr	Asp	Ile	Tyr	Ile	Pro	Asp	Leu	Pro	Arg	Ala	Ala		
385					390					395					400		
Arg	Leu	Cys	Leu	Ser	Ile	Cys	Ser	Val	Lys	Gly	Arg	Lys	Gly	Ala	Lys		
			405					410						415			
Glu	Glu	His	Cys	Pro	Leu	Ala	Trp	Gly	Asn	Ile	Asn	Leu	Phe	Asp	Tyr		
			420					425					430				
Thr	Asp	Thr	Leu	Val	Ser	Gly	Lys	Met	Ala	Leu	Asn	Leu	Trp	Pro	Val		

865					870					875				880	
Leu	Lys	Asp	Lys	Asn	Lys	Gly	Glu	Ile	Tyr	Asp	Ala	Ala	Ile	Asp	Leu
				885					890					895	
Phe	Thr	Arg	Ser	Cys	Ala	Gly	Tyr	Cys	Val	Ala	Thr	Phe	Ile	Leu	Gly
			900					905					910		
Ile	Gly	Asp	Arg	His	Asn	Ser	Asn	Ile	Met	Val	Lys	Asp	Asp	Gly	Gln
		915					920					925			
Leu	Phe	His	Ile	Asp	Phe	Gly	His	Phe	Leu	Asp	His	Lys	Lys	Lys	Lys
	930					935				940					
Phe	Gly	Tyr	Lys	Arg	Glu	Arg	Val	Pro	Phe	Val	Leu	Thr	Gln	Asp	Phe
945					950					955					960
Leu	Ile	Val	Ile	Ser	Lys	Gly	Ala	Gln	Glu	Cys	Thr	Lys	Thr	Arg	Glu
				965					970					975	
Phe	Glu	Arg	Phe	Gln	Glu	Met	Cys	Tyr	Lys	Ala	Tyr	Leu	Ala	Ile	Arg
			980					985					990		
Gln	His	Ala	Asn	Leu	Phe	Ile	Asn	Leu	Phe	Ser	Met	Met	Leu	Gly	Ser
		995					1000					1005			
Gly	Met	Pro	Glu	Leu	Gln	Ser	Phe	Asp	Asp	Ile	Ala	Tyr	Ile	Arg	Lys
	1010					1015					1020				
Thr	Leu	Ala	Leu	Asp	Lys	Thr	Glu	Gln	Glu	Ala	Leu	Glu	Tyr	Phe	Met
1025					1030					1035					1040
Lys	Gln	Met	Asn	Asp	Ala	His	His	Gly	Gly	Trp	Thr	Thr	Lys	Met	Asp
			1045					1050						1055	
Trp	Ile	Phe	His	Thr	Ile	Lys	Gln	His	Ala	Leu	Asn				
			1060				1065								

<210> 181
 <211> 862
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >PMS2|ENSG00000122512|ENST00000265849|2589

<400> 181															
Met	Glu	Arg	Ala	Glu	Ser	Ser	Ser	Thr	Glu	Pro	Ala	Lys	Ala	Ile	Lys
1				5					10					15	
Pro	Ile	Asp	Arg	Lys	Ser	Val	His	Gln	Ile	Cys	Ser	Gly	Gln	Val	Val
		20						25					30		
Leu	Ser	Leu	Ser	Thr	Ala	Val	Lys	Glu	Leu	Val	Glu	Asn	Ser	Leu	Asp
		35					40					45			
Ala	Gly	Ala	Thr	Asn	Ile	Asp	Leu	Lys	Leu	Lys	Asp	Tyr	Gly	Val	Asp
	50					55					60				
Leu	Ile	Glu	Val	Ser	Asp	Asn	Gly	Cys	Gly	Val	Glu	Glu	Glu	Asn	Phe
65					70					75					80
Glu	Gly	Leu	Thr	Leu	Lys	His	His	Thr	Ser	Lys	Ile	Gln	Glu	Phe	Ala
			85					90						95	
Asp	Leu	Thr	Gln	Val	Glu	Thr	Phe	Gly	Phe	Arg	Gly	Glu	Ala	Leu	Ser
			100					105					110		
Ser	Leu	Cys	Ala	Leu	Ser	Asp	Val	Thr	Ile	Ser	Thr	Cys	His	Ala	Ser
		115					120					125			
Ala	Lys	Val	Gly	Thr	Arg	Leu	Met	Phe	Asp	His	Asn	Gly	Lys	Ile	Ile
	130						135				140				

Gln	Lys	Thr	Pro	Tyr	Pro	Arg	Pro	Arg	Gly	Thr	Thr	Val	Ser	Val	Gln
145					150					155					160
Gln	Leu	Phe	Ser	Thr	Leu	Pro	Val	Arg	His	Lys	Glu	Phe	Gln	Arg	Asn
				165					170					175	
Ile	Lys	Lys	Glu	Tyr	Ala	Lys	Met	Val	Gln	Val	Leu	His	Ala	Tyr	Cys
			180					185					190		
Ile	Ile	Ser	Ala	Gly	Ile	Arg	Val	Ser	Cys	Thr	Asn	Gln	Leu	Gly	Gln
		195					200				205				
Gly	Lys	Arg	Gln	Pro	Val	Val	Cys	Thr	Gly	Gly	Ser	Pro	Ser	Ile	Lys
	210					215					220				
Glu	Asn	Ile	Gly	Ser	Val	Phe	Gly	Gln	Lys	Gln	Leu	Gln	Ser	Leu	Ile
225					230					235					240
Pro	Phe	Val	Gln	Leu	Pro	Pro	Ser	Asp	Ser	Val	Cys	Glu	Glu	Tyr	Gly
				245					250					255	
Leu	Ser	Cys	Ser	Asp	Ala	Leu	His	Asn	Leu	Phe	Tyr	Ile	Ser	Gly	Phe
			260					265					270		
Ile	Ser	Gln	Cys	Thr	His	Gly	Val	Gly	Arg	Ser	Ser	Thr	Asp	Arg	Gln
		275					280					285			
Phe	Phe	Phe	Ile	Asn	Arg	Arg	Pro	Cys	Asp	Pro	Ala	Lys	Val	Cys	Arg
	290					295					300				
Leu	Val	Asn	Glu	Val	Tyr	His	Met	Tyr	Asn	Arg	His	Gln	Tyr	Pro	Phe
305					310					315					320
Val	Val	Leu	Asn	Ile	Ser	Val	Asp	Ser	Glu	Cys	Val	Asp	Ile	Asn	Val
			325						330					335	
Thr	Pro	Asp	Lys	Arg	Gln	Ile	Leu	Leu	Gln	Glu	Glu	Lys	Leu	Leu	Leu
			340					345					350		
Ala	Val	Leu	Lys	Thr	Ser	Leu	Ile	Gly	Met	Phe	Asp	Ser	Asp	Val	Asn
		355					360				365				
Lys	Leu	Asn	Val	Ser	Gln	Gln	Pro	Leu	Leu	Asp	Val	Glu	Gly	Asn	Leu
	370					375				380					
Ile	Lys	Met	His	Ala	Ala	Asp	Leu	Glu	Lys	Pro	Met	Val	Glu	Lys	Gln
385					390					395					400
Asp	Gln	Ser	Pro	Ser	Leu	Arg	Thr	Gly	Glu	Glu	Lys	Lys	Asp	Val	Ser
			405						410					415	
Ile	Ser	Arg	Leu	Arg	Glu	Ala	Phe	Ser	Leu	Arg	His	Thr	Thr	Glu	Asn
			420					425					430		
Lys	Pro	His	Ser	Pro	Lys	Thr	Pro	Glu	Pro	Arg	Arg	Ser	Pro	Leu	Gly
		435					440					445			
Gln	Lys	Arg	Gly	Met	Leu	Ser	Ser	Ser	Thr	Ser	Gly	Ala	Ile	Ser	Asp
	450					455					460				
Lys	Gly	Val	Leu	Arg	Pro	Gln	Lys	Glu	Ala	Val	Ser	Ser	Ser	His	Gly
465					470					475					480
Pro	Ser	Asp	Pro	Thr	Asp	Arg	Ala	Glu	Val	Glu	Lys	Asp	Ser	Gly	His
				485					490					495	
Gly	Ser	Thr	Ser	Val	Asp	Ser	Glu	Gly	Phe	Ser	Ile	Pro	Asp	Thr	Gly
			500					505					510		
Ser	His	Cys	Ser	Ser	Glu	Tyr	Ala	Ala	Ser	Ser	Pro	Gly	Asp	Arg	Gly
		515					520					525			
Ser	Gln	Glu	His	Val	Asp	Ser	Gln	Glu	Lys	Ala	Pro	Lys	Thr	Asp	Asp
	530					535					540				
Ser	Phe	Ser	Asp	Val	Asp	Cys	His	Ser	Asn	Gln	Glu	Asp	Thr	Gly	Cys
545					550					555					560
Lys	Phe	Arg	Val	Leu	Pro	Gln	Pro	Thr	Asn	Leu	Ala	Thr	Pro	Asn	Thr
				565					570					575	

Lys	Arg	Phe	Lys	Lys	Glu	Glu	Ile	Leu	Ser	Ser	Ser	Asp	Ile	Cys	Gln	
			580					585					590			
Lys	Leu	Val	Asn	Thr	Gln	Asp	Met	Ser	Ala	Ser	Gln	Val	Asp	Val	Ala	
		595					600					605				
Val	Lys	Ile	Asn	Lys	Lys	Val	Val	Pro	Leu	Asp	Phe	Ser	Met	Ser	Ser	
	610					615					620					
Leu	Ala	Lys	Arg	Ile	Lys	Gln	Leu	His	His	Glu	Ala	Gln	Gln	Ser	Glu	
625					630					635					640	
Gly	Glu	Gln	Asn	Tyr	Arg	Lys	Phe	Arg	Ala	Lys	Ile	Cys	Pro	Gly	Glu	
			645					650						655		
Asn	Gln	Ala	Ala	Glu	Asp	Glu	Leu	Arg	Lys	Glu	Ile	Ser	Lys	Thr	Met	
		660						665					670			
Phe	Ala	Glu	Met	Glu	Ile	Ile	Gly	Gln	Phe	Asn	Leu	Gly	Phe	Ile	Ile	
		675					680					685				
Thr	Lys	Leu	Asn	Glu	Asp	Ile	Phe	Ile	Val	Asp	Gln	His	Ala	Thr	Asp	
	690					695					700					
Glu	Lys	Tyr	Asn	Phe	Glu	Met	Leu	Gln	Gln	His	Thr	Val	Leu	Gln	Gly	
705					710					715					720	
Gln	Arg	Leu	Ile	Ala	Pro	Gln	Thr	Leu	Asn	Leu	Thr	Ala	Val	Asn	Glu	
			725					730						735		
Ala	Val	Leu	Ile	Glu	Asn	Leu	Glu	Ile	Phe	Arg	Lys	Asn	Gly	Phe	Asp	
		740					745						750			
Phe	Val	Ile	Asp	Glu	Asn	Ala	Pro	Val	Thr	Glu	Arg	Ala	Lys	Leu	Ile	
	755					760						765				
Ser	Leu	Pro	Thr	Ser	Lys	Asn	Trp	Thr	Phe	Gly	Pro	Gln	Asp	Val	Asp	
	770					775					780					
Glu	Leu	Ile	Phe	Met	Leu	Ser	Asp	Ser	Pro	Gly	Val	Met	Cys	Arg	Pro	
785					790					795					800	
Ser	Arg	Val	Lys	Gln	Met	Phe	Ala	Ser	Arg	Ala	Cys	Arg	Lys	Ser	Val	
			805				810							815		
Met	Ile	Gly	Thr	Ala	Leu	Asn	Thr	Ser	Glu	Met	Lys	Lys	Leu	Ile	Thr	
		820					825						830			
His	Met	Gly	Glu	Met	Asp	His	Pro	Trp	Asn	Cys	Pro	His	Gly	Arg	Pro	
	835					840						845				
Thr	Met	Arg	His	Ile	Ala	Asn	Leu	Gly	Val	Ile	Ser	Gln	Asn			
	850					855					860					

<210> 182
 <211> 1462
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >POLA1|ENSG00000101868|ENST00000379059|4389

Met	Ala	Pro	Val	His	Gly	Asp	Asp	Ser	Leu	Ser	Asp	Ser	Gly	Ser	Phe	
1				5				10					15			
Val	Ser	Ser	Arg	Ala	Arg	Arg	Glu	Lys	Lys	Ser	Lys	Lys	Gly	Arg	Gln	
			20				25						30			
Glu	Ala	Leu	Glu	Arg	Leu	Lys	Lys	Ala	Lys	Ala	Gly	Glu	Lys	Tyr	Lys	
	35					40					45					
Tyr	Glu	Val	Glu	Asp	Phe	Thr	Gly	Val	Tyr	Glu	Glu	Val	Asp	Glu	Glu	

50		55		60											
Gln Tyr Ser Lys Leu Val	Gln Ala Arg Gln Asp Asp Asp Trp Ile Val														
65	70	75	80												
Asp Asp Asp Gly Ile Gly Tyr Val Glu Asp Gly Arg Glu Ile Phe Asp															
85	90	95													
Asp Asp Leu Glu Asp Asp Ala Leu Asp Ala Asp Glu Lys Gly Lys Asp															
100	105	110													
Gly Lys Ala Arg Asn Lys Asp Lys Arg Asn Val Lys Lys Leu Ala Val															
115	120	125													
Thr Lys Pro Asn Asn Ile Lys Ser Met Phe Ile Ala Cys Ala Gly Lys															
130	135	140													
Lys Thr Ala Asp Lys Ala Val Asp Leu Ser Lys Asp Gly Leu Leu Gly															
145	150	155													
Asp Ile Leu Gln Asp Leu Asn Thr Glu Thr Pro Gln Ile Thr Pro Pro															
165	170	175													
Pro Val Met Ile Leu Lys Lys Lys Arg Ser Ile Gly Ala Ser Pro Asn															
180	185	190													
Pro Phe Ser Val His Thr Ala Thr Ala Val Pro Ser Gly Lys Ile Ala															
195	200	205													
Ser Pro Val Ser Arg Lys Glu Pro Pro Leu Thr Pro Val Pro Leu Lys															
210	215	220													
Arg Ala Glu Phe Ala Gly Asp Asp Val Gln Val Glu Ser Thr Glu Glu															
225	230	235													
Glu Gln Glu Ser Gly Ala Met Glu Phe Glu Asp Gly Asp Phe Asp Glu															
245	250	255													
Pro Met Glu Val Glu Glu Val Asp Leu Glu Pro Met Ala Ala Lys Ala															
260	265	270													
Trp Asp Lys Glu Ser Glu Pro Ala Glu Glu Val Lys Gln Glu Ala Asp															
275	280	285													
Ser Gly Lys Gly Thr Val Ser Tyr Leu Gly Ser Phe Leu Pro Asp Val															
290	295	300													
Ser Cys Trp Asp Ile Asp Gln Glu Gly Asp Ser Ser Phe Ser Val Gln															
305	310	315													
Glu Val Gln Val Asp Ser Ser His Leu Pro Leu Val Lys Gly Ala Asp															
325	330	335													
Glu Glu Gln Val Phe His Phe Tyr Trp Leu Asp Ala Tyr Glu Asp Gln															
340	345	350													
Tyr Asn Gln Pro Gly Val Val Phe Leu Phe Gly Lys Val Trp Ile Glu															
355	360	365													
Ser Ala Glu Thr His Val Ser Cys Cys Val Met Val Lys Asn Ile Glu															
370	375	380													
Arg Thr Leu Tyr Phe Leu Pro Arg Glu Met Lys Ile Asp Leu Asn Thr															
385	390	395													
Gly Lys Glu Thr Gly Thr Pro Ile Ser Met Lys Asp Val Tyr Glu Glu															
405	410	415													
Phe Asp Glu Lys Ile Ala Thr Lys Tyr Lys Ile Met Lys Phe Lys Ser															
420	425	430													
Lys Pro Val Glu Lys Asn Tyr Ala Phe Glu Ile Pro Asp Val Pro Glu															
435	440	445													
Lys Ser Glu Tyr Leu Glu Val Lys Tyr Ser Ala Glu Met Pro Gln Leu															
450	455	460													
Pro Gln Asp Leu Lys Gly Glu Thr Phe Ser His Val Phe Gly Thr Asn															
465	470	475													
Thr Ser Ser Leu Glu Leu Phe Leu Met Asn Arg Lys Ile Lys Gly Pro															

1345		1350		1355		1360
Pro	Leu	Gln	Phe	Ser	Arg	Thr
		1365		1370		1375
Ala	Thr	Leu	Gln	Pro	Glu	Tyr
		1380		1385		1390
Cys	Phe	Tyr	Arg	Tyr	Ile	Phe
		1395		1400		1405
Thr	Thr	Asp	His	Glu	Lys	Asp
		1410		1415		1420
Lys	Val	Leu	Gln	Asp	Tyr	Arg
		1425		1430		1435
Leu	Ser	Arg	Ser	Gly	Tyr	Ser
				1445		1450
Gly	Cys	Ala	Val	Lys	Ser	
				1460		

<210> 183
 <211> 335
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >POLB|ENSG00000070501|ENST00000265421|1008

<400> 183

Met	Ser	Lys	Arg	Lys	Ala	Pro	Gln	Glu	Thr	Leu	Asn	Gly	Gly	Ile	Thr
1				5				10						15	
Asp	Met	Leu	Thr	Glu	Leu	Ala	Asn	Phe	Glu	Lys	Asn	Val	Ser	Gln	Ala
			20					25					30		
Ile	His	Lys	Tyr	Asn	Ala	Tyr	Arg	Lys	Ala	Ala	Ser	Val	Ile	Ala	Lys
		35					40					45			
Tyr	Pro	His	Lys	Ile	Lys	Ser	Gly	Ala	Glu	Ala	Lys	Lys	Leu	Pro	Gly
		50				55					60				
Val	Gly	Thr	Lys	Ile	Ala	Glu	Lys	Ile	Asp	Glu	Phe	Leu	Ala	Thr	Gly
65					70				75						80
Lys	Leu	Arg	Lys	Leu	Glu	Lys	Ile	Arg	Gln	Asp	Asp	Thr	Ser	Ser	Ser
				85				90						95	
Ile	Asn	Phe	Leu	Thr	Arg	Val	Ser	Gly	Ile	Gly	Pro	Ser	Ala	Ala	Arg
			100					105					110		
Lys	Phe	Val	Asp	Glu	Gly	Ile	Lys	Thr	Leu	Glu	Asp	Leu	Arg	Lys	Asn
		115					120					125			
Glu	Asp	Lys	Leu	Asn	His	His	Gln	Arg	Ile	Gly	Leu	Lys	Tyr	Phe	Gly
		130				135					140				
Asp	Phe	Glu	Lys	Arg	Ile	Pro	Arg	Glu	Glu	Met	Leu	Gln	Met	Gln	Asp
145					150					155					160
Ile	Val	Leu	Asn	Glu	Val	Lys	Lys	Val	Asp	Ser	Glu	Tyr	Ile	Ala	Thr
				165					170					175	
Val	Cys	Gly	Ser	Phe	Arg	Arg	Gly	Ala	Glu	Ser	Ser	Gly	Asp	Met	Asp
			180					185					190		
Val	Leu	Leu	Thr	His	Pro	Ser	Phe	Thr	Ser	Glu	Ser	Thr	Lys	Gln	Pro
		195					200					205			
Lys	Leu	Leu	His	Gln	Val	Val	Glu	Gln	Leu	Gln	Lys	Val	His	Phe	Ile
		210					215				220				

Thr	Asp	Thr	Leu	Ser	Lys	Gly	Glu	Thr	Lys	Phe	Met	Gly	Val	Cys	Gln
225					230					235					240
Leu	Pro	Ser	Lys	Asn	Asp	Glu	Lys	Glu	Tyr	Pro	His	Arg	Arg	Ile	Asp
				245					250					255	
Ile	Arg	Leu	Ile	Pro	Lys	Asp	Gln	Tyr	Tyr	Cys	Gly	Val	Leu	Tyr	Phe
			260					265					270		
Thr	Gly	Ser	Asp	Ile	Phe	Asn	Lys	Asn	Met	Arg	Ala	His	Ala	Leu	Glu
		275					280					285			
Lys	Gly	Phe	Thr	Ile	Asn	Glu	Tyr	Thr	Ile	Arg	Pro	Leu	Gly	Val	Thr
	290					295				300					
Gly	Val	Ala	Gly	Glu	Pro	Leu	Pro	Val	Asp	Ser	Glu	Lys	Asp	Ile	Phe
305					310				315					320	
Asp	Tyr	Ile	Gln	Trp	Lys	Tyr	Arg	Glu	Pro	Lys	Asp	Arg	Ser	Glu	
			325					330						335	

<210> 184

<211> 713

<212> PRT

<213> Homo sapiens

<220>

<223> >POLH|ENSG00000170734|ENST00000372236|2142

<400> 184

Met	Ala	Thr	Gly	Gln	Asp	Arg	Val	Val	Ala	Leu	Val	Asp	Met	Asp	Cys
1				5					10					15	
Phe	Phe	Val	Gln	Val	Glu	Gln	Arg	Gln	Asn	Pro	His	Leu	Arg	Asn	Lys
			20					25					30		
Pro	Cys	Ala	Val	Val	Gln	Tyr	Lys	Ser	Trp	Lys	Gly	Gly	Gly	Ile	Ile
		35					40					45			
Ala	Val	Ser	Tyr	Glu	Ala	Arg	Ala	Phe	Gly	Val	Thr	Arg	Ser	Met	Trp
	50					55					60				
Ala	Asp	Asp	Ala	Lys	Lys	Leu	Cys	Pro	Asp	Leu	Leu	Leu	Ala	Gln	Val
65					70					75				80	
Arg	Glu	Ser	Arg	Gly	Lys	Ala	Asn	Leu	Thr	Lys	Tyr	Arg	Glu	Ala	Ser
			85						90				95		
Val	Glu	Val	Met	Glu	Ile	Met	Ser	Arg	Phe	Ala	Val	Ile	Glu	Arg	Ala
		100					105					110			
Ser	Ile	Asp	Glu	Ala	Tyr	Val	Asp	Leu	Thr	Ser	Ala	Val	Gln	Glu	Arg
		115					120				125				
Leu	Gln	Lys	Leu	Gln	Gly	Gln	Pro	Ile	Ser	Ala	Asp	Leu	Leu	Pro	Ser
	130					135					140				
Thr	Tyr	Ile	Glu	Gly	Leu	Pro	Gln	Gly	Pro	Thr	Thr	Ala	Glu	Glu	Thr
145					150					155				160	
Val	Gln	Lys	Glu	Gly	Met	Arg	Lys	Gln	Gly	Leu	Phe	Gln	Trp	Leu	Asp
				165					170					175	
Ser	Leu	Gln	Ile	Asp	Asn	Leu	Thr	Ser	Pro	Asp	Leu	Gln	Leu	Thr	Val
			180					185					190		
Gly	Ala	Val	Ile	Val	Glu	Glu	Met	Arg	Ala	Ala	Ile	Glu	Arg	Glu	Thr
	195						200					205			
Gly	Phe	Gln	Cys	Ser	Ala	Gly	Ile	Ser	His	Asn	Lys	Val	Leu	Ala	Lys
	210					215					220				
Leu	Ala	Cys	Gly	Leu	Asn	Lys	Pro	Asn	Arg	Gln	Thr	Leu	Val	Ser	His

225					230					235				240	
Gly	Ser	Val	Pro	Gln	Leu	Phe	Ser	Gln	Met	Pro	Ile	Arg	Lys	Ile	Arg
				245					250					255	
Ser	Leu	Gly	Gly	Lys	Leu	Gly	Ala	Ser	Val	Ile	Glu	Ile	Leu	Gly	Ile
			260					265					270		
Glu	Tyr	Met	Gly	Glu	Leu	Thr	Gln	Phe	Thr	Glu	Ser	Gln	Leu	Gln	Ser
		275					280					285			
His	Phe	Gly	Glu	Lys	Asn	Gly	Ser	Trp	Leu	Tyr	Ala	Met	Cys	Arg	Gly
	290					295					300				
Ile	Glu	His	Asp	Pro	Val	Lys	Pro	Arg	Gln	Leu	Pro	Lys	Thr	Ile	Gly
305					310				315					320	
Cys	Ser	Lys	Asn	Phe	Pro	Gly	Lys	Thr	Ala	Leu	Ala	Thr	Arg	Glu	Gln
			325						330					335	
Val	Gln	Trp	Trp	Leu	Leu	Gln	Leu	Ala	Gln	Glu	Leu	Glu	Glu	Arg	Leu
			340					345					350		
Thr	Lys	Asp	Arg	Asn	Asp	Asn	Asp	Arg	Val	Ala	Thr	Gln	Leu	Val	Val
		355					360					365			
Ser	Ile	Arg	Val	Gln	Gly	Asp	Lys	Arg	Leu	Ser	Ser	Leu	Arg	Arg	Cys
	370					375					380				
Cys	Ala	Leu	Thr	Arg	Tyr	Asp	Ala	His	Lys	Met	Ser	His	Asp	Ala	Phe
385					390					395					400
Thr	Val	Ile	Lys	Asn	Cys	Asn	Thr	Ser	Gly	Ile	Gln	Thr	Glu	Trp	Ser
			405						410					415	
Pro	Pro	Leu	Thr	Met	Leu	Phe	Leu	Cys	Ala	Thr	Lys	Phe	Ser	Ala	Ser
			420					425					430		
Ala	Pro	Ser	Ser	Ser	Thr	Asp	Ile	Thr	Ser	Phe	Leu	Ser	Ser	Asp	Pro
		435				440						445			
Ser	Ser	Leu	Pro	Lys	Val	Pro	Val	Thr	Ser	Ser	Glu	Ala	Lys	Thr	Gln
	450					455					460				
Gly	Ser	Gly	Pro	Ala	Val	Thr	Ala	Thr	Lys	Lys	Ala	Thr	Thr	Ser	Leu
465					470						475				480
Glu	Ser	Phe	Phe	Gln	Lys	Ala	Ala	Glu	Arg	Gln	Lys	Val	Lys	Glu	Ala
			485						490					495	
Ser	Leu	Ser	Ser	Leu	Thr	Ala	Pro	Thr	Gln	Ala	Pro	Met	Ser	Asn	Ser
			500					505					510		
Pro	Ser	Lys	Pro	Ser	Leu	Pro	Phe	Gln	Thr	Ser	Gln	Ser	Thr	Gly	Thr
		515					520					525			
Glu	Pro	Phe	Phe	Lys	Gln	Lys	Ser	Leu	Leu	Leu	Lys	Gln	Lys	Gln	Leu
	530					535					540				
Asn	Asn	Ser	Ser	Val	Ser	Ser	Pro	Gln	Gln	Asn	Pro	Trp	Ser	Asn	Cys
545					550					555					560
Lys	Ala	Leu	Pro	Asn	Ser	Leu	Pro	Thr	Glu	Tyr	Pro	Gly	Cys	Val	Pro
			565						570					575	
Val	Cys	Glu	Gly	Val	Ser	Lys	Leu	Glu	Glu	Ser	Ser	Lys	Ala	Thr	Pro
			580					585					590		
Ala	Glu	Met	Asp	Leu	Ala	His	Asn	Ser	Gln	Ser	Met	His	Ala	Ser	Ser
		595					600					605			
Ala	Ser	Lys	Ser	Val	Leu	Glu	Val	Thr	Gln	Lys	Ala	Thr	Pro	Asn	Pro
	610					615					620				
Ser	Leu	Leu	Ala	Ala	Glu	Asp	Gln	Val	Pro	Cys	Glu	Lys	Cys	Gly	Ser
625					630					635					640
Leu	Val	Pro	Val	Trp	Asp	Met	Pro	Glu	His	Met	Asp	Tyr	His	Phe	Ala
			645						650					655	
Leu	Glu	Leu	Gln	Lys	Ser	Phe	Leu	Gln	Pro	His	Ser	Ser	Asn	Pro	Gln

			660					665					670				
Val	Val	Ser	Ala	Val	Ser	His	Gln	Gly	Lys	Arg	Asn	Pro	Lys	Ser	Pro		
		675					680					685					
Leu	Ala	Cys	Thr	Asn	Lys	Arg	Pro	Arg	Pro	Glu	Gly	Met	Gln	Thr	Leu		
	690					695					700						
Glu	Ser	Phe	Phe	Lys	Pro	Leu	Thr	His									
705					710												

<210> 185

<211> 575

<212> PRT

<213> Homo sapiens

<220>

<223> >POLL|ENSG00000166169|ENST00000370162|1728

<400> 185

Met	Asp	Pro	Arg	Gly	Ile	Leu	Lys	Ala	Phe	Pro	Lys	Arg	Gln	Lys	Ile		
1				5				10					15				
His	Ala	Asp	Ala	Ser	Ser	Lys	Val	Leu	Ala	Lys	Ile	Pro	Arg	Arg	Glu		
		20						25					30				
Glu	Gly	Glu	Glu	Ala	Glu	Glu	Trp	Leu	Ser	Ser	Leu	Arg	Ala	His	Val		
	35					40					45						
Val	Arg	Thr	Gly	Ile	Gly	Arg	Ala	Arg	Ala	Glu	Leu	Phe	Glu	Lys	Gln		
	50				55					60							
Ile	Val	Gln	His	Gly	Gly	Gln	Leu	Cys	Pro	Ala	Gln	Gly	Pro	Gly	Val		
65				70					75						80		
Thr	His	Ile	Val	Val	Asp	Glu	Gly	Met	Asp	Tyr	Glu	Arg	Ala	Leu	Arg		
			85					90						95			
Leu	Leu	Arg	Leu	Pro	Gln	Leu	Pro	Pro	Gly	Ala	Gln	Leu	Val	Lys	Ser		
		100						105					110				
Ala	Trp	Leu	Ser	Leu	Cys	Leu	Gln	Glu	Arg	Arg	Leu	Val	Asp	Val	Ala		
	115						120					125					
Gly	Phe	Ser	Ile	Phe	Ile	Pro	Ser	Arg	Tyr	Leu	Asp	His	Pro	Gln	Pro		
	130					135					140						
Ser	Lys	Ala	Glu	Gln	Asp	Ala	Ser	Ile	Pro	Pro	Gly	Thr	His	Glu	Ala		
145				150					155						160		
Leu	Leu	Gln	Thr	Ala	Leu	Ser	Pro	Pro	Pro	Pro	Pro	Thr	Arg	Pro	Val		
			165					170						175			
Ser	Pro	Pro	Gln	Lys	Ala	Lys	Glu	Ala	Pro	Asn	Thr	Gln	Ala	Gln	Pro		
		180						185					190				
Ile	Ser	Asp	Asp	Glu	Ala	Ser	Asp	Gly	Glu	Glu	Thr	Gln	Val	Ser	Ala		
	195						200					205					
Ala	Asp	Leu	Glu	Ala	Leu	Ile	Ser	Gly	His	Tyr	Pro	Thr	Ser	Leu	Glu		
	210					215					220						
Gly	Asp	Cys	Glu	Pro	Ser	Pro	Ala	Pro	Ala	Val	Leu	Asp	Lys	Trp	Val		
225				230						235					240		
Cys	Ala	Gln	Pro	Ser	Ser	Gln	Lys	Ala	Thr	Asn	His	Asn	Leu	His	Ile		
			245						250					255			
Thr	Glu	Lys	Leu	Glu	Val	Leu	Ala	Lys	Ala	Tyr	Ser	Val	Gln	Gly	Asp		
	260						265						270				
Lys	Trp	Arg	Ala	Leu	Gly	Tyr	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Lys	Ser		
	275						280					285					

Phe	His	Lys	Pro	Val	Thr	Ser	Tyr	Gln	Glu	Ala	Cys	Ser	Ile	Pro	Gly
290						295					300				
Ile	Gly	Lys	Arg	Met	Ala	Glu	Lys	Ile	Ile	Glu	Ile	Leu	Glu	Ser	Gly
305					310					315					320
His	Leu	Arg	Lys	Leu	Asp	His	Ile	Ser	Glu	Ser	Val	Pro	Val	Leu	Glu
				325					330					335	
Leu	Phe	Ser	Asn	Ile	Trp	Gly	Ala	Gly	Thr	Lys	Thr	Ala	Gln	Met	Trp
			340					345					350		
Tyr	Gln	Gln	Gly	Phe	Arg	Ser	Leu	Glu	Asp	Ile	Arg	Ser	Gln	Ala	Ser
	355					360					365				
Leu	Thr	Thr	Gln	Gln	Ala	Ile	Gly	Leu	Lys	His	Tyr	Ser	Asp	Phe	Leu
370						375					380				
Glu	Arg	Met	Pro	Arg	Glu	Glu	Ala	Thr	Glu	Ile	Glu	Gln	Thr	Val	Gln
385					390					395					400
Lys	Ala	Ala	Gln	Ala	Phe	Asn	Ser	Gly	Leu	Leu	Cys	Val	Ala	Cys	Gly
			405					410						415	
Ser	Tyr	Arg	Arg	Gly	Lys	Ala	Thr	Cys	Gly	Asp	Val	Asp	Val	Leu	Ile
		420						425				430			
Thr	His	Pro	Asp	Gly	Arg	Ser	His	Arg	Gly	Ile	Phe	Ser	Arg	Leu	Leu
		435					440					445			
Asp	Ser	Leu	Arg	Gln	Glu	Gly	Phe	Leu	Thr	Asp	Asp	Leu	Val	Ser	Gln
450						455					460				
Glu	Glu	Asn	Gly	Gln	Gln	Gln	Lys	Tyr	Leu	Gly	Val	Cys	Arg	Leu	Pro
465					470					475					480
Gly	Pro	Gly	Arg	Arg	His	Arg	Arg	Leu	Asp	Ile	Ile	Val	Val	Pro	Tyr
			485						490					495	
Ser	Glu	Phe	Ala	Cys	Ala	Leu	Leu	Tyr	Phe	Thr	Gly	Ser	Ala	His	Phe
			500					505					510		
Asn	Arg	Ser	Met	Arg	Ala	Leu	Ala	Lys	Thr	Lys	Gly	Met	Ser	Leu	Ser
		515					520					525			
Glu	His	Ala	Leu	Ser	Thr	Ala	Val	Val	Arg	Asn	Thr	His	Gly	Cys	Lys
	530					535					540				
Val	Gly	Pro	Gly	Arg	Val	Leu	Pro	Thr	Pro	Thr	Glu	Lys	Asp	Val	Phe
545					550					555					560
Arg	Leu	Leu	Gly	Leu	Pro	Tyr	Arg	Glu	Pro	Ala	Glu	Arg	Asp	Trp	
				565					570					575	

<210> 186

<211> 900

<212> PRT

<213> Homo sapiens

<220>

<223> >POLN|ENSG00000130997|ENST00000511885|2703

<400> 186

Met	Glu	Asn	Tyr	Glu	Ala	Leu	Val	Gly	Phe	Asp	Leu	Cys	Asn	Thr	Pro
1				5					10					15	
Leu	Ser	Ser	Val	Ala	Gln	Lys	Ile	Met	Ser	Ala	Met	His	Ser	Gly	Asp
			20					25					30		
Leu	Val	Asp	Ser	Lys	Thr	Trp	Gly	Lys	Ser	Thr	Glu	Thr	Met	Glu	Val
	35						40					45			
Ile	Asn	Lys	Ser	Ser	Val	Lys	Tyr	Ser	Val	Gln	Leu	Glu	Asp	Arg	Lys

Page 535

Page 536

<210> 187
 <211> 2590
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >POLQ|ENSG00000051341|ENST00000264233|7773

<400> 187
 Met Asn Leu Leu Arg Arg Ser Gly Lys Arg Arg Arg Ser Glu Ser Gly
 1 5 10 15
 Ser Asp Ser Phe Ser Gly Ser Gly Gly Asp Ser Ser Ala Ser Pro Gln
 20 25 30
 Phe Leu Ser Gly Ser Val Leu Ser Pro Pro Pro Gly Leu Gly Arg Cys
 35 40 45
 Leu Lys Ala Ala Ala Ala Gly Glu Cys Lys Pro Thr Val Pro Asp Tyr
 50 55 60
 Glu Arg Asp Lys Leu Leu Leu Ala Asn Trp Gly Leu Pro Lys Ala Val
 65 70 75 80
 Leu Glu Lys Tyr His Ser Phe Gly Val Lys Lys Met Phe Glu Trp Gln
 85 90 95
 Ala Glu Cys Leu Leu Leu Gly Gln Val Leu Glu Gly Lys Asn Leu Val
 100 105 110
 Tyr Ser Ala Pro Thr Ser Ala Gly Lys Thr Leu Val Ala Glu Leu Leu
 115 120 125
 Ile Leu Lys Arg Val Leu Glu Met Arg Lys Lys Ala Leu Phe Ile Leu
 130 135 140
 Pro Phe Val Ser Val Ala Lys Glu Lys Lys Tyr Tyr Leu Gln Ser Leu
 145 150 155 160
 Phe Gln Glu Val Gly Ile Lys Val Asp Gly Tyr Met Gly Ser Thr Ser
 165 170 175
 Pro Ser Arg His Phe Ser Ser Leu Asp Ile Ala Val Cys Thr Ile Glu
 180 185 190
 Arg Ala Asn Gly Leu Ile Asn Arg Leu Ile Glu Glu Asn Lys Met Asp
 195 200 205
 Leu Leu Gly Met Val Val Val Asp Glu Leu His Met Leu Gly Asp Ser
 210 215 220
 His Arg Gly Tyr Leu Leu Glu Leu Leu Leu Thr Lys Ile Cys Tyr Ile
 225 230 235 240
 Thr Arg Lys Ser Ala Ser Cys Gln Ala Asp Leu Ala Ser Ser Leu Ser
 245 250 255
 Asn Ala Val Gln Ile Val Gly Met Ser Ala Thr Leu Pro Asn Leu Glu
 260 265 270
 Leu Val Ala Ser Trp Leu Asn Ala Glu Leu Tyr His Thr Asp Phe Arg
 275 280 285
 Pro Val Pro Leu Leu Glu Ser Val Lys Val Gly Asn Ser Ile Tyr Asp
 290 295 300
 Ser Ser Met Lys Leu Val Arg Glu Phe Glu Pro Met Leu Gln Val Lys
 305 310 315 320
 Gly Asp Glu Asp His Val Val Ser Leu Cys Tyr Glu Thr Ile Cys Asp
 325 330 335
 Asn His Ser Val Leu Leu Phe Cys Pro Ser Lys Lys Trp Cys Glu Lys
 340 345 350

Leu	Ala	Asp	Ile	Ile	Ala	Arg	Glu	Phe	Tyr	Asn	Leu	His	His	Gln	Ala		
		355					360					365					
Glu	Gly	Leu	Val	Lys	Pro	Ser	Glu	Cys	Pro	Pro	Val	Ile	Leu	Glu	Gln		
	370					375					380						
Lys	Glu	Leu	Leu	Glu	Val	Met	Asp	Gln	Leu	Arg	Arg	Leu	Pro	Ser	Gly		
385					390					395					400		
Leu	Asp	Ser	Val	Leu	Gln	Lys	Thr	Val	Pro	Trp	Gly	Val	Ala	Phe	His		
			405						410					415			
His	Ala	Gly	Leu	Thr	Phe	Glu	Glu	Arg	Asp	Ile	Ile	Glu	Gly	Ala	Phe		
		420						425					430				
Arg	Gln	Gly	Leu	Ile	Arg	Val	Leu	Ala	Ala	Thr	Ser	Thr	Leu	Ser	Ser		
		435					440					445					
Gly	Val	Asn	Leu	Pro	Ala	Arg	Arg	Val	Ile	Ile	Arg	Thr	Pro	Ile	Phe		
	450					455					460						
Gly	Gly	Arg	Pro	Leu	Asp	Ile	Leu	Thr	Tyr	Lys	Gln	Met	Val	Gly	Arg		
465					470					475					480		
Ala	Gly	Arg	Lys	Gly	Val	Asp	Thr	Val	Gly	Glu	Ser	Ile	Leu	Ile	Cys		
			485						490					495			
Lys	Asn	Ser	Glu	Lys	Ser	Lys	Gly	Ile	Ala	Leu	Leu	Gln	Gly	Ser	Leu		
			500					505					510				
Lys	Pro	Val	Arg	Ser	Cys	Leu	Gln	Arg	Arg	Glu	Gly	Glu	Glu	Val	Thr		
		515					520					525					
Gly	Ser	Met	Ile	Arg	Ala	Ile	Leu	Glu	Ile	Ile	Val	Gly	Gly	Val	Ala		
	530					535					540						
Ser	Thr	Ser	Gln	Asp	Met	His	Thr	Tyr	Ala	Ala	Cys	Thr	Phe	Leu	Ala		
545					550				555						560		
Ala	Ser	Met	Lys	Glu	Gly	Lys	Gln	Gly	Ile	Gln	Arg	Asn	Gln	Glu	Ser		
			565						570					575			
Val	Gln	Leu	Gly	Ala	Ile	Glu	Ala	Cys	Val	Met	Trp	Leu	Leu	Glu	Asn		
		580						585					590				
Glu	Phe	Ile	Gln	Ser	Thr	Glu	Ala	Ser	Asp	Gly	Thr	Glu	Gly	Lys	Val		
	595						600					605					
Tyr	His	Pro	Thr	His	Leu	Gly	Ser	Ala	Thr	Leu	Ser	Ser	Ser	Leu	Ser		
	610					615					620						
Pro	Ala	Asp	Thr	Leu	Asp	Ile	Phe	Ala	Asp	Leu	Gln	Arg	Ala	Met	Lys		
625					630					635					640		
Gly	Phe	Val	Leu	Glu	Asn	Asp	Leu	His	Ile	Leu	Tyr	Leu	Val	Thr	Pro		
			645						650					655			
Met	Phe	Glu	Asp	Trp	Thr	Thr	Ile	Asp	Trp	Tyr	Arg	Phe	Phe	Cys	Leu		
			660					665					670				
Trp	Glu	Lys	Leu	Pro	Thr	Ser	Met	Lys	Arg	Val	Ala	Glu	Leu	Val	Gly		
		675					680					685					
Val	Glu	Glu	Gly	Phe	Leu	Ala	Arg	Cys	Val	Lys	Gly	Lys	Val	Val	Ala		
	690					695					700						
Arg	Thr	Glu	Arg	Gln	His	Arg	Gln	Met	Ala	Ile	His	Lys	Arg	Phe	Phe		
705					710					715					720		
Thr	Ser	Leu	Val	Leu	Leu	Asp	Leu	Ile	Ser	Glu	Val	Pro	Leu	Arg	Glu		
			725						730					735			
Ile	Asn	Gln	Lys	Tyr	Gly	Cys	Asn	Arg	Gly	Gln	Ile	Gln	Ser	Leu	Gln		
		740						745					750				
Gln	Ser	Ala	Ala	Val	Tyr	Ala	Gly	Met	Ile	Thr	Val	Phe	Ser	Asn	Arg		
	755						760					765					
Leu	Gly	Trp	His	Asn	Met	Glu	Leu	Leu	Leu	Ser	Gln	Phe	Gln	Lys	Arg		
	770					775					780						

Leu	Thr	Phe	Gly	Ile	Gln	Arg	Glu	Leu	Cys	Asp	Leu	Val	Arg	Val	Ser	785	790	795	800
Leu	Leu	Asn	Ala	Gln	Arg	Ala	Arg	Val	Leu	Tyr	Ala	Ser	Gly	Phe	His	805	810		815
Thr	Val	Ala	Asp	Leu	Ala	Arg	Ala	Asn	Ile	Val	Glu	Val	Glu	Val	Ile	820	825		830
Leu	Lys	Asn	Ala	Val	Pro	Phe	Lys	Ser	Ala	Arg	Lys	Ala	Val	Asp	Glu	835	840		845
Glu	Glu	Glu	Ala	Val	Glu	Glu	Arg	Arg	Asn	Met	Arg	Thr	Ile	Trp	Val	850	855		860
Thr	Gly	Arg	Lys	Gly	Leu	Thr	Glu	Arg	Glu	Ala	Ala	Ala	Leu	Ile	Val	865	870		875
Glu	Glu	Ala	Arg	Met	Ile	Leu	Gln	Gln	Asp	Leu	Val	Glu	Met	Gly	Val	885	890		895
Gln	Trp	Asn	Pro	Cys	Ala	Leu	Leu	His	Ser	Ser	Thr	Cys	Ser	Leu	Thr	900	905		910
His	Ser	Glu	Ser	Glu	Val	Lys	Glu	His	Thr	Phe	Ile	Ser	Gln	Thr	Lys	915	920		925
Ser	Ser	Tyr	Lys	Lys	Leu	Thr	Ser	Lys	Asn	Lys	Ser	Asn	Thr	Ile	Phe	930	935		940
Ser	Asp	Ser	Tyr	Ile	Lys	His	Ser	Pro	Asn	Ile	Val	Gln	Asp	Leu	Asn	945	950		955
Lys	Ser	Arg	Glu	His	Thr	Ser	Ser	Phe	Asn	Cys	Asn	Phe	Gln	Asn	Gly	965	970		975
Asn	Gln	Glu	His	Gln	Thr	Cys	Ser	Ile	Phe	Arg	Ala	Arg	Lys	Arg	Ala	980	985		990
Ser	Leu	Asp	Ile	Asn	Lys	Glu	Lys	Pro	Gly	Ala	Ser	Gln	Asn	Glu	Gly	995	1000		1005
Lys	Thr	Ser	Asp	Lys	Lys	Val	Val	Gln	Thr	Phe	Ser	Gln	Lys	Thr	Lys	1010	1015		1020
Lys	Ala	Pro	Leu	Asn	Phe	Asn	Ser	Glu	Lys	Met	Ser	Arg	Ser	Phe	Arg	1025	1030		1035
Ser	Trp	Lys	Arg	Arg	Lys	His	Leu	Lys	Arg	Ser	Arg	Asp	Ser	Ser	Pro	1045	1050		1055
Leu	Lys	Asp	Ser	Gly	Ala	Cys	Arg	Ile	His	Leu	Gln	Gly	Gln	Thr	Leu	1060	1065		1070
Ser	Asn	Pro	Ser	Leu	Cys	Glu	Asp	Pro	Phe	Thr	Leu	Asp	Glu	Lys	Lys	1075	1080		1085
Thr	Glu	Phe	Arg	Asn	Ser	Gly	Pro	Phe	Ala	Lys	Asn	Val	Ser	Leu	Ser	1090	1095		1100
Gly	Lys	Glu	Lys	Asp	Asn	Lys	Thr	Ser	Phe	Pro	Leu	Gln	Ile	Lys	Gln	1105	1110		1115
Asn	Cys	Ser	Trp	Asn	Ile	Thr	Leu	Thr	Asn	Asp	Asn	Phe	Val	Glu	His	1125	1130		1135
Ile	Val	Thr	Gly	Ser	Gln	Ser	Lys	Asn	Val	Thr	Cys	Gln	Ala	Thr	Ser	1140	1145		1150
Val	Val	Ser	Glu	Lys	Gly	Arg	Gly	Val	Ala	Val	Glu	Ala	Glu	Lys	Ile	1155	1160		1165
Asn	Glu	Val	Leu	Ile	Gln	Asn	Gly	Ser	Lys	Asn	Gln	Asn	Val	Tyr	Met	1170	1175		1180
Lys	His	His	Asp	Ile	His	Pro	Ile	Asn	Gln	Tyr	Leu	Arg	Lys	Gln	Ser	1185	1190		1195
His	Glu	Gln	Thr	Ser	Thr	Ile	Thr	Lys	Gln	Lys	Asn	Ile	Ile	Glu	Arg	1205	1210		1215

Gln Met Pro Cys Glu Ala Val Ser Ser Tyr Ile Asn Arg Asp Ser Asn
 1220 1225 1230
 Val Thr Ile Asn Cys Glu Arg Ile Lys Leu Asn Thr Glu Glu Asn Lys
 1235 1240 1245
 Pro Ser His Phe Gln Ala Leu Gly Asp Asp Ile Ser Arg Thr Val Ile
 1250 1255 1260
 Pro Ser Glu Val Leu Pro Ser Ala Gly Ala Phe Ser Lys Ser Glu Gly
 1265 1270 1275 1280
 Gln His Glu Asn Phe Leu Asn Ile Ser Arg Leu Gln Glu Lys Thr Gly
 1285 1290 1295
 Thr Tyr Thr Thr Asn Lys Thr Lys Asn Asn His Val Ser Asp Leu Gly
 1300 1305 1310
 Leu Val Leu Cys Asp Phe Glu Asp Ser Phe Tyr Leu Asp Thr Gln Ser
 1315 1320 1325
 Glu Lys Ile Ile Gln Gln Met Ala Thr Glu Asn Ala Lys Leu Gly Ala
 1330 1335 1340
 Lys Asp Thr Asn Leu Ala Ala Gly Ile Met Gln Lys Ser Leu Val Gln
 1345 1350 1355 1360
 Gln Asn Ser Met Asn Ser Phe Gln Lys Glu Cys His Ile Pro Phe Pro
 1365 1370 1375
 Ala Glu Gln His Pro Leu Gly Ala Thr Lys Ile Asp His Leu Asp Leu
 1380 1385 1390
 Lys Thr Val Gly Thr Met Lys Gln Ser Ser Asp Ser His Gly Val Asp
 1395 1400 1405
 Ile Leu Thr Pro Glu Ser Pro Ile Phe His Ser Pro Ile Leu Leu Glu
 1410 1415 1420
 Glu Asn Gly Leu Phe Leu Lys Lys Asn Glu Val Ser Val Thr Asp Ser
 1425 1430 1435 1440
 Gln Leu Asn Ser Phe Leu Gln Gly Tyr Gln Thr Gln Glu Thr Val Lys
 1445 1450 1455
 Pro Val Ile Leu Leu Ile Pro Gln Lys Arg Thr Pro Thr Gly Val Glu
 1460 1465 1470
 Gly Glu Cys Leu Pro Val Pro Glu Thr Ser Leu Asn Met Ser Asp Ser
 1475 1480 1485
 Leu Leu Phe Asp Ser Phe Ser Asp Asp Tyr Leu Val Lys Glu Gln Leu
 1490 1495 1500
 Pro Asp Met Gln Met Lys Glu Pro Leu Pro Ser Glu Val Thr Ser Asn
 1505 1510 1515 1520
 His Phe Ser Asp Ser Leu Cys Leu Gln Glu Asp Leu Ile Lys Lys Ser
 1525 1530 1535
 Asn Val Asn Glu Asn Gln Asp Thr His Gln Gln Leu Thr Cys Ser Asn
 1540 1545 1550
 Asp Glu Ser Ile Ile Phe Ser Glu Met Asp Ser Val Gln Met Val Glu
 1555 1560 1565
 Ala Leu Asp Asn Val Asp Ile Phe Pro Val Gln Glu Lys Asn His Thr
 1570 1575 1580
 Val Val Ser Pro Arg Ala Leu Glu Leu Ser Asp Pro Val Leu Asp Glu
 1585 1590 1595 1600
 His His Gln Gly Asp Gln Asp Gly Gly Asp Gln Asp Glu Arg Ala Glu
 1605 1610 1615
 Lys Ser Lys Leu Thr Gly Thr Arg Gln Asn His Ser Phe Ile Trp Ser
 1620 1625 1630
 Gly Ala Ser Phe Asp Leu Ser Pro Gly Leu Gln Arg Ile Leu Asp Lys
 1635 1640 1645

Val	Ser	Ser	Pro	Leu	Glu	Asn	Glu	Lys	Leu	Lys	Ser	Met	Thr	Ile	Asn	
1650						1655					1660					
Phe	Ser	Ser	Leu	Asn	Arg	Lys	Asn	Thr	Glu	Leu	Asn	Glu	Glu	Gln	Glu	
1665				1670					1675						1680	
Val	Ile	Ser	Asn	Leu	Glu	Thr	Lys	Gln	Val	Gln	Gly	Ile	Ser	Phe	Ser	
			1685						1690					1695		
Ser	Asn	Asn	Glu	Val	Lys	Ser	Lys	Ile	Glu	Met	Leu	Glu	Asn	Asn	Ala	
			1700					1705					1710			
Asn	His	Asp	Glu	Thr	Ser	Ser	Leu	Leu	Pro	Arg	Lys	Glu	Ser	Asn	Ile	
	1715						1720					1725				
Val	Asp	Asp	Asn	Gly	Leu	Ile	Pro	Pro	Thr	Pro	Ile	Pro	Thr	Ser	Ala	
1730						1735					1740					
Ser	Lys	Leu	Thr	Phe	Pro	Gly	Ile	Leu	Glu	Thr	Pro	Val	Asn	Pro	Trp	
1745				1750					1755						1760	
Lys	Thr	Asn	Asn	Val	Leu	Gln	Pro	Gly	Glu	Ser	Tyr	Leu	Phe	Gly	Ser	
			1765						1770					1775		
Pro	Ser	Asp	Ile	Lys	Asn	His	Asp	Leu	Ser	Pro	Gly	Ser	Arg	Asn	Gly	
		1780						1785					1790			
Phe	Lys	Asp	Asn	Ser	Pro	Ile	Ser	Asp	Thr	Ser	Phe	Ser	Leu	Gln	Leu	
	1795						1800					1805				
Ser	Gln	Asp	Gly	Leu	Gln	Leu	Thr	Pro	Ala	Ser	Ser	Ser	Ser	Glu	Ser	
	1810					1815					1820					
Leu	Ser	Ile	Ile	Asp	Val	Ala	Ser	Asp	Gln	Asn	Leu	Phe	Gln	Thr	Phe	
1825				1830					1835						1840	
Ile	Lys	Glu	Trp	Arg	Cys	Lys	Lys	Arg	Phe	Ser	Ile	Ser	Leu	Ala	Cys	
			1845						1850					1855		
Glu	Lys	Ile	Arg	Ser	Leu	Thr	Ser	Ser	Lys	Thr	Ala	Thr	Ile	Gly	Ser	
		1860						1865					1870			
Arg	Phe	Lys	Gln	Ala	Ser	Ser	Pro	Gln	Glu	Ile	Pro	Ile	Arg	Asp	Asp	
	1875						1880					1885				
Gly	Phe	Pro	Ile	Lys	Gly	Cys	Asp	Asp	Thr	Leu	Val	Val	Gly	Leu	Ala	
	1890					1895					1900					
Val	Cys	Trp	Gly	Gly	Arg	Asp	Ala	Tyr	Tyr	Phe	Ser	Leu	Gln	Lys	Glu	
1905					1910					1915					1920	
Gln	Lys	His	Ser	Glu	Ile	Ser	Ala	Ser	Leu	Val	Pro	Pro	Ser	Leu	Asp	
			1925						1930					1935		
Pro	Ser	Leu	Thr	Leu	Lys	Asp	Arg	Met	Trp	Tyr	Leu	Gln	Ser	Cys	Leu	
		1940						1945					1950			
Arg	Lys	Glu	Ser	Asp	Lys	Glu	Cys	Ser	Val	Val	Ile	Tyr	Asp	Phe	Ile	
	1955						1960					1965				
Gln	Ser	Tyr	Lys	Ile	Leu	Leu	Leu	Ser	Cys	Gly	Ile	Ser	Leu	Glu	Gln	
	1970					1975					1980					
Ser	Tyr	Glu	Asp	Pro	Lys	Val	Ala	Cys	Trp	Leu	Leu	Asp	Pro	Asp	Ser	
1985					1990					1995					2000	
Gln	Glu	Pro	Thr	Leu	His	Ser	Ile	Val	Thr	Ser	Phe	Leu	Pro	His	Glu	
			2005						2010					2015		
Leu	Pro	Leu	Leu	Glu	Gly	Met	Glu	Thr	Ser	Gln	Gly	Ile	Gln	Ser	Leu	
		2020						2025					2030			
Gly	Leu	Asn	Ala	Gly	Ser	Glu	His	Ser	Gly	Arg	Tyr	Arg	Ala	Ser	Val	
		2035					2040					2045				
Glu	Ser	Ile	Leu	Ile	Phe	Asn	Ser	Met	Asn	Gln	Leu	Asn	Ser	Leu	Leu	
	2050					2055					2060					
Gln	Lys	Glu	Asn	Leu	Gln	Asp	Val	Phe	Arg	Lys	Val	Glu	Met	Pro	Ser	
2065					2070					2075					2080	

Gln	Tyr	Cys	Leu	Ala	Leu	Leu	Glu	Leu	Asn	Gly	Ile	Gly	Phe	Ser	Thr		
				2085					2090					2095			
Ala	Glu	Cys	Glu	Ser	Gln	Lys	His	Ile	Met	Gln	Ala	Lys	Leu	Asp	Ala		
			2100					2105					2110				
Ile	Glu	Thr	Gln	Ala	Tyr	Gln	Leu	Ala	Gly	His	Ser	Phe	Ser	Phe	Thr		
		2115					2120					2125					
Ser	Ser	Asp	Asp	Ile	Ala	Glu	Val	Leu	Phe	Leu	Glu	Leu	Lys	Leu	Pro		
	2130					2135					2140						
Pro	Asn	Arg	Glu	Met	Lys	Asn	Gln	Gly	Ser	Lys	Lys	Thr	Leu	Gly	Ser		
2145					2150					2155					2160		
Thr	Arg	Arg	Gly	Ile	Asp	Asn	Gly	Arg	Lys	Leu	Arg	Leu	Gly	Arg	Gln		
			2165					2170						2175			
Phe	Ser	Thr	Ser	Lys	Asp	Val	Leu	Asn	Lys	Leu	Lys	Ala	Leu	His	Pro		
			2180					2185					2190				
Leu	Pro	Gly	Leu	Ile	Leu	Glu	Trp	Arg	Arg	Ile	Thr	Asn	Ala	Ile	Thr		
		2195					2200					2205					
Lys	Val	Val	Phe	Pro	Leu	Gln	Arg	Glu	Lys	Cys	Leu	Asn	Pro	Phe	Leu		
	2210					2215					2220						
Gly	Met	Glu	Arg	Ile	Tyr	Pro	Val	Ser	Gln	Ser	His	Thr	Ala	Thr	Gly		
2225					2230					2235					2240		
Arg	Ile	Thr	Phe	Thr	Glu	Pro	Asn	Ile	Gln	Asn	Val	Pro	Arg	Asp	Phe		
			2245						2250					2255			
Glu	Ile	Lys	Met	Pro	Thr	Leu	Val	Gly	Glu	Ser	Pro	Pro	Ser	Gln	Ala		
			2260					2265					2270				
Val	Gly	Lys	Gly	Leu	Leu	Pro	Met	Gly	Arg	Gly	Lys	Tyr	Lys	Lys	Gly		
	2275					2280						2285					
Phe	Ser	Val	Asn	Pro	Arg	Cys	Gln	Ala	Gln	Met	Glu	Glu	Arg	Ala	Ala		
	2290					2295					2300						
Asp	Arg	Gly	Met	Pro	Phe	Ser	Ile	Ser	Met	Arg	His	Ala	Phe	Val	Pro		
2305					2310				2315						2320		
Phe	Pro	Gly	Gly	Ser	Ile	Leu	Ala	Ala	Asp	Tyr	Ser	Gln	Leu	Glu	Leu		
			2325						2330					2335			
Arg	Ile	Leu	Ala	His	Leu	Ser	His	Asp	Arg	Arg	Leu	Ile	Gln	Val	Leu		
			2340					2345					2350				
Asn	Thr	Gly	Ala	Asp	Val	Phe	Arg	Ser	Ile	Ala	Ala	Glu	Trp	Lys	Met		
		2355					2360					2365					
Ile	Glu	Pro	Glu	Ser	Val	Gly	Asp	Asp	Leu	Arg	Gln	Gln	Ala	Lys	Gln		
	2370					2375					2380						
Ile	Cys	Tyr	Gly	Ile	Ile	Tyr	Gly	Met	Gly	Ala	Lys	Ser	Leu	Gly	Glu		
2385					2390				2395						2400		
Gln	Met	Gly	Ile	Lys	Glu	Asn	Asp	Ala	Ala	Cys	Tyr	Ile	Asp	Ser	Phe		
			2405					2410					2415				
Lys	Ser	Arg	Tyr	Thr	Gly	Ile	Asn	Gln	Phe	Met	Thr	Glu	Thr	Val	Lys		
			2420				2425					2430					
Asn	Cys	Lys	Arg	Asp	Gly	Phe	Val	Gln	Thr	Ile	Leu	Gly	Arg	Arg	Arg		
	2435						2440					2445					
Tyr	Leu	Pro	Gly	Ile	Lys	Asp	Asn	Asn	Pro	Tyr	Arg	Lys	Ala	His	Ala		
	2450					2455					2460						
Glu	Arg	Gln	Ala	Ile	Asn	Thr	Ile	Val	Gln	Gly	Ser	Ala	Ala	Asp	Ile		
2465					2470					2475					2480		
Val	Lys	Ile	Ala	Thr	Val	Asn	Ile	Gln	Lys	Gln	Leu	Glu	Thr	Phe	His		
			2485					2490						2495			
Ser	Thr	Phe	Lys	Ser	His	Gly	His	Arg	Glu	Gly	Met	Leu	Gln	Ser	Asp		
			2500					2505					2510				

Gln	Thr	Gly	Leu	Ser	Arg	Lys	Arg	Lys	Leu	Gln	Gly	Met	Phe	Cys	Pro
		2515					2520					2525			
Ile	Arg	Gly	Gly	Phe	Phe	Ile	Leu	Gln	Leu	His	Asp	Glu	Leu	Leu	Tyr
	2530					2535					2540				
Glu	Val	Ala	Glu	Glu	Asp	Val	Val	Gln	Val	Ala	Gln	Ile	Val	Lys	Asn
2545					2550					2555					2560
Glu	Met	Glu	Ser	Ala	Val	Lys	Leu	Ser	Val	Lys	Leu	Lys	Val	Lys	Val
				2565					2570					2575	
Lys	Ile	Gly	Ala	Ser	Trp	Gly	Glu	Leu	Lys	Asp	Phe	Asp	Val		
			2580					2585					2590		

<210> 188

<211> 4127

<212> PRT

<213> Homo sapiens

<220>

<223> >PRKDC|ENSG00000253729|ENST00000314191|12384

<400> 188

Met	Ala	Gly	Ser	Gly	Ala	Gly	Val	Arg	Cys	Ser	Leu	Leu	Arg	Leu	Gln
1				5					10					15	
Glu	Thr	Leu	Ser	Ala	Ala	Asp	Arg	Cys	Gly	Ala	Ala	Leu	Ala	Gly	His
		20						25					30		
Gln	Leu	Ile	Arg	Gly	Leu	Gly	Gln	Glu	Cys	Val	Leu	Ser	Ser	Ser	Pro
	35						40					45			
Ala	Val	Leu	Ala	Leu	Gln	Thr	Ser	Leu	Val	Phe	Ser	Arg	Asp	Phe	Gly
	50					55				60					
Leu	Leu	Val	Phe	Val	Arg	Lys	Ser	Leu	Asn	Ser	Ile	Glu	Phe	Arg	Glu
65					70				75						80
Cys	Arg	Glu	Glu	Ile	Leu	Lys	Phe	Leu	Cys	Ile	Phe	Leu	Glu	Lys	Met
				85					90					95	
Gly	Gln	Lys	Ile	Ala	Pro	Tyr	Ser	Val	Glu	Ile	Lys	Asn	Thr	Cys	Thr
			100					105					110		
Ser	Val	Tyr	Thr	Lys	Asp	Arg	Ala	Ala	Lys	Cys	Lys	Ile	Pro	Ala	Leu
		115					120					125			
Asp	Leu	Leu	Ile	Lys	Leu	Leu	Gln	Thr	Phe	Arg	Ser	Ser	Arg	Leu	Met
	130					135					140				
Asp	Glu	Phe	Lys	Ile	Gly	Glu	Leu	Phe	Ser	Lys	Phe	Tyr	Gly	Glu	Leu
145					150				155						160
Ala	Leu	Lys	Lys	Lys	Ile	Pro	Asp	Thr	Val	Leu	Glu	Lys	Val	Tyr	Glu
				165					170					175	
Leu	Leu	Gly	Leu	Leu	Gly	Glu	Val	His	Pro	Ser	Glu	Met	Ile	Asn	Asn
			180					185					190		
Ala	Glu	Asn	Leu	Phe	Arg	Ala	Phe	Leu	Gly	Glu	Leu	Lys	Thr	Gln	Met
		195					200					205			
Thr	Ser	Ala	Val	Arg	Glu	Pro	Lys	Leu	Pro	Val	Leu	Ala	Gly	Cys	Leu
		210				215					220				
Lys	Gly	Leu	Ser	Ser	Leu	Leu	Cys	Asn	Phe	Thr	Lys	Ser	Met	Glu	Glu
225					230				235						240
Asp	Pro	Gln	Thr	Ser	Arg	Glu	Ile	Phe	Asn	Phe	Val	Leu	Lys	Ala	Ile
				245					250					255	
Arg	Pro	Gln	Ile	Asp	Leu	Lys	Arg	Tyr	Ala	Val	Pro	Ser	Ala	Gly	Leu

690		695		700
Ala Leu Phe Val Lys	Phe Gly Lys Glu Val	Ala Val Lys Met Lys Gln		
705	710	715		720
Tyr Lys Asp Glu Leu	Leu Ala Ser Cys Leu	Thr Phe Leu Leu Ser Leu		
	725	730		735
Pro His Asn Ile Ile	Glu Leu Asp Val	Arg Ala Tyr Val Pro Ala Leu		
	740	745		750
Gln Met Ala Phe Lys	Leu Gly Leu Ser Tyr Thr	Pro Leu Ala Glu Val		
	755	760		765
Gly Leu Asn Ala Leu	Glu Glu Trp Ser Ile Tyr	Ile Asp Arg His Val		
	770	775		780
Met Gln Pro Tyr Tyr	Lys Asp Ile Leu Pro	Cys Leu Asp Gly Tyr Leu		
785	790	795		800
Lys Thr Ser Ala Leu	Ser Asp Glu Thr Lys	Asn Asn Trp Glu Val Ser		
	805	810		815
Ala Leu Ser Arg Ala	Ala Gln Lys Gly Phe	Asn Lys Val Val Leu Lys		
	820	825		830
His Leu Lys Lys Thr	Lys Asn Leu Ser Ser	Asn Glu Ala Ile Ser Leu		
	835	840		845
Glu Glu Ile Arg Ile	Arg Val Val Gln Met	Leu Gly Ser Leu Gly Gly		
	850	855		860
Gln Ile Asn Lys Asn	Leu Leu Thr Val Thr	Ser Ser Asp Glu Met Met		
865	870	875		880
Lys Ser Tyr Val Ala	Trp Asp Arg Glu Lys	Arg Leu Ser Phe Ala Val		
	885	890		895
Pro Phe Arg Glu Met	Lys Pro Val Ile Phe	Leu Asp Val Phe Leu Pro		
	900	905		910
Arg Val Thr Glu Leu	Ala Leu Thr Ala Ser	Asp Arg Gln Thr Lys Val		
	915	920		925
Ala Ala Cys Glu Leu	Leu His Ser Met Val	Met Phe Met Leu Gly Lys		
	930	935		940
Ala Thr Gln Met Pro	Glu Gly Gly Gln Gly	Ala Pro Pro Met Tyr Gln		
945	950	955		960
Leu Tyr Lys Arg Thr	Phe Pro Val Leu Leu	Arg Leu Ala Cys Asp Val		
	965	970		975
Asp Gln Val Thr Arg	Gln Leu Tyr Glu Pro	Leu Val Met Gln Leu Ile		
	980	985		990
His Trp Phe Thr Asn	Asn Lys Lys Phe Glu	Ser Gln Asp Thr Val Ala		
	995	1000		1005
Leu Leu Glu Ala Ile	Leu Asp Gly Ile Val	Asp Pro Val Asp Ser Thr		
	1010	1015		1020
Leu Arg Asp Phe Cys	Gly Arg Cys Ile Arg	Glu Phe Leu Lys Trp Ser		
1025	1030	1035		1040
Ile Lys Gln Ile Thr	Pro Gln Gln Gln Glu	Lys Ser Pro Val Asn Thr		
	1045	1050		1055
Lys Ser Leu Phe Lys	Arg Leu Tyr Ser Leu	Ala Leu His Pro Asn Ala		
	1060	1065		1070
Phe Lys Arg Leu Gly	Ala Ser Leu Ala Phe	Asn Asn Ile Tyr Arg Glu		
	1075	1080		1085
Phe Arg Glu Glu Glu	Ser Leu Val Glu Gln	Phe Val Phe Glu Ala Leu		
	1090	1095		1100
Val Ile Tyr Met Glu	Ser Leu Ala Leu Ala	His Ala Asp Glu Lys Ser		
1105	1110	1115		1120
Leu Gly Thr Ile Gln	Gln Cys Cys Asp Ala	Ile Asp His Leu Cys Arg		

1125										1130					1135				
Ile	Ile	Glu	Lys	Lys	His	Val	Ser	Leu	Asn	Lys	Ala	Lys	Lys	Arg	Arg				
1140										1145					1150				
Leu	Pro	Arg	Gly	Phe	Pro	Pro	Ser	Ala	Ser	Leu	Cys	Leu	Leu	Asp	Leu				
1155										1160					1165				
Val	Lys	Trp	Leu	Leu	Ala	His	Cys	Gly	Arg	Pro	Gln	Thr	Glu	Cys	Arg				
1170										1175					1180				
His	Lys	Ser	Ile	Glu	Leu	Phe	Tyr	Lys	Phe	Val	Pro	Leu	Leu	Pro	Gly				
1185										1190					1195				
Asn	Arg	Ser	Pro	Asn	Leu	Trp	Leu	Lys	Asp	Val	Leu	Lys	Glu	Glu	Gly				
1205										1210					1215				
Val	Ser	Phe	Leu	Ile	Asn	Thr	Phe	Glu	Gly	Gly	Gly	Cys	Gly	Gln	Pro				
1220										1225					1230				
Ser	Gly	Ile	Leu	Ala	Gln	Pro	Thr	Leu	Leu	Tyr	Arg	Gly	Pro	Phe	Ser				
1235										1240					1245				
Leu	Gln	Ala	Thr	Leu	Cys	Trp	Leu	Asp	Leu	Leu	Leu	Ala	Ala	Leu	Glu				
1250										1255					1260				
Cys	Tyr	Asn	Thr	Phe	Ile	Gly	Glu	Arg	Thr	Val	Gly	Ala	Leu	Gln	Val				
1265										1270					1275				
Leu	Gly	Thr	Glu	Ala	Gln	Ser	Ser	Leu	Leu	Lys	Ala	Val	Ala	Phe	Phe				
1285										1290					1295				
Leu	Glu	Ser	Ile	Ala	Met	His	Asp	Ile	Ile	Ala	Ala	Glu	Lys	Cys	Phe				
1300										1305					1310				
Gly	Thr	Gly	Ala	Ala	Gly	Asn	Arg	Thr	Ser	Pro	Gln	Glu	Gly	Glu	Arg				
1315										1320					1325				
Tyr	Asn	Tyr	Ser	Lys	Cys	Thr	Val	Val	Val	Arg	Ile	Met	Glu	Phe	Thr				
1330										1335					1340				
Thr	Thr	Leu	Leu	Asn	Thr	Ser	Pro	Glu	Gly	Trp	Lys	Leu	Leu	Lys	Lys				
1345										1350					1355				
Asp	Leu	Cys	Asn	Thr	His	Leu	Met	Arg	Val	Leu	Val	Gln	Thr	Leu	Cys				
1365										1370					1375				
Glu	Pro	Ala	Ser	Ile	Gly	Phe	Asn	Ile	Gly	Asp	Val	Gln	Val	Met	Ala				
1380										1385					1390				
His	Leu	Pro	Asp	Val	Cys	Val	Asn	Leu	Met	Lys	Ala	Leu	Lys	Met	Ser				
1395										1400					1405				
Pro	Tyr	Lys	Asp	Ile	Leu	Glu	Thr	His	Leu	Arg	Glu	Lys	Ile	Thr	Ala				
1410										1415					1420				
Gln	Ser	Ile	Glu	Glu	Leu	Cys	Ala	Val	Asn	Leu	Tyr	Gly	Pro	Asp	Ala				
1425										1430					1435				
Gln	Val	Asp	Arg	Ser	Arg	Leu	Ala	Ala	Val	Val	Ser	Ala	Cys	Lys	Gln				
1445										1450					1455				
Leu	His	Arg	Ala	Gly	Leu	Leu	His	Asn	Ile	Leu	Pro	Ser	Gln	Ser	Thr				
1460										1465					1470				
Asp	Leu	His	Ser	Val	Gly	Thr	Glu	Leu	Leu	Ser	Leu	Val	Tyr	Lys					
1475										1480					1485				
Gly	Ile	Ala	Pro	Gly	Asp	Glu	Arg	Gln	Cys	Leu	Pro	Ser	Leu	Asp	Leu				
1490										1495					1500				
Ser	Cys	Lys	Gln	Leu	Ala	Ser	Gly	Leu	Leu	Glu	Leu	Ala	Phe	Ala	Phe				
1505										1510					1515				
Gly	Gly	Leu	Cys	Glu	Arg	Leu	Val	Ser	Leu	Leu	Leu	Asn	Pro	Ala	Val				
1525										1530					1535				
Leu	Ser	Thr	Ala	Ser	Leu	Gly	Ser	Ser	Gln	Gly	Ser	Val	Ile	His	Phe				
1540										1545					1550				
Ser	His	Gly	Glu	Tyr	Phe	Tyr	Ser	Leu	Phe	Ser	Glu	Thr	Ile	Asn	Thr				

1555					1560					1565					
Glu	Leu	Leu	Lys	Asn	Leu	Asp	Leu	Ala	Val	Leu	Glu	Leu	Met	Gln	Ser
1570					1575					1580					
Ser	Val	Asp	Asn	Thr	Lys	Met	Val	Ser	Ala	Val	Leu	Asn	Gly	Met	Leu
1585					1590					1595					1600
Asp	Gln	Ser	Phe	Arg	Glu	Arg	Ala	Asn	Gln	Lys	His	Gln	Gly	Leu	Lys
1605					1610					1615					
Leu	Ala	Thr	Thr	Ile	Leu	Gln	His	Trp	Lys	Lys	Cys	Asp	Ser	Trp	Trp
1620					1625					1630					
Ala	Lys	Asp	Ser	Pro	Leu	Glu	Thr	Lys	Met	Ala	Val	Leu	Ala	Leu	Leu
1635					1640					1645					
Ala	Lys	Ile	Leu	Gln	Ile	Asp	Ser	Ser	Val	Ser	Phe	Asn	Thr	Ser	His
1650					1655					1660					
Gly	Ser	Phe	Pro	Glu	Val	Phe	Thr	Thr	Tyr	Ile	Ser	Leu	Leu	Ala	Asp
1665					1670					1675					1680
Thr	Lys	Leu	Asp	Leu	His	Leu	Lys	Gly	Gln	Ala	Val	Thr	Leu	Leu	Pro
1685					1690					1695					
Phe	Phe	Thr	Ser	Leu	Thr	Gly	Gly	Ser	Leu	Glu	Glu	Leu	Arg	Arg	Val
1700					1705					1710					
Leu	Glu	Gln	Leu	Ile	Val	Ala	His	Phe	Pro	Met	Gln	Ser	Arg	Glu	Phe
1715					1720					1725					
Pro	Pro	Gly	Thr	Pro	Arg	Phe	Asn	Asn	Tyr	Val	Asp	Cys	Met	Lys	Lys
1730					1735					1740					
Phe	Leu	Asp	Ala	Leu	Glu	Leu	Ser	Gln	Ser	Pro	Met	Leu	Leu	Glu	Leu
1745					1750					1755					1760
Met	Thr	Glu	Val	Leu	Cys	Arg	Glu	Gln	Gln	His	Val	Met	Glu	Glu	Leu
1765					1770					1775					
Phe	Gln	Ser	Ser	Phe	Arg	Arg	Ile	Ala	Arg	Arg	Gly	Ser	Cys	Val	Thr
1780					1785					1790					
Gln	Val	Gly	Leu	Leu	Glu	Ser	Val	Tyr	Glu	Met	Phe	Arg	Lys	Asp	Asp
1795					1800					1805					
Pro	Arg	Leu	Ser	Phe	Thr	Arg	Gln	Ser	Phe	Val	Asp	Arg	Ser	Leu	Leu
1810					1815					1820					
Thr	Leu	Leu	Trp	His	Cys	Ser	Leu	Asp	Ala	Leu	Arg	Glu	Phe	Phe	Ser
1825					1830					1835					1840
Thr	Ile	Val	Val	Asp	Ala	Ile	Asp	Val	Leu	Lys	Ser	Arg	Phe	Thr	Lys
1845					1850					1855					
Leu	Asn	Glu	Ser	Thr	Phe	Asp	Thr	Gln	Ile	Thr	Lys	Lys	Met	Gly	Tyr
1860					1865					1870					
Tyr	Lys	Ile	Leu	Asp	Val	Met	Tyr	Ser	Arg	Leu	Pro	Lys	Asp	Asp	Val
1875					1880					1885					
His	Ala	Lys	Glu	Ser	Lys	Ile	Asn	Gln	Val	Phe	His	Gly	Ser	Cys	Ile
1890					1895					1900					
Thr	Glu	Gly	Asn	Glu	Leu	Thr	Lys	Thr	Leu	Ile	Lys	Leu	Cys	Tyr	Asp
1905					1910					1915					1920
Ala	Phe	Thr	Glu	Asn	Met	Ala	Gly	Glu	Asn	Gln	Leu	Leu	Glu	Arg	Arg
1925					1930					1935					
Arg	Leu	Tyr	His	Cys	Ala	Ala	Tyr	Asn	Cys	Ala	Ile	Ser	Val	Ile	Cys
1940					1945					1950					
Cys	Val	Phe	Asn	Glu	Leu	Lys	Phe	Tyr	Gln	Gly	Phe	Leu	Phe	Ser	Glu
1955					1960					1965					
Lys	Pro	Glu	Lys	Asn	Leu	Leu	Ile	Phe	Glu	Asn	Leu	Ile	Asp	Leu	Lys
1970					1975					1980					
Arg	Arg	Tyr	Asn	Phe	Pro	Val	Glu	Val	Glu	Val	Pro	Met	Glu	Arg	Lys

1985					1990					1995				2000
Lys	Lys	Tyr	Ile	Glu	Ile	Arg	Lys	Glu	Ala	Arg	Glu	Ala	Ala	Asn Gly
				2005					2010					2015
Asp	Ser	Asp	Gly	Pro	Ser	Tyr	Met	Ser	Ser	Leu	Ser	Tyr	Leu	Ala Asp
			2020					2025					2030	
Ser	Thr	Leu	Ser	Glu	Glu	Met	Ser	Gln	Phe	Asp	Phe	Ser	Thr	Gly Val
		2035					2040					2045		
Gln	Ser	Tyr	Ser	Tyr	Ser	Ser	Gln	Asp	Pro	Arg	Pro	Ala	Thr	Gly Arg
	2050				2055						2060			
Phe	Arg	Arg	Arg	Glu	Gln	Arg	Asp	Pro	Thr	Val	His	Asp	Asp	Val Leu
2065					2070					2075				2080
Glu	Leu	Glu	Met	Asp	Glu	Leu	Asn	Arg	His	Glu	Cys	Met	Ala	Pro Leu
			2085					2090						2095
Thr	Ala	Leu	Val	Lys	His	Met	His	Arg	Ser	Leu	Gly	Pro	Pro	Gln Gly
			2100					2105						2110
Glu	Glu	Asp	Ser	Val	Pro	Arg	Asp	Leu	Pro	Ser	Trp	Met	Lys	Phe Leu
		2115					2120					2125		
His	Gly	Lys	Leu	Gly	Asn	Pro	Ile	Val	Pro	Leu	Asn	Ile	Arg	Leu Phe
	2130				2135						2140			
Leu	Ala	Lys	Leu	Val	Ile	Asn	Thr	Glu	Glu	Val	Phe	Arg	Pro	Tyr Ala
2145					2150					2155				2160
Lys	His	Trp	Leu	Ser	Pro	Leu	Leu	Gln	Leu	Ala	Ala	Ser	Glu	Asn Asn
			2165					2170						2175
Gly	Gly	Glu	Gly	Ile	His	Tyr	Met	Val	Val	Glu	Ile	Val	Ala	Thr Ile
			2180					2185					2190	
Leu	Ser	Trp	Thr	Gly	Leu	Ala	Thr	Pro	Thr	Gly	Val	Pro	Lys	Asp Glu
		2195					2200					2205		
Val	Leu	Ala	Asn	Arg	Leu	Leu	Asn	Phe	Leu	Met	Lys	His	Val	Phe His
	2210				2215						2220			
Pro	Lys	Arg	Ala	Val	Phe	Arg	His	Asn	Leu	Glu	Ile	Ile	Lys	Thr Leu
2225					2230					2235				2240
Val	Glu	Cys	Trp	Lys	Asp	Cys	Leu	Ser	Ile	Pro	Tyr	Arg	Leu	Ile Phe
			2245					2250						2255
Glu	Lys	Phe	Ser	Gly	Lys	Asp	Pro	Asn	Ser	Lys	Asp	Asn	Ser	Val Gly
			2260					2265					2270	
Ile	Gln	Leu	Leu	Gly	Ile	Val	Met	Ala	Asn	Asp	Leu	Pro	Pro	Tyr Asp
		2275					2280						2285	
Pro	Gln	Cys	Gly	Ile	Gln	Ser	Ser	Glu	Tyr	Phe	Gln	Ala	Leu	Val Asn
	2290				2295						2300			
Asn	Met	Ser	Phe	Val	Arg	Tyr	Lys	Glu	Val	Tyr	Ala	Ala	Ala	Ala Glu
2305					2310					2315				2320
Val	Leu	Gly	Leu	Ile	Leu	Arg	Tyr	Val	Met	Glu	Arg	Lys	Asn	Ile Leu
			2325					2330						2335
Glu	Glu	Ser	Leu	Cys	Glu	Leu	Val	Ala	Lys	Gln	Leu	Lys	Gln	His Gln
			2340					2345					2350	
Asn	Thr	Met	Glu	Asp	Lys	Phe	Ile	Val	Cys	Leu	Asn	Lys	Val	Thr Lys
		2355					2360					2365		
Ser	Phe	Pro	Pro	Leu	Ala	Asp	Arg	Phe	Met	Asn	Ala	Val	Phe	Phe Leu
		2370				2375					2380			
Leu	Pro	Lys	Phe	His	Gly	Val	Leu	Lys	Thr	Leu	Cys	Leu	Glu	Val Val
2385					2390					2395				2400
Leu	Cys	Arg	Val	Glu	Gly	Met	Thr	Glu	Leu	Tyr	Phe	Gln	Leu	Lys Ser
			2405					2410						2415
Lys	Asp	Phe	Val	Gln	Val	Met	Arg	His	Arg	Asp	Asp	Glu	Arg	Gln Lys

2850	2855	2860
Ala Ala Leu Leu Ser	Leu Asp Pro Ala	Ala Val Ser Ala Gly Cys Leu
2865	2870	2875
Ala Ser Leu Gln Gln	Pro Val Gly Ile	Arg Leu Leu Glu Glu Ala Leu
2885	2890	2895
Leu Arg Leu Leu Pro	Ala Glu Leu Pro	Ala Lys Arg Val Arg Gly Lys
2900	2905	2910
Ala Arg Leu Pro Pro	Asp Val Leu Arg Trp	Val Glu Leu Ala Lys Leu
2915	2920	2925
Tyr Arg Ser Ile Gly	Glu Tyr Asp Val	Leu Arg Gly Ile Phe Thr Ser
2930	2935	2940
Glu Ile Gly Thr Lys	Gln Ile Thr Gln	Ser Ala Leu Leu Ala Glu Ala
2945	2950	2955
Arg Ser Asp Tyr Ser	Glu Ala Ala Lys	Gln Tyr Asp Glu Ala Leu Asn
2965	2970	2975
Lys Gln Asp Trp Val	Asp Gly Glu Pro	Thr Glu Ala Glu Lys Asp Phe
2980	2985	2990
Trp Glu Leu Ala Ser	Leu Asp Cys Tyr	Asn His Leu Ala Glu Trp Lys
2995	3000	3005
Ser Leu Glu Tyr Cys	Ser Thr Ala Ser	Ile Asp Ser Glu Asn Pro Pro
3010	3015	3020
Asp Leu Asn Lys Ile	Trp Ser Glu Pro	Phe Tyr Gln Glu Thr Tyr Leu
3025	3030	3035
Pro Tyr Met Ile Arg	Ser Lys Leu Lys	Leu Leu Leu Gln Gly Glu Ala
3045	3050	3055
Asp Gln Ser Leu Leu	Thr Phe Ile Asp	Lys Ala Met His Gly Glu Leu
3060	3065	3070
Gln Lys Ala Ile Leu	Glu Leu His Tyr	Ser Gln Glu Leu Ser Leu Leu
3075	3080	3085
Tyr Leu Leu Gln Asp	Asp Val Asp Arg	Ala Lys Tyr Tyr Ile Gln Asn
3090	3095	3100
Gly Ile Gln Ser Phe	Met Gln Asn Tyr	Ser Ser Ile Asp Val Leu Leu
3105	3110	3115
His Gln Ser Arg Leu	Thr Lys Leu Gln	Ser Val Gln Ala Leu Thr Glu
3125	3130	3135
Ile Gln Glu Phe Ile	Ser Phe Ile Ser	Lys Gln Gly Asn Leu Ser Ser
3140	3145	3150
Gln Val Pro Leu Lys	Arg Leu Leu Asn	Thr Trp Thr Asn Arg Tyr Pro
3155	3160	3165
Asp Ala Lys Met Asp	Pro Met Asn Ile	Trp Asp Asp Ile Ile Thr Asn
3170	3175	3180
Arg Cys Phe Phe Leu	Ser Lys Ile Glu	Glu Lys Leu Thr Pro Leu Pro
3185	3190	3195
Glu Asp Asn Ser Met	Asn Val Asp Gln	Asp Gly Asp Pro Ser Asp Arg
3205	3210	3215
Met Glu Val Gln Glu	Gln Glu Glu Asp	Ile Ser Ser Leu Ile Arg Ser
3220	3225	3230
Cys Lys Phe Ser Met	Lys Met Lys Met	Ile Asp Ser Ala Arg Lys Gln
3235	3240	3245
Asn Asn Phe Ser Leu	Ala Met Lys Leu	Leu Lys Glu Leu His Lys Glu
3250	3255	3260
Ser Lys Thr Arg Asp	Asp Trp Leu Val	Ser Trp Val Gln Ser Tyr Cys
3265	3270	3275
Arg Leu Ser His Cys	Arg Ser Arg Ser	Gln Gly Cys Ser Glu Gln Val

3715	3720	3725
Ala Ser Leu Arg Arg Pro Lys Arg Ile Ile Ile Arg Gly His Asp Glu		
3730	3735	3740
Arg Glu His Pro Phe Leu Val Lys Gly Gly Glu Asp Leu Arg Gln Asp		
3745	3750	3755
Gln Arg Val Glu Gln Leu Phe Gln Val Met Asn Gly Ile Leu Ala Gln		3760
3765	3770	3775
Asp Ser Ala Cys Ser Gln Arg Ala Leu Gln Leu Arg Thr Tyr Ser Val		
3780	3785	3790
Val Pro Met Thr Ser Arg Leu Gly Leu Ile Glu Trp Leu Glu Asn Thr		
3795	3800	3805
Val Thr Leu Lys Asp Leu Leu Leu Asn Thr Met Ser Gln Glu Glu Lys		
3810	3815	3820
Ala Ala Tyr Leu Ser Asp Pro Arg Ala Pro Pro Cys Glu Tyr Lys Asp		
3825	3830	3835
Trp Leu Thr Lys Met Ser Gly Lys His Asp Val Gly Ala Tyr Met Leu		3840
3845	3850	3855
Met Tyr Lys Gly Ala Asn Arg Thr Glu Thr Val Thr Ser Phe Arg Lys		
3860	3865	3870
Arg Glu Ser Lys Val Pro Ala Asp Leu Leu Lys Arg Ala Phe Val Arg		
3875	3880	3885
Met Ser Thr Ser Pro Glu Ala Phe Leu Ala Leu Arg Ser His Phe Ala		
3890	3895	3900
Ser Ser His Ala Leu Ile Cys Ile Ser His Trp Ile Leu Gly Ile Gly		
3905	3910	3915
Asp Arg His Leu Asn Asn Phe Met Val Ala Met Glu Thr Gly Gly Val		3920
3925	3930	3935
Ile Gly Ile Asp Phe Gly His Ala Phe Gly Ser Ala Thr Gln Phe Leu		
3940	3945	3950
Pro Val Pro Glu Leu Met Pro Phe Arg Leu Thr Arg Gln Phe Ile Asn		
3955	3960	3965
Leu Met Leu Pro Met Lys Glu Thr Gly Leu Met Tyr Ser Ile Met Val		
3970	3975	3980
His Ala Leu Arg Ala Phe Arg Ser Asp Pro Gly Leu Leu Thr Asn Thr		
3985	3990	3995
Met Asp Val Phe Val Lys Glu Pro Ser Phe Asp Trp Lys Asn Phe Glu		4000
4005	4010	4015
Gln Lys Met Leu Lys Lys Gly Gly Ser Trp Ile Gln Glu Ile Asn Val		
4020	4025	4030
Ala Glu Lys Asn Trp Tyr Pro Arg Gln Lys Ile Cys Tyr Ala Lys Arg		
4035	4040	4045
Lys Leu Ala Gly Ala Asn Pro Ala Val Ile Thr Cys Asp Glu Leu Leu		
4050	4055	4060
Leu Gly His Glu Lys Ala Pro Ala Phe Arg Asp Tyr Val Ala Val Ala		
4065	4070	4075
Arg Gly Ser Lys Asp His Asn Ile Arg Ala Gln Glu Pro Glu Ser Gly		
4085	4090	4095
Leu Ser Glu Glu Thr Gln Val Lys Cys Leu Met Asp Gln Ala Thr Asp		
4100	4105	4110
Pro Asn Ile Leu Gly Arg Thr Trp Glu Gly Trp Glu Pro Trp Met		
4115	4120	4125

<210> 189

<211> 403

<212> PRT
 <213> Homo sapiens

<220>
 <223> >PTEN|ENSG00000171862|ENST00000371953|1212

<400> 189

Met	Thr	Ala	Ile	Ile	Lys	Glu	Ile	Val	Ser	Arg	Asn	Lys	Arg	Arg	Tyr	1	5	10	15
Gln	Glu	Asp	Gly	Phe	Asp	Leu	Asp	Leu	Thr	Tyr	Ile	Tyr	Pro	Asn	Ile	20	25	30	
Ile	Ala	Met	Gly	Phe	Pro	Ala	Glu	Arg	Leu	Glu	Gly	Val	Tyr	Arg	Asn	35	40	45	
Asn	Ile	Asp	Asp	Val	Val	Arg	Phe	Leu	Asp	Ser	Lys	His	Lys	Asn	His	50	55	60	
Tyr	Lys	Ile	Tyr	Asn	Leu	Cys	Ala	Glu	Arg	His	Tyr	Asp	Thr	Ala	Lys	65	70	75	
Phe	Asn	Cys	Arg	Val	Ala	Gln	Tyr	Pro	Phe	Glu	Asp	His	Asn	Pro	Pro	85	90	95	
Gln	Leu	Glu	Leu	Ile	Lys	Pro	Phe	Cys	Glu	Asp	Leu	Asp	Gln	Trp	Leu	100	105	110	
Ser	Glu	Asp	Asp	Asn	His	Val	Ala	Ala	Ile	His	Cys	Lys	Ala	Gly	Lys	115	120	125	
Gly	Arg	Thr	Gly	Val	Met	Ile	Cys	Ala	Tyr	Leu	Leu	His	Arg	Gly	Lys	130	135	140	
Phe	Leu	Lys	Ala	Gln	Glu	Ala	Leu	Asp	Phe	Tyr	Gly	Glu	Val	Arg	Thr	145	150	155	
Arg	Asp	Lys	Lys	Gly	Val	Thr	Ile	Pro	Ser	Gln	Arg	Arg	Tyr	Val	Tyr	165	170	175	
Tyr	Tyr	Ser	Tyr	Leu	Leu	Lys	Asn	His	Leu	Asp	Tyr	Arg	Pro	Val	Ala	180	185	190	
Leu	Leu	Phe	His	Lys	Met	Met	Phe	Glu	Thr	Ile	Pro	Met	Phe	Ser	Gly	195	200	205	
Gly	Thr	Cys	Asn	Pro	Gln	Phe	Val	Val	Cys	Gln	Leu	Lys	Val	Lys	Ile	210	215	220	
Tyr	Ser	Ser	Asn	Ser	Gly	Pro	Thr	Arg	Arg	Glu	Asp	Lys	Phe	Met	Tyr	225	230	235	
Phe	Glu	Phe	Pro	Gln	Pro	Leu	Pro	Val	Cys	Gly	Asp	Ile	Lys	Val	Glu	245	250	255	
Phe	Phe	His	Lys	Gln	Asn	Lys	Met	Leu	Lys	Lys	Asp	Lys	Met	Phe	His	260	265	270	
Phe	Trp	Val	Asn	Thr	Phe	Phe	Ile	Pro	Gly	Pro	Glu	Glu	Thr	Ser	Glu	275	280	285	
Lys	Val	Glu	Asn	Gly	Ser	Leu	Cys	Asp	Gln	Glu	Ile	Asp	Ser	Ile	Cys	290	295	300	
Ser	Ile	Glu	Arg	Ala	Asp	Asn	Asp	Lys	Glu	Tyr	Leu	Val	Leu	Thr	Leu	305	310	315	
Thr	Lys	Asn	Asp	Leu	Asp	Lys	Ala	Asn	Lys	Asp	Lys	Ala	Asn	Arg	Tyr	325	330	335	
Phe	Ser	Pro	Asn	Phe	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Lys	Thr	Val	Glu	340	345	350	
Glu	Pro	Ser	Asn	Pro	Glu	Ala	Ser	Ser	Ser	Thr	Ser	Val	Thr	Pro	Asp	355	360	365	

Val	Ser	Asp	Asn	Glu	Pro	Asp	His	Tyr	Arg	Tyr	Ser	Asp	Thr	Thr	Asp
370						375					380				
Ser	Asp	Pro	Glu	Asn	Glu	Pro	Phe	Asp	Glu	Asp	Gln	His	Thr	Gln	Ile
385					390					395					400
Thr	Lys	Val													

<210> 190
 <211> 681
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >RAD17|ENSG00000152942|ENST00000509734|2046

<400> 190

Met	Ser	Lys	Thr	Phe	Leu	Arg	Pro	Lys	Val	Ser	Ser	Thr	Lys	Val	Thr
1				5				10					15		
Asp	Trp	Val	Asp	Pro	Ser	Phe	Asp	Asp	Phe	Leu	Glu	Cys	Ser	Gly	Val
		20					25					30			
Ser	Thr	Ile	Thr	Ala	Thr	Ser	Leu	Gly	Val	Asn	Asn	Ser	Ser	His	Arg
		35				40					45				
Arg	Lys	Asn	Gly	Pro	Ser	Thr	Leu	Glu	Ser	Ser	Arg	Phe	Pro	Ala	Arg
	50				55					60					
Lys	Arg	Gly	Asn	Leu	Ser	Ser	Leu	Glu	Gln	Ile	Tyr	Gly	Leu	Glu	Asn
65				70					75					80	
Ser	Lys	Glu	Tyr	Leu	Ser	Glu	Asn	Glu	Pro	Trp	Val	Asp	Lys	Tyr	Lys
				85				90					95		
Pro	Glu	Thr	Gln	His	Glu	Leu	Ala	Val	His	Lys	Lys	Lys	Ile	Glu	Glu
			100					105					110		
Val	Glu	Thr	Trp	Leu	Lys	Ala	Gln	Val	Leu	Glu	Arg	Gln	Pro	Lys	Gln
		115					120					125			
Gly	Gly	Ser	Ile	Leu	Leu	Ile	Thr	Gly	Pro	Pro	Gly	Cys	Gly	Lys	Thr
	130					135					140				
Thr	Thr	Leu	Lys	Ile	Leu	Ser	Lys	Glu	His	Gly	Ile	Gln	Val	Gln	Glu
145					150				155						160
Trp	Ile	Asn	Pro	Val	Leu	Pro	Asp	Phe	Gln	Lys	Asp	Asp	Phe	Lys	Gly
				165				170						175	
Met	Phe	Asn	Thr	Glu	Ser	Ser	Phe	His	Met	Phe	Pro	Tyr	Gln	Ser	Gln
			180					185					190		
Ile	Ala	Val	Phe	Lys	Glu	Phe	Leu	Leu	Arg	Ala	Thr	Lys	Tyr	Asn	Lys
		195					200					205			
Leu	Gln	Met	Leu	Gly	Asp	Asp	Leu	Arg	Thr	Asp	Lys	Lys	Ile	Ile	Leu
	210				215						220				
Val	Glu	Asp	Leu	Pro	Asn	Gln	Phe	Tyr	Arg	Asp	Ser	His	Thr	Leu	His
225					230					235					240
Glu	Val	Leu	Arg	Lys	Tyr	Val	Arg	Ile	Gly	Arg	Cys	Pro	Leu	Ile	Phe
				245					250					255	
Ile	Ile	Ser	Asp	Ser	Leu	Ser	Gly	Asp	Asn	Asn	Gln	Arg	Leu	Leu	Phe
			260				265						270		
Pro	Lys	Glu	Ile	Gln	Glu	Glu	Cys	Ser	Ile	Ser	Asn	Ile	Ser	Phe	Asn
		275					280					285			
Pro	Val	Ala	Pro	Thr	Ile	Met	Met	Lys	Phe	Leu	Asn	Arg	Ile	Val	Thr

290	Ile	Glu	Ala	Asn	Lys	Asn	Gly	Gly	Lys	Ile	Thr	Val	Pro	Asp	Lys	Thr
305	Ser	Leu	Glu	Leu	Leu	Cys	Gln	Gly	Cys	Ser	Gly	Asp	Ile	Arg	Ser	Ala
				325						330					335	
	Ile	Asn	Ser	Leu	Gln	Phe	Ser	Ser	Ser	Lys	Gly	Glu	Asn	Asn	Leu	Arg
				340						345					350	
	Pro	Arg	Lys	Lys	Gly	Met	Ser	Leu	Lys	Ser	Asp	Ala	Val	Leu	Ser	Lys
			355					360					365			
	Ser	Lys	Arg	Arg	Lys	Lys	Pro	Asp	Arg	Val	Phe	Glu	Asn	Gln	Glu	Val
		370					375					380				
	Gln	Ala	Ile	Gly	Gly	Lys	Asp	Val	Ser	Leu	Phe	Leu	Phe	Arg	Ala	Leu
385						390					395					400
	Gly	Lys	Ile	Leu	Tyr	Cys	Lys	Arg	Ala	Ser	Leu	Thr	Glu	Leu	Asp	Ser
					405					410					415	
	Pro	Arg	Leu	Pro	Ser	His	Leu	Ser	Glu	Tyr	Glu	Arg	Asp	Thr	Leu	Leu
			420						425					430		
	Val	Glu	Pro	Glu	Glu	Val	Val	Glu	Met	Ser	His	Met	Pro	Gly	Asp	Leu
			435					440					445			
	Phe	Asn	Leu	Tyr	Leu	His	Gln	Asn	Tyr	Ile	Asp	Phe	Phe	Met	Glu	Ile
		450					455					460				
	Asp	Asp	Ile	Val	Arg	Ala	Ser	Glu	Phe	Leu	Ser	Phe	Ala	Asp	Ile	Leu
465						470					475					480
	Ser	Gly	Asp	Trp	Asn	Thr	Arg	Ser	Leu	Leu	Arg	Glu	Tyr	Ser	Thr	Ser
				485						490					495	
	Ile	Ala	Thr	Arg	Gly	Val	Met	His	Ser	Asn	Lys	Ala	Arg	Gly	Tyr	Ala
			500						505					510		
	His	Cys	Gln	Gly	Gly	Gly	Ser	Ser	Phe	Arg	Pro	Leu	His	Lys	Pro	Gln
		515						520				525				
	Trp	Phe	Leu	Ile	Asn	Lys	Lys	Tyr	Arg	Glu	Asn	Cys	Leu	Ala	Ala	Lys
		530					535					540				
	Ala	Leu	Phe	Pro	Asp	Phe	Cys	Leu	Pro	Ala	Leu	Cys	Leu	Gln	Thr	Gln
545						550					555					560
	Leu	Leu	Pro	Tyr	Leu	Ala	Leu	Leu	Thr	Ile	Pro	Met	Arg	Asn	Gln	Ala
				565						570					575	
	Gln	Ile	Ser	Phe	Ile	Gln	Asp	Ile	Gly	Arg	Leu	Pro	Leu	Lys	Arg	His
			580						585					590		
	Phe	Gly	Arg	Leu	Lys	Met	Glu	Ala	Leu	Thr	Asp	Arg	Glu	His	Gly	Met
		595					600					605				
	Ile	Asp	Pro	Asp	Ser	Gly	Asp	Glu	Ala	Gln	Leu	Asn	Gly	Gly	His	Ser
		610					615					620				
	Ala	Glu	Glu	Ser	Leu	Gly	Glu	Pro	Thr	Gln	Ala	Thr	Val	Pro	Glu	Thr
625						630					635					640
	Trp	Ser	Leu	Pro	Leu	Ser	Gln	Asn	Ser	Ala	Ser	Glu	Leu	Pro	Ala	Ser
				645						650					655	
	Gln	Pro	Gln	Pro	Phe	Ser	Ala	Gln	Gly	Asp	Met	Glu	Glu	Asn	Ile	Ile
			660						665					670		
	Ile	Glu	Asp	Tyr	Glu	Ser	Asp	Gly	Thr							
		675						680								

<210> 191

<211> 495

<212> PRT

<213> Homo sapiens

<220>

<223> >RAD18|ENSG00000070950|ENST00000264926|1488

<400> 191

Met	Asp	Ser	Leu	Ala	Glu	Ser	Arg	Trp	Pro	Pro	Gly	Leu	Ala	Val	Met
1			5					10					15		
Lys	Thr	Ile	Asp	Asp	Leu	Leu	Arg	Cys	Gly	Ile	Cys	Phe	Glu	Tyr	Phe
		20						25				30			
Asn	Ile	Ala	Met	Ile	Ile	Pro	Gln	Cys	Ser	His	Asn	Tyr	Cys	Ser	Leu
		35				40					45				
Cys	Ile	Arg	Lys	Phe	Leu	Ser	Tyr	Lys	Thr	Gln	Cys	Pro	Thr	Cys	Cys
	50					55				60					
Val	Thr	Val	Thr	Glu	Pro	Asp	Leu	Lys	Asn	Asn	Arg	Ile	Leu	Asp	Glu
65				70					75					80	
Leu	Val	Lys	Ser	Leu	Asn	Phe	Ala	Arg	Asn	His	Leu	Leu	Gln	Phe	Ala
			85					90					95		
Leu	Glu	Ser	Pro	Ala	Lys	Ser	Pro	Ala	Ser	Ser	Ser	Ser	Lys	Asn	Leu
			100					105					110		
Ala	Val	Lys	Val	Tyr	Thr	Pro	Val	Ala	Ser	Arg	Gln	Ser	Leu	Lys	Gln
		115				120					125				
Gly	Ser	Arg	Leu	Met	Asp	Asn	Phe	Leu	Ile	Arg	Glu	Met	Ser	Gly	Ser
	130					135					140				
Thr	Ser	Glu	Leu	Leu	Ile	Lys	Glu	Asn	Lys	Ser	Lys	Phe	Ser	Pro	Gln
145				150					155					160	
Lys	Glu	Ala	Ser	Pro	Ala	Ala	Lys	Thr	Lys	Glu	Thr	Arg	Ser	Val	Glu
			165					170					175		
Glu	Ile	Ala	Pro	Asp	Pro	Ser	Glu	Ala	Lys	Arg	Pro	Glu	Pro	Pro	Ser
		180					185					190			
Thr	Ser	Thr	Leu	Lys	Gln	Val	Thr	Lys	Val	Asp	Cys	Pro	Val	Cys	Gly
	195					200					205				
Val	Asn	Ile	Pro	Glu	Ser	His	Ile	Asn	Lys	His	Leu	Asp	Ser	Cys	Leu
	210					215					220				
Ser	Arg	Glu	Glu	Lys	Lys	Glu	Ser	Leu	Arg	Ser	Ser	Val	His	Lys	Arg
225				230					235					240	
Lys	Pro	Leu	Pro	Lys	Thr	Val	Tyr	Asn	Leu	Ser	Asp	Arg	Asp	Leu	
			245					250					255		
Lys	Lys	Lys	Leu	Lys	Glu	His	Gly	Leu	Ser	Ile	Gln	Gly	Asn	Lys	Gln
		260					265					270			
Gln	Leu	Ile	Lys	Arg	His	Gln	Glu	Phe	Val	His	Met	Tyr	Asn	Ala	Gln
	275					280					285				
Cys	Asp	Ala	Leu	His	Pro	Lys	Ser	Ala	Ala	Glu	Ile	Val	Arg	Glu	Ile
	290					295					300				
Glu	Asn	Ile	Glu	Lys	Thr	Arg	Met	Arg	Leu	Glu	Ala	Ser	Lys	Leu	Asn
305				310					315					320	
Glu	Ser	Val	Met	Val	Phe	Thr	Lys	Asp	Gln	Thr	Glu	Lys	Glu	Ile	Asp
			325					330					335		
Glu	Ile	His	Ser	Lys	Tyr	Arg	Lys	Lys	His	Lys	Ser	Glu	Phe	Gln	Leu
	340						345					350			
Leu	Val	Asp	Gln	Ala	Arg	Lys	Gly	Tyr	Lys	Lys	Ile	Ala	Gly	Met	Ser
	355					360					365				
Gln	Lys	Thr	Val	Thr	Ile	Thr	Lys	Glu	Asp	Glu	Ser	Thr	Glu	Lys	Leu
	370					375					380				

Ser	Ser	Val	Cys	Met	Gly	Gln	Glu	Asp	Asn	Met	Thr	Ser	Val	Thr	Asn
385					390					395					400
His	Phe	Ser	Gln	Ser	Lys	Leu	Asp	Ser	Pro	Glu	Glu	Leu	Glu	Pro	Asp
				405					410						415
Arg	Glu	Glu	Asp	Ser	Ser	Ser	Cys	Ile	Asp	Ile	Gln	Glu	Val	Leu	Ser
			420					425					430		
Ser	Ser	Glu	Ser	Asp	Ser	Cys	Asn	Ser	Ser	Ser	Ser	Asp	Ile	Ile	Arg
		435					440					445			
Asp	Leu	Leu	Glu	Glu	Glu	Glu	Ala	Trp	Glu	Ala	Ser	His	Lys	Asn	Asp
	450					455				460					
Leu	Gln	Asp	Thr	Glu	Ile	Ser	Pro	Arg	Gln	Asn	Arg	Arg	Thr	Arg	Ala
465					470					475					480
Ala	Glu	Ser	Ala	Glu	Ile	Glu	Pro	Arg	Asn	Lys	Arg	Asn	Arg	Asn	
				485					490						495

<210> 192

<211> 1312

<212> PRT

<213> Homo sapiens

<220>

<223> >RAD50|ENSG00000113522|ENST00000265335|3939

<400> 192

Met	Ser	Arg	Ile	Glu	Lys	Met	Ser	Ile	Leu	Gly	Val	Arg	Ser	Phe	Gly
1				5					10					15	
Ile	Glu	Asp	Lys	Asp	Lys	Gln	Ile	Ile	Thr	Phe	Phe	Ser	Pro	Leu	Thr
			20				25						30		
Ile	Leu	Val	Gly	Pro	Asn	Gly	Ala	Gly	Lys	Thr	Thr	Ile	Ile	Glu	Cys
		35				40						45			
Leu	Lys	Tyr	Ile	Cys	Thr	Gly	Asp	Phe	Pro	Pro	Gly	Thr	Lys	Gly	Asn
	50					55					60				
Thr	Phe	Val	His	Asp	Pro	Lys	Val	Ala	Gln	Glu	Thr	Asp	Val	Arg	Ala
65				70					75					80	
Gln	Ile	Arg	Leu	Gln	Phe	Arg	Asp	Val	Asn	Gly	Glu	Leu	Ile	Ala	Val
			85					90						95	
Gln	Arg	Ser	Met	Val	Cys	Thr	Gln	Lys	Ser	Lys	Lys	Thr	Glu	Phe	Lys
		100					105						110		
Thr	Leu	Glu	Gly	Val	Ile	Thr	Arg	Thr	Lys	His	Gly	Glu	Lys	Val	Ser
	115					120					125				
Leu	Ser	Ser	Lys	Cys	Ala	Glu	Ile	Asp	Arg	Glu	Met	Ile	Ser	Ser	Leu
	130					135				140					
Gly	Val	Ser	Lys	Ala	Val	Leu	Asn	Asn	Val	Ile	Phe	Cys	His	Gln	Glu
145				150					155					160	
Asp	Ser	Asn	Trp	Pro	Leu	Ser	Glu	Gly	Lys	Ala	Leu	Lys	Gln	Lys	Phe
			165					170						175	
Asp	Glu	Ile	Phe	Ser	Ala	Thr	Arg	Tyr	Ile	Lys	Ala	Leu	Glu	Thr	Leu
		180					185						190		
Arg	Gln	Val	Arg	Gln	Thr	Gln	Gly	Gln	Lys	Val	Lys	Glu	Tyr	Gln	Met
	195					200						205			
Glu	Leu	Lys	Tyr	Leu	Lys	Gln	Tyr	Lys	Glu	Lys	Ala	Cys	Glu	Ile	Arg
	210					215					220				
Asp	Gln	Ile	Thr	Ser	Lys	Glu	Ala	Gln	Leu	Thr	Ser	Ser	Lys	Glu	Ile

225					230					235				240
Val	Lys	Ser	Tyr	Glu	Asn	Glu	Leu	Asp	Pro	Leu	Lys	Asn	Arg	Leu
				245					250					255
Glu	Ile	Glu	His	Asn	Leu	Ser	Lys	Ile	Met	Lys	Leu	Asp	Asn	Glu
			260					265						270
Lys	Ala	Leu	Asp	Ser	Arg	Lys	Lys	Gln	Met	Glu	Lys	Asp	Asn	Ser
			275					280						285
Leu	Glu	Glu	Lys	Met	Glu	Lys	Val	Phe	Gln	Gly	Thr	Asp	Glu	Gln
	290					295					300			
Asn	Asp	Leu	Tyr	His	Asn	His	Gln	Arg	Thr	Val	Arg	Glu	Lys	Glu
305					310					315				320
Lys	Leu	Val	Asp	Cys	His	Arg	Glu	Leu	Glu	Lys	Leu	Asn	Lys	Glu
				325					330					335
Arg	Leu	Leu	Asn	Gln	Glu	Lys	Ser	Glu	Leu	Leu	Val	Glu	Gln	Gly
			340					345						350
Leu	Gln	Leu	Gln	Ala	Asp	Arg	His	Gln	Glu	His	Ile	Arg	Ala	Arg
			355				360					365		
Ser	Leu	Ile	Gln	Ser	Leu	Ala	Thr	Gln	Leu	Glu	Leu	Asp	Gly	Phe
	370					375						380		
Arg	Gly	Pro	Phe	Ser	Glu	Arg	Gln	Ile	Lys	Asn	Phe	His	Lys	Leu
385					390					395				400
Arg	Glu	Arg	Gln	Glu	Gly	Glu	Ala	Lys	Thr	Ala	Asn	Gln	Leu	Met
				405					410					415
Asp	Phe	Ala	Glu	Lys	Glu	Thr	Leu	Lys	Gln	Lys	Gln	Ile	Asp	Glu
			420					425					430	
Arg	Asp	Lys	Lys	Thr	Gly	Leu	Gly	Arg	Ile	Ile	Glu	Leu	Lys	Ser
		435					440				445			
Ile	Leu	Ser	Lys	Lys	Gln	Asn	Glu	Leu	Lys	Asn	Val	Lys	Tyr	Glu
	450					455					460			
Gln	Gln	Leu	Glu	Gly	Ser	Ser	Asp	Arg	Ile	Leu	Glu	Leu	Asp	Gln
465					470					475				480
Leu	Ile	Lys	Ala	Glu	Arg	Glu	Leu	Ser	Lys	Ala	Glu	Lys	Asn	Ser
			485						490					495
Val	Glu	Thr	Leu	Lys	Met	Glu	Val	Ile	Ser	Leu	Gln	Asn	Glu	Lys
			500					505						510
Asp	Leu	Asp	Arg	Thr	Leu	Arg	Lys	Leu	Asp	Gln	Glu	Met	Glu	Gln
		515					520					525		
Asn	His	His	Thr	Thr	Thr	Arg	Thr	Gln	Met	Glu	Met	Leu	Thr	Lys
	530					535					540			
Lys	Ala	Asp	Lys	Asp	Glu	Gln	Ile	Arg	Lys	Ile	Lys	Ser	Arg	His
545					550					555				560
Asp	Glu	Leu	Thr	Ser	Leu	Leu	Gly	Tyr	Phe	Pro	Asn	Lys	Lys	Gln
				565					570					575
Glu	Asp	Trp	Leu	His	Ser	Lys	Ser	Lys	Glu	Ile	Asn	Gln	Thr	Arg
			580					585					590	
Arg	Leu	Ala	Lys	Leu	Asn	Lys	Glu	Leu	Ala	Ser	Ser	Glu	Gln	Asn
		595					600					605		
Asn	His	Ile	Asn	Asn	Glu	Leu	Lys	Arg	Lys	Glu	Glu	Gln	Leu	Ser
	610					615					620			
Tyr	Glu	Asp	Lys	Leu	Phe	Asp	Val	Cys	Gly	Ser	Gln	Asp	Phe	Glu
625					630					635				640
Asp	Leu	Asp	Arg	Leu	Lys	Glu	Glu	Ile	Glu	Lys	Ser	Ser	Lys	Gln
			645						650					655
Ala	Met	Leu	Ala	Gly	Ala	Thr	Ala	Val	Tyr	Ser	Gln	Phe	Ile	Thr

1090	1095	1100
Arg Glu Met Met Ile Val Met Arg Thr Thr Glu Leu Val Asn Lys Asp		
1105	1110	1115
Leu Asp Ile Tyr Tyr Lys Thr Leu Asp Gln Ala Ile Met Lys Phe His		1120
	1125	1130
Ser Met Lys Met Glu Glu Ile Asn Lys Ile Ile Arg Asp Leu Trp Arg		1135
	1140	1145
Ser Thr Tyr Arg Gly Gln Asp Ile Glu Tyr Ile Glu Ile Arg Ser Asp		1150
	1155	1160
Ala Asp Glu Asn Val Ser Ala Ser Asp Lys Arg Arg Asn Tyr Asn Tyr		1165
	1170	1175
Arg Val Val Met Leu Lys Gly Asp Thr Ala Leu Asp Met Arg Gly Arg		1180
1185	1190	1195
Cys Ser Ala Gly Gln Lys Val Leu Ala Ser Leu Ile Ile Arg Leu Ala		1200
	1205	1210
Leu Ala Glu Thr Phe Cys Leu Asn Cys Gly Ile Ile Ala Leu Asp Glu		1215
	1220	1225
Pro Thr Thr Asn Leu Asp Arg Glu Asn Ile Glu Ser Leu Ala His Ala		1230
	1235	1240
Leu Val Glu Ile Ile Lys Ser Arg Ser Gln Gln Arg Asn Phe Gln Leu		1245
	1250	1255
Leu Val Ile Thr His Asp Glu Asp Phe Val Glu Leu Leu Gly Arg Ser		1260
1265	1270	1275
Glu Tyr Val Glu Lys Phe Tyr Arg Ile Lys Lys Asn Ile Asp Gln Cys		1280
	1285	1290
Ser Glu Ile Val Lys Cys Ser Val Ser Ser Leu Gly Phe Asn Val His		1295
	1300	1305
		1310

<210> 193
 <211> 340
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >RAD51|ENSG00000051180|ENST00000382643|1023

<400> 193
Met Ala Met Gln Met Gln Leu Glu Ala Asn Ala Asp Thr Ser Val Glu
1 5 10 15
Glu Glu Ser Phe Gly Pro Gln Pro Ile Ser Arg Leu Glu Gln Cys Gly
20 25 30
Ile Asn Ala Asn Asp Val Lys Lys Leu Glu Glu Ala Gly Phe His Thr
35 40 45
Val Glu Ala Val Ala Tyr Ala Pro Lys Lys Glu Leu Ile Asn Ile Lys
50 55 60
Gly Ile Ser Glu Ala Lys Ala Asp Lys Ile Leu Thr Glu Ser Arg Ser
65 70 75 80
Val Ala Arg Leu Glu Cys Asn Ser Val Ile Leu Val Tyr Cys Thr Leu
85 90 95
Arg Leu Ser Gly Ser Ser Asp Ser Pro Ala Ser Ala Ser Arg Val Val
100 105 110

Gly	Thr	Thr	Gly	Gly	Ile	Glu	Thr	Gly	Ser	Ile	Thr	Glu	Met	Phe	Gly
		115					120					125			
Glu	Phe	Arg	Thr	Gly	Lys	Thr	Gln	Ile	Cys	His	Thr	Leu	Ala	Val	Thr
	130					135					140				
Cys	Gln	Leu	Pro	Ile	Asp	Arg	Gly	Gly	Gly	Glu	Gly	Lys	Ala	Met	Tyr
145					150					155					160
Ile	Asp	Thr	Glu	Gly	Thr	Phe	Arg	Pro	Glu	Arg	Leu	Leu	Ala	Val	Ala
			165						170					175	
Glu	Arg	Tyr	Gly	Leu	Ser	Gly	Ser	Asp	Val	Leu	Asp	Asn	Val	Ala	Tyr
			180					185					190		
Ala	Arg	Ala	Phe	Asn	Thr	Asp	His	Gln	Thr	Gln	Leu	Leu	Tyr	Gln	Ala
		195				200						205			
Ser	Ala	Met	Met	Val	Glu	Ser	Arg	Tyr	Ala	Leu	Leu	Ile	Val	Asp	Ser
	210					215					220				
Ala	Thr	Ala	Leu	Tyr	Arg	Thr	Asp	Tyr	Ser	Gly	Arg	Gly	Glu	Leu	Ser
225					230					235					240
Ala	Arg	Gln	Met	His	Leu	Ala	Arg	Phe	Leu	Arg	Met	Leu	Leu	Arg	Leu
				245				250						255	
Ala	Asp	Glu	Phe	Gly	Val	Ala	Val	Val	Ile	Thr	Asn	Gln	Val	Val	Ala
			260					265					270		
Gln	Val	Asp	Gly	Ala	Ala	Met	Phe	Ala	Ala	Asp	Pro	Lys	Lys	Pro	Ile
		275					280					285			
Gly	Gly	Asn	Ile	Ile	Ala	His	Ala	Ser	Thr	Thr	Arg	Leu	Tyr	Leu	Arg
	290					295					300				
Lys	Gly	Arg	Gly	Glu	Thr	Arg	Ile	Cys	Lys	Ile	Tyr	Asp	Ser	Pro	Cys
305					310					315					320
Leu	Pro	Glu	Ala	Glu	Ala	Met	Phe	Ala	Ile	Asn	Ala	Asp	Gly	Val	Gly
				325					330					335	
Asp	Ala	Lys	Asp												
			340												

<210> 194
 <211> 418
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >RAD52|ENSG00000002016|ENST00000358495|1257

<400> 194

Met	Ser	Gly	Thr	Glu	Glu	Ala	Ile	Leu	Gly	Gly	Arg	Asp	Ser	His	Pro
1				5					10					15	
Ala	Ala	Gly	Gly	Gly	Ser	Val	Leu	Cys	Phe	Gly	Gln	Cys	Gln	Tyr	Thr
		20						25					30		
Ala	Glu	Glu	Tyr	Gln	Ala	Ile	Gln	Lys	Ala	Leu	Arg	Gln	Arg	Leu	Gly
		35					40					45			
Pro	Glu	Tyr	Ile	Ser	Ser	Arg	Met	Ala	Gly	Gly	Gly	Gln	Lys	Val	Cys
	50					55					60				
Tyr	Ile	Glu	Gly	His	Arg	Val	Ile	Asn	Leu	Ala	Asn	Glu	Met	Phe	Gly
65				70					75					80	
Tyr	Asn	Gly	Trp	Ala	His	Ser	Ile	Thr	Gln	Gln	Asn	Val	Asp	Phe	Val
				85					90				95		
Asp	Leu	Asn	Asn	Gly	Lys	Phe	Tyr	Val	Gly	Val	Cys	Ala	Phe	Val	Arg

Lys	Pro	Lys	Phe	Ile	Pro	Pro	Gly	Arg	Ser	Asn	Pro	Gly	Leu	Asn	Glu		
			20					25					30				
Glu	Ile	Thr	Lys	Leu	Asn	Pro	Asp	Ile	Lys	Leu	Phe	Glu	Gly	Val	Ala		
		35					40					45					
Ile	Asn	Asn	Thr	Phe	Leu	Pro	Ser	Gln	Asn	Asp	Leu	Arg	Ile	Cys	Ser		
	50					55					60						
Leu	Asn	Leu	Pro	Ser	Glu	Glu	Ser	Thr	Arg	Glu	Ile	Asn	Asn	Arg	Asp		
65					70					75					80		
Asn	Cys	Ser	Gly	Lys	Tyr	Cys	Phe	Glu	Ala	Pro	Thr	Leu	Ala	Thr	Leu		
			85					90						95			
Asp	Pro	Pro	His	Thr	Val	His	Ser	Ala	Pro	Lys	Glu	Val	Ala	Val	Ser		
			100					105					110				
Lys	Glu	Gln	Glu	Glu	Lys	Ser	Asp	Ser	Leu	Val	Lys	Tyr	Phe	Ser	Val		
		115					120					125					
Val	Trp	Cys	Lys	Pro	Ser	Lys	Lys	Lys	His	Lys	Lys	Trp	Glu	Gly	Asp		
	130					135						140					
Ala	Val	Leu	Ile	Val	Lys	Gly	Lys	Ser	Phe	Ile	Leu	Lys	Asn	Leu	Glu		
145					150					155					160		
Gly	Lys	Asp	Ile	Gly	Arg	Gly	Ile	Gly	Tyr	Lys	Phe	Lys	Glu	Leu	Glu		
				165					170					175			
Lys	Ile	Glu	Glu	Gly	Gln	Thr	Leu	Met	Ile	Cys	Gly	Lys	Glu	Ile	Glu		
		180						185					190				
Val	Met	Gly	Val	Ile	Ser	Pro	Asp	Asp	Phe	Ser	Ser	Gly	Arg	Cys	Phe		
		195					200					205					
Gln	Leu	Gly	Gly	Gly	Ser	Thr	Ala	Ile	Ser	His	Ser	Ser	Gln	Val	Ala		
	210					215					220						
Arg	Lys	Cys	Phe	Ser	Asn	Pro	Phe	Lys	Ser	Val	Cys	Lys	Pro	Ser	Ser		
225					230					235					240		
Lys	Glu	Asn	Arg	Gln	Asn	Asp	Phe	Gln	Asn	Cys	Lys	Pro	Arg	His	Asp		
				245				250						255			
Pro	Tyr	Thr	Pro	Asn	Ser	Leu	Val	Met	Pro	Arg	Pro	Asp	Lys	Asn	His		
			260					265					270				
Gln	Trp	Val	Phe	Asn	Lys	Asn	Cys	Phe	Pro	Leu	Val	Asp	Val	Val	Ile		
		275					280					285					
Asp	Pro	Tyr	Leu	Val	Tyr	His	Leu	Arg	Pro	His	Gln	Lys	Glu	Gly	Ile		
	290					295					300						
Ile	Phe	Leu	Tyr	Glu	Cys	Val	Met	Gly	Met	Arg	Met	Asn	Gly	Arg	Cys		
305					310				315						320		
Gly	Ala	Ile	Leu	Ala	Asp	Glu	Met	Gly	Leu	Gly	Lys	Thr	Leu	Gln	Cys		
				325				330						335			
Ile	Ser	Leu	Ile	Trp	Thr	Leu	Gln	Cys	Gln	Gly	Pro	Tyr	Gly	Gly	Lys		
			340					345					350				
Pro	Val	Ile	Lys	Lys	Thr	Leu	Ile	Val	Thr	Pro	Gly	Ser	Leu	Val	Asn		
		355					360					365					
Asn	Trp	Lys	Lys	Glu	Phe	Gln	Lys	Trp	Leu	Gly	Ser	Glu	Arg	Ile	Lys		
	370					375					380						
Ile	Phe	Thr	Val	Asp	Gln	Asp	His	Lys	Val	Glu	Glu	Phe	Ile	Lys	Ser		
385					390					395					400		
Ile	Phe	Tyr	Ser	Val	Leu	Ile	Ile	Ser	Tyr	Glu	Met	Leu	Leu	Arg	Ser		
				405					410					415			
Leu	Asp	Gln	Ile	Lys	Asn	Ile	Lys	Phe	Asp	Leu	Leu	Ile	Cys	Asp	Glu		
			420					425					430				
Gly	His	Arg	Leu	Lys	Asn	Ser	Ala	Ile	Lys	Thr	Thr	Thr	Ala	Leu	Ile		
		435					440						445				

Ser	Leu	Ser	Cys	Glu	Lys	Arg	Ile	Ile	Leu	Thr	Gly	Thr	Pro	Ile	Gln
450						455					460				
Asn	Asp	Leu	Gln	Glu	Phe	Phe	Ala	Leu	Ile	Asp	Phe	Val	Asn	Pro	Gly
465					470					475					480
Ile	Leu	Gly	Ser	Leu	Ser	Ser	Tyr	Arg	Lys	Ile	Tyr	Glu	Glu	Pro	Ile
				485					490					495	
Ile	Leu	Ser	Arg	Glu	Pro	Ser	Ala	Ser	Glu	Glu	Glu	Lys	Glu	Leu	Gly
			500					505					510		
Glu	Arg	Arg	Ala	Ala	Glu	Leu	Thr	Cys	Leu	Thr	Gly	Leu	Phe	Ile	Leu
		515					520					525			
Arg	Arg	Thr	Gln	Glu	Ile	Ile	Asn	Lys	Tyr	Leu	Pro	Pro	Lys	Ile	Glu
		530				535					540				
Asn	Val	Val	Phe	Cys	Arg	Pro	Gly	Ala	Leu	Gln	Ile	Glu	Leu	Tyr	Arg
545					550					555					560
Lys	Leu	Leu	Asn	Ser	Gln	Val	Val	Arg	Phe	Cys	Leu	Gln	Gly	Leu	Leu
				565					570					575	
Glu	Asn	Ser	Pro	His	Leu	Ile	Cys	Ile	Gly	Ala	Leu	Lys	Lys	Leu	Cys
			580					585					590		
Asn	His	Pro	Cys	Leu	Leu	Phe	Asn	Ser	Ile	Lys	Glu	Lys	Glu	Cys	Ser
		595					600					605			
Ser	Thr	Cys	Asp	Lys	Asn	Glu	Glu	Lys	Ser	Leu	Tyr	Lys	Gly	Leu	Leu
					615						620				
Ser	Val	Phe	Pro	Ala	Asp	Tyr	Asn	Pro	Leu	Leu	Phe	Thr	Glu	Lys	Glu
625					630					635					640
Ser	Gly	Lys	Leu	Gln	Val	Leu	Ser	Lys	Leu	Leu	Ala	Val	Ile	His	Glu
				645					650					655	
Leu	Arg	Pro	Thr	Glu	Lys	Val	Val	Leu	Val	Ser	Asn	Tyr	Thr	Gln	Thr
			660					665					670		
Leu	Asn	Ile	Leu	Gln	Glu	Val	Cys	Lys	Arg	His	Gly	Tyr	Ala	Tyr	Thr
		675					680					685			
Arg	Leu	Asp	Gly	Gln	Thr	Pro	Ile	Ser	Gln	Arg	Gln	Gln	Ile	Val	Asp
		690				695					700				
Gly	Phe	Asn	Ser	Gln	His	Ser	Ser	Phe	Phe	Ile	Phe	Leu	Leu	Ser	Ser
705					710					715					720
Lys	Ala	Gly	Gly	Val	Gly	Leu	Asn	Leu	Ile	Gly	Gly	Ser	His	Leu	Ile
				725					730					735	
Leu	Tyr	Asp	Ile	Asp	Trp	Asn	Pro	Ala	Thr	Asp	Ile	Gln	Ala	Met	Ser
			740					745					750		
Arg	Val	Trp	Arg	Asp	Gly	Gln	Lys	Tyr	Pro	Val	His	Ile	Tyr	Arg	Leu
		755				760						765			
Leu	Thr	Thr	Gly	Thr	Ile	Glu	Glu	Lys	Ile	Tyr	Gln	Arg	Gln	Ile	Ser
					775						780				
Lys	Gln	Gly	Leu	Cys	Gly	Ala	Val	Val	Asp	Leu	Thr	Lys	Thr	Ser	Glu
785					790				795						800
His	Ile	Gln	Phe	Ser	Val	Glu	Glu	Leu	Lys	Asn	Leu	Phe	Thr	Leu	His
				805					810					815	
Glu	Ser	Ser	Asp	Cys	Val	Thr	His	Asp	Leu	Leu	Asp	Cys	Glu	Cys	Thr
			820					825					830		
Gly	Glu	Glu	Val	His	Thr	Gly	Asp	Ser	Leu	Glu	Lys	Phe	Ile	Val	Ser
			835				840					845			
Arg	Asp	Cys	Gln	Leu	Gly	Pro	His	His	Gln	Lys	Ser	Asn	Ser	Leu	Lys
		850				855					860				
Pro	Leu	Ser	Met	Ser	Gln	Leu	Lys	Gln	Trp	Lys	His	Phe	Ser	Gly	Asp
865					870					875					880

His	Leu	Asn	Leu	Thr	Asp	Pro	Phe	Leu	Glu	Arg	Ile	Thr	Glu	Asn	Val
				885					890					895	
Ser	Phe	Ile	Phe	Gln	Asn	Ile	Thr	Thr	Gln	Ala	Thr	Gly	Thr		
			900					905					910		

<210> 196
 <211> 747
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >RAD54L|ENSG00000085999|ENST00000371975|2244

<400> 196

Met	Arg	Arg	Ser	Leu	Ala	Pro	Ser	Gln	Leu	Ala	Lys	Arg	Lys	Pro	Glu
1				5					10					15	
Gly	Arg	Ser	Cys	Asp	Asp	Glu	Asp	Trp	Gln	Pro	Gly	Leu	Val	Thr	Pro
			20					25					30		
Arg	Lys	Arg	Lys	Ser	Ser	Ser	Glu	Thr	Gln	Ile	Gln	Glu	Cys	Phe	Leu
			35				40					45			
Ser	Pro	Phe	Arg	Lys	Pro	Leu	Ser	Gln	Leu	Thr	Asn	Gln	Pro	Pro	Cys
			50			55					60				
Leu	Asp	Ser	Ser	Gln	His	Glu	Ala	Phe	Ile	Arg	Ser	Ile	Leu	Ser	Lys
65				70					75						80
Pro	Phe	Lys	Val	Pro	Ile	Pro	Asn	Tyr	Gln	Gly	Pro	Leu	Gly	Ser	Arg
			85						90					95	
Ala	Leu	Gly	Leu	Lys	Arg	Ala	Gly	Val	Arg	Arg	Ala	Leu	His	Asp	Pro
			100					105					110		
Leu	Glu	Lys	Asp	Ala	Leu	Val	Leu	Tyr	Glu	Pro	Pro	Pro	Leu	Ser	Ala
			115				120					125			
His	Asp	Gln	Leu	Lys	Leu	Asp	Lys	Glu	Lys	Leu	Pro	Val	His	Val	Val
						135					140				
Val	Asp	Pro	Ile	Leu	Ser	Lys	Val	Leu	Arg	Pro	His	Gln	Arg	Glu	Gly
145					150					155					160
Val	Lys	Phe	Leu	Trp	Glu	Cys	Val	Thr	Ser	Arg	Arg	Ile	Pro	Gly	Ser
			165						170					175	
His	Gly	Cys	Ile	Met	Ala	Asp	Glu	Met	Gly	Leu	Gly	Lys	Thr	Leu	Gln
			180					185					190		
Cys	Ile	Thr	Leu	Met	Trp	Thr	Leu	Leu	Arg	Gln	Ser	Pro	Glu	Cys	Lys
			195				200					205			
Pro	Glu	Ile	Asp	Lys	Ala	Val	Val	Val	Ser	Pro	Ser	Ser	Leu	Val	Lys
			210			215						220			
Asn	Trp	Tyr	Asn	Glu	Val	Gly	Lys	Trp	Leu	Gly	Gly	Arg	Ile	Gln	Pro
225					230					235					240
Leu	Ala	Ile	Asp	Gly	Gly	Ser	Lys	Asp	Glu	Ile	Asp	Gln	Lys	Leu	Glu
			245						250					255	
Gly	Phe	Met	Asn	Gln	Arg	Gly	Ala	Arg	Val	Ser	Ser	Pro	Ile	Leu	Ile
			260					265					270		
Ile	Ser	Tyr	Glu	Thr	Phe	Arg	Leu	His	Val	Gly	Val	Leu	Gln	Lys	Gly
			275				280					285			
Ser	Val	Gly	Leu	Val	Ile	Cys	Asp	Glu	Gly	His	Arg	Leu	Lys	Asn	Ser
			290			295					300				
Glu	Asn	Gln	Thr	Tyr	Gln	Ala	Leu	Asp	Ser	Leu	Asn	Thr	Ser	Arg	Arg

305					310					315				320
Val	Leu	Ile	Ser	Gly	Thr	Pro	Ile	Gln	Asn	Asp	Leu	Leu	Glu	Tyr
				325					330					335
Ser	Leu	Val	His	Phe	Val	Asn	Ser	Gly	Ile	Leu	Gly	Thr	Ala	His
			340					345					350	
Phe	Lys	Lys	His	Phe	Glu	Leu	Pro	Ile	Leu	Lys	Gly	Arg	Asp	Ala
		355					360					365		
Ala	Ser	Glu	Ala	Asp	Arg	Gln	Leu	Gly	Glu	Glu	Arg	Leu	Arg	Glu
	370					375					380			
Thr	Ser	Ile	Val	Asn	Arg	Cys	Leu	Ile	Arg	Arg	Thr	Ser	Asp	Ile
385					390					395				400
Ser	Lys	Tyr	Leu	Pro	Val	Lys	Ile	Glu	Gln	Val	Val	Cys	Cys	Arg
				405					410					415
Thr	Pro	Leu	Gln	Thr	Glu	Leu	Tyr	Lys	Arg	Phe	Leu	Arg	Gln	Ala
			420					425					430	
Pro	Ala	Glu	Glu	Leu	Leu	Glu	Gly	Lys	Met	Ser	Val	Ser	Ser	Leu
	435						440					445		
Ser	Ile	Thr	Ser	Leu	Lys	Lys	Leu	Cys	Asn	His	Pro	Ala	Leu	Ile
	450					455					460			
Asp	Lys	Cys	Val	Glu	Glu	Glu	Asp	Gly	Phe	Val	Gly	Ala	Leu	Asp
465					470					475				480
Phe	Pro	Pro	Gly	Tyr	Ser	Ser	Lys	Ala	Leu	Glu	Pro	Gln	Leu	Ser
				485					490					495
Lys	Met	Leu	Val	Leu	Asp	Tyr	Ile	Leu	Ala	Val	Thr	Arg	Ser	Arg
		500						505				510		
Ser	Asp	Lys	Val	Val	Leu	Val	Ser	Asn	Tyr	Thr	Gln	Thr	Leu	Asp
	515						520					525		
Phe	Glu	Lys	Leu	Cys	Arg	Ala	Arg	Arg	Tyr	Leu	Tyr	Val	Arg	Leu
	530					535					540			
Gly	Thr	Met	Ser	Ile	Lys	Lys	Arg	Ala	Lys	Val	Val	Glu	Arg	Phe
545					550					555				560
Ser	Pro	Ser	Ser	Pro	Asp	Phe	Val	Phe	Met	Leu	Ser	Ser	Lys	Ala
				565					570					575
Gly	Cys	Gly	Leu	Asn	Leu	Ile	Gly	Ala	Asn	Arg	Leu	Val	Met	Phe
			580					585					590	
Pro	Asp	Trp	Asn	Pro	Ala	Asn	Asp	Glu	Gln	Ala	Met	Ala	Arg	Val
	595						600					605		
Arg	Asp	Gly	Gln	Lys	Lys	Thr	Cys	Tyr	Ile	Tyr	Arg	Leu	Leu	Ser
	610					615					620			
Gly	Thr	Ile	Glu	Glu	Lys	Ile	Phe	Gln	Arg	Gln	Ser	His	Lys	Lys
625					630					635				640
Leu	Ser	Ser	Cys	Val	Val	Asp	Glu	Glu	Gln	Asp	Val	Glu	Arg	His
				645					650					655
Ser	Leu	Gly	Glu	Leu	Lys	Glu	Leu	Phe	Ile	Leu	Asp	Glu	Ala	Ser
			660					665					670	
Ser	Asp	Thr	His	Asp	Arg	Leu	His	Cys	Arg	Arg	Cys	Val	Asn	Ser
	675						680					685		
Gln	Ile	Arg	Pro	Pro	Pro	Asp	Gly	Ser	Asp	Cys	Thr	Ser	Asp	Leu
	690					695					700			
Gly	Trp	Asn	His	Cys	Thr	Asp	Lys	Trp	Gly	Leu	Arg	Asp	Glu	Val
705					710					715				720
Gln	Ala	Ala	Trp	Asp	Ala	Ala	Ser	Thr	Ala	Ile	Thr	Phe	Val	Phe
				725					730					735
Gln	Arg	Ser	His	Glu	Glu	Gln	Arg	Gly	Leu	Arg				

740

745

<210> 197

<211> 391

<212> PRT

<213> Homo sapiens

<220>

<223> >RAD9A|ENSG00000172613|ENST00000307980|1176

<400> 197

```

Met Lys Cys Leu Val Thr Gly Gly Asn Val Lys Val Leu Gly Lys Ala
1      5      10      15
Val His Ser Leu Ser Arg Ile Gly Asp Glu Leu Tyr Leu Glu Pro Leu
20     25     30
Glu Asp Gly Leu Ser Leu Arg Thr Val Asn Ser Ser Arg Ser Ala Tyr
35     40     45
Ala Cys Phe Leu Phe Ala Pro Leu Phe Phe Gln Gln Tyr Gln Ala Ala
50     55     60
Thr Pro Gly Gln Asp Leu Leu Arg Cys Lys Ile Leu Met Lys Ser Phe
65     70     75     80
Leu Ser Val Phe Arg Ser Leu Ala Met Leu Glu Lys Thr Val Glu Lys
85     90     95
Cys Cys Ile Ser Leu Asn Gly Arg Ser Ser Arg Leu Val Val Gln Leu
100    105    110
His Cys Lys Phe Gly Val Arg Lys Thr His Asn Leu Ser Phe Gln Asp
115    120    125
Cys Glu Ser Leu Gln Ala Val Phe Asp Pro Ala Ser Cys Pro His Met
130    135    140
Leu Arg Ala Pro Ala Arg Val Leu Gly Glu Ala Val Leu Pro Phe Ser
145    150    155    160
Pro Ala Leu Ala Glu Val Thr Leu Gly Ile Gly Arg Gly Arg Arg Val
165    170    175
Ile Leu Arg Ser Tyr His Glu Glu Glu Ala Asp Ser Thr Ala Lys Ala
180    185    190
Met Val Thr Glu Met Cys Leu Gly Glu Glu Asp Phe Gln Gln Leu Gln
195    200    205
Ala Gln Glu Gly Val Ala Ile Thr Phe Cys Leu Lys Glu Phe Arg Gly
210    215    220
Leu Leu Ser Phe Ala Glu Ser Ala Asn Leu Asn Leu Ser Ile His Phe
225    230    235    240
Asp Ala Pro Gly Arg Pro Ala Ile Phe Thr Ile Lys Asp Ser Leu Leu
245    250    255
Asp Gly His Phe Val Leu Ala Thr Leu Ser Asp Thr Asp Ser His Ser
260    265    270
Gln Asp Leu Gly Ser Pro Glu Arg His Gln Pro Val Pro Gln Leu Gln
275    280    285
Ala His Ser Thr Pro His Pro Asp Asp Phe Ala Asn Asp Asp Ile Asp
290    295    300
Ser Tyr Met Ile Ala Met Glu Thr Thr Ile Gly Asn Glu Gly Ser Arg
305    310    315    320
Val Leu Pro Ser Ile Ser Leu Ser Pro Gly Pro Gln Pro Pro Lys Ser
325    330    335

```


Pro Gly Pro His Ser Glu Glu Glu Asp Glu Ala Glu Pro Ser Thr Val
340 345 350
Pro Gly Thr Pro Pro Pro Lys Lys Phe Arg Ser Leu Phe Phe Gly Ser
355 360 365
Ile Leu Ala Pro Val Arg Ser Pro Gln Gly Pro Ser Pro Val Leu Ala
370 375 380
Glu Asp Ser Glu Gly Glu Gly
385 390

<210> 198
<211> 928
<212> PRT
<213> Homo sapiens

<220>
<223> >RB1|ENSG00000139687|ENST00000267163|2787

<400> 198
Met Pro Pro Lys Thr Pro Arg Lys Thr Ala Ala Thr Ala Ala Ala Ala
1 5 10 15
Ala Ala Glu Pro Pro Ala Pro Pro Pro Pro Pro Pro Glu Glu Asp
20 25 30
Pro Glu Gln Asp Ser Gly Pro Glu Asp Leu Pro Leu Val Arg Leu Glu
35 40 45
Phe Glu Glu Thr Glu Glu Pro Asp Phe Thr Ala Leu Cys Gln Lys Leu
50 55 60
Lys Ile Pro Asp His Val Arg Glu Arg Ala Trp Leu Thr Trp Glu Lys
65 70 75 80
Val Ser Ser Val Asp Gly Val Leu Gly Gly Tyr Ile Gln Lys Lys Lys
85 90 95
Glu Leu Trp Gly Ile Cys Ile Phe Ile Ala Ala Val Asp Leu Asp Glu
100 105 110
Met Ser Phe Thr Phe Thr Glu Leu Gln Lys Asn Ile Glu Ile Ser Val
115 120 125
His Lys Phe Phe Asn Leu Leu Lys Glu Ile Asp Thr Ser Thr Lys Val
130 135 140
Asp Asn Ala Met Ser Arg Leu Leu Lys Lys Tyr Asp Val Leu Phe Ala
145 150 155 160
Leu Phe Ser Lys Leu Glu Arg Thr Cys Glu Leu Ile Tyr Leu Thr Gln
165 170 175
Pro Ser Ser Ser Ile Ser Thr Glu Ile Asn Ser Ala Leu Val Leu Lys
180 185 190
Val Ser Trp Ile Thr Phe Leu Leu Ala Lys Gly Glu Val Leu Gln Met
195 200 205
Glu Asp Asp Leu Val Ile Ser Phe Gln Leu Met Leu Cys Val Leu Asp
210 215 220
Tyr Phe Ile Lys Leu Ser Pro Pro Met Leu Leu Lys Glu Pro Tyr Lys
225 230 235 240
Thr Ala Val Ile Pro Ile Asn Gly Ser Pro Arg Thr Pro Arg Arg Gly
245 250 255
Gln Asn Arg Ser Ala Arg Ile Ala Lys Gln Leu Glu Asn Asp Thr Arg
260 265 270
Ile Ile Glu Val Leu Cys Lys Glu His Glu Cys Asn Ile Asp Glu Val

705					710					715					720
Phe	Lys	Ile	Ile	Val	Thr	Ala	Tyr	Lys	Asp	Leu	Pro	His	Ala	Val	Gln
				725					730					735	
Glu	Thr	Phe	Lys	Arg	Val	Leu	Ile	Lys	Glu	Glu	Glu	Tyr	Asp	Ser	Ile
			740					745					750		
Ile	Val	Phe	Tyr	Asn	Ser	Val	Phe	Met	Gln	Arg	Leu	Lys	Thr	Asn	Ile
		755					760					765			
Leu	Gln	Tyr	Ala	Ser	Thr	Arg	Pro	Pro	Thr	Leu	Ser	Pro	Ile	Pro	His
	770					775					780				
Ile	Pro	Arg	Ser	Pro	Tyr	Lys	Phe	Pro	Ser	Ser	Pro	Leu	Arg	Ile	Pro
785					790				795					800	
Gly	Gly	Asn	Ile	Tyr	Ile	Ser	Pro	Leu	Lys	Ser	Pro	Tyr	Lys	Ile	Ser
			805						810					815	
Glu	Gly	Leu	Pro	Thr	Pro	Thr	Lys	Met	Thr	Pro	Arg	Ser	Arg	Ile	Leu
		820					825						830		
Val	Ser	Ile	Gly	Glu	Ser	Phe	Gly	Thr	Ser	Glu	Lys	Phe	Gln	Lys	Ile
	835					840						845			
Asn	Gln	Met	Val	Cys	Asn	Ser	Asp	Arg	Val	Leu	Lys	Arg	Ser	Ala	Glu
	850				855						860				
Gly	Ser	Asn	Pro	Pro	Lys	Pro	Leu	Lys	Lys	Leu	Arg	Phe	Asp	Ile	Glu
865					870					875				880	
Gly	Ser	Asp	Glu	Ala	Asp	Gly	Ser	Lys	His	Leu	Pro	Gly	Glu	Ser	Lys
			885					890						895	
Phe	Gln	Gln	Lys	Leu	Ala	Glu	Met	Thr	Ser	Thr	Arg	Thr	Arg	Met	Gln
		900					905					910			
Lys	Gln	Lys	Met	Asn	Asp	Ser	Met	Asp	Thr	Ser	Asn	Lys	Glu	Glu	Lys
	915					920						925			

<210> 199
 <211> 3130
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >REV3L|ENSG00000009413|ENST00000358835|9393

<400> 199															
Met	Phe	Ser	Val	Arg	Ile	Val	Thr	Ala	Asp	Tyr	Tyr	Met	Ala	Ser	Pro
1			5						10				15		
Leu	Gln	Gly	Leu	Asp	Thr	Cys	Gln	Ser	Pro	Leu	Thr	Gln	Ala	Pro	Val
			20				25						30		
Lys	Lys	Val	Pro	Val	Val	Arg	Val	Phe	Gly	Ala	Thr	Pro	Ala	Gly	Gln
	35					40						45			
Lys	Thr	Cys	Leu	His	Leu	His	Gly	Ile	Phe	Pro	Tyr	Leu	Tyr	Val	Pro
	50				55						60				
Tyr	Asp	Gly	Tyr	Gly	Gln	Gln	Pro	Glu	Ser	Tyr	Leu	Ser	Gln	Met	Ala
65				70					75					80	
Phe	Ser	Ile	Asp	Arg	Ala	Leu	Asn	Val	Ala	Leu	Gly	Asn	Pro	Ser	Ser
			85					90					95		
Thr	Ala	Gln	His	Val	Phe	Lys	Val	Ser	Leu	Val	Ser	Gly	Met	Pro	Phe
		100					105						110		

Tyr	Gly	Tyr	His	Glu	Lys	Glu	Arg	His	Phe	Met	Lys	Ile	Tyr	Leu	Tyr		
		115					120					125					
Asn	Pro	Thr	Met	Val	Lys	Arg	Ile	Cys	Glu	Leu	Leu	Gln	Ser	Gly	Ala		
	130					135					140						
Ile	Met	Asn	Lys	Phe	Tyr	Gln	Pro	His	Glu	Ala	His	Ile	Pro	Tyr	Leu		
145					150					155					160		
Leu	Gln	Leu	Phe	Ile	Asp	Tyr	Asn	Leu	Tyr	Gly	Met	Asn	Leu	Ile	Asn		
			165						170					175			
Leu	Ala	Ala	Val	Lys	Phe	Arg	Lys	Ala	Arg	Arg	Lys	Ser	Asn	Thr	Leu		
		180						185					190				
His	Ala	Thr	Gly	Ser	Cys	Lys	Asn	His	Leu	Ser	Gly	Asn	Ser	Leu	Ala		
	195						200					205					
Asp	Thr	Leu	Phe	Arg	Trp	Glu	Gln	Asp	Glu	Ile	Pro	Ser	Ser	Leu	Ile		
	210					215					220						
Leu	Glu	Gly	Val	Glu	Pro	Gln	Ser	Thr	Cys	Glu	Leu	Glu	Val	Asp	Ala		
225					230					235					240		
Val	Ala	Ala	Asp	Ile	Leu	Asn	Arg	Leu	Asp	Ile	Glu	Ala	Gln	Ile	Gly		
			245						250					255			
Gly	Asn	Pro	Gly	Leu	Gln	Ala	Ile	Trp	Glu	Asp	Glu	Lys	Gln	Arg	Arg		
		260					265						270				
Arg	Asn	Arg	Asn	Glu	Thr	Ser	Gln	Met	Ser	Gln	Pro	Glu	Ser	Gln	Asp		
	275						280					285					
His	Arg	Phe	Val	Pro	Ala	Thr	Glu	Ser	Glu	Lys	Lys	Phe	Gln	Lys	Arg		
	290					295					300						
Leu	Gln	Glu	Ile	Leu	Lys	Gln	Asn	Asp	Phe	Ser	Val	Thr	Leu	Ser	Gly		
305					310				315						320		
Ser	Val	Asp	Tyr	Ser	Asp	Gly	Ser	Gln	Glu	Phe	Ser	Ala	Glu	Leu	Thr		
			325						330					335			
Leu	His	Ser	Glu	Val	Leu	Ser	Pro	Glu	Met	Leu	Gln	Cys	Thr	Pro	Ala		
		340						345					350				
Asn	Met	Val	Glu	Val	His	Lys	Asp	Lys	Glu	Ser	Ser	Lys	Gly	His	Thr		
	355						360					365					
Arg	His	Lys	Val	Glu	Glu	Ala	Leu	Ile	Asn	Glu	Glu	Ala	Ile	Leu	Asn		
	370					375					380						
Leu	Met	Glu	Asn	Ser	Gln	Thr	Phe	Gln	Pro	Leu	Thr	Gln	Arg	Leu	Ser		
385					390					395					400		
Glu	Ser	Pro	Val	Phe	Met	Asp	Ser	Ser	Pro	Asp	Glu	Ala	Leu	Val	His		
			405						410					415			
Leu	Leu	Ala	Gly	Leu	Glu	Ser	Asp	Gly	Tyr	Arg	Gly	Glu	Arg	Asn	Arg		
		420						425				430					
Met	Pro	Ser	Pro	Cys	Arg	Ser	Phe	Gly	Asn	Asn	Lys	Tyr	Pro	Gln	Asn		
	435						440					445					
Ser	Asp	Asp	Glu	Glu	Asn	Glu	Pro	Gln	Ile	Glu	Lys	Glu	Glu	Met	Glu		
	450					455					460						
Leu	Ser	Leu	Val	Met	Ser	Gln	Arg	Trp	Asp	Ser	Asn	Ile	Glu	Glu	His		
465					470					475					480		
Cys	Ala	Lys	Lys	Arg	Ser	Leu	Cys	Arg	Asn	Thr	His	Arg	Ser	Ser	Thr		
			485						490					495			
Glu	Asp	Asp	Asp	Ser	Ser	Ser	Gly	Glu	Glu	Met	Glu	Trp	Ser	Asp	Asn		
		500					505						510				
Ser	Leu	Leu	Leu	Ala	Ser	Leu	Ser	Ile	Pro	Gln	Leu	Asp	Gly	Thr	Ala		
	515					520					525						
Asp	Glu	Asn	Ser	Asp	Asn	Pro	Leu	Asn	Asn	Glu	Asn	Ser	Arg	Thr	His		
	530					535					540						

Ser	Ser	Val	Ile	Ala	Thr	Ser	Lys	Leu	Ser	Val	Lys	Pro	Ser	Ile	Phe
545					550					555					560
His	Lys	Asp	Ala	Ala	Thr	Leu	Glu	Pro	Ser	Ser	Ser	Ala	Lys	Ile	Thr
				565					570						575
Phe	Gln	Cys	Lys	His	Thr	Ser	Ala	Leu	Ser	Ser	His	Val	Leu	Asn	Lys
			580					585						590	
Glu	Asp	Leu	Ile	Glu	Asp	Leu	Ser	Gln	Thr	Asn	Lys	Asn	Thr	Glu	Lys
		595				600						605			
Gly	Leu	Asp	Asn	Ser	Val	Thr	Ser	Phe	Thr	Asn	Glu	Ser	Thr	Tyr	Ser
	610					615					620				
Met	Lys	Tyr	Pro	Gly	Ser	Leu	Ser	Ser	Thr	Val	His	Ser	Glu	Asn	Ser
625					630					635					640
His	Lys	Glu	Asn	Ser	Lys	Lys	Glu	Ile	Leu	Pro	Val	Ser	Ser	Cys	Glu
				645					650						655
Ser	Ser	Ile	Phe	Asp	Tyr	Glu	Glu	Asp	Ile	Pro	Ser	Val	Thr	Arg	Gln
			660					665						670	
Val	Pro	Ser	Arg	Lys	Tyr	Thr	Asn	Ile	Arg	Lys	Ile	Glu	Lys	Asp	Ser
		675					680						685		
Pro	Phe	Ile	His	Met	His	Arg	His	Pro	Asn	Glu	Asn	Thr	Leu	Gly	Lys
	690					695						700			
Asn	Ser	Phe	Asn	Phe	Ser	Asp	Leu	Asn	His	Ser	Lys	Asn	Lys	Val	Ser
705					710					715					720
Ser	Glu	Gly	Asn	Glu	Lys	Gly	Asn	Ser	Thr	Ala	Leu	Ser	Ser	Leu	Phe
				725					730						735
Pro	Ser	Ser	Phe	Thr	Glu	Asn	Cys	Glu	Leu	Leu	Ser	Cys	Ser	Gly	Glu
			740					745					750		
Asn	Arg	Thr	Met	Val	His	Ser	Leu	Asn	Ser	Thr	Ala	Asp	Glu	Ser	Gly
		755					760					765			
Leu	Asn	Lys	Leu	Lys	Ile	Arg	Tyr	Glu	Glu	Phe	Gln	Glu	His	Lys	Thr
	770					775					780				
Glu	Lys	Pro	Ser	Leu	Ser	Gln	Gln	Ala	Ala	His	Tyr	Met	Phe	Phe	Pro
785					790					795					800
Ser	Val	Val	Leu	Ser	Asn	Cys	Leu	Thr	Arg	Pro	Gln	Lys	Leu	Ser	Pro
				805					810						815
Val	Thr	Tyr	Lys	Leu	Gln	Pro	Gly	Asn	Lys	Pro	Ser	Arg	Leu	Lys	Leu
			820					825						830	
Asn	Lys	Arg	Lys	Leu	Ala	Gly	His	Gln	Glu	Thr	Ser	Thr	Lys	Ser	Ser
		835					840						845		
Glu	Thr	Gly	Ser	Thr	Lys	Asp	Asn	Phe	Ile	Gln	Asn	Asn	Pro	Cys	Asn
	850					855					860				
Ser	Asn	Pro	Glu	Lys	Asp	Asn	Ala	Leu	Ala	Ser	Asp	Leu	Thr	Lys	Thr
865					870					875					880
Thr	Arg	Gly	Ala	Phe	Glu	Asn	Lys	Thr	Pro	Thr	Asp	Gly	Phe	Ile	Asp
				885					890						895
Cys	His	Phe	Gly	Asp	Gly	Thr	Leu	Glu	Thr	Glu	Gln	Ser	Phe	Gly	Leu
			900					905						910	
Tyr	Gly	Asn	Lys	Tyr	Thr	Leu	Arg	Ala	Lys	Arg	Lys	Val	Asn	Tyr	Glu
		915					920					925			
Thr	Glu	Asp	Ser	Glu	Ser	Ser	Phe	Val	Thr	His	Asn	Ser	Lys	Ile	Ser
	930					935						940			
Leu	Pro	His	Pro	Met	Glu	Ile	Gly	Glu	Ser	Leu	Asp	Gly	Thr	Leu	Lys
945					950					955					960
Ser	Arg	Lys	Arg	Arg	Lys	Met	Ser	Lys	Lys	Leu	Pro	Pro	Val	Ile	Ile
				965					970						975

Lys Tyr Ile Ile Ile Asn Arg Phe Arg Gly Arg Lys Asn Met Leu Val
 980 985 990
 Lys Leu Gly Lys Ile Asp Ser Lys Glu Lys Gln Val Ile Leu Thr Glu
 995 1000 1005
 Glu Lys Met Glu Leu Tyr Lys Lys Leu Ala Pro Leu Lys Asp Phe Trp
 1010 1015 1020
 Pro Lys Val Pro Asp Ser Pro Ala Thr Lys Tyr Pro Ile Tyr Pro Leu
 1025 1030 1035 1040
 Thr Pro Lys Lys Ser His Arg Arg Lys Ser Lys His Lys Ser Ala Lys
 1045 1050 1055
 Lys Lys Thr Gly Lys Gln Gln Arg Thr Asn Asn Glu Asn Ile Lys Arg
 1060 1065 1070
 Thr Leu Ser Phe Arg Lys Lys Arg Ser His Ala Ile Leu Ser Pro Pro
 1075 1080 1085
 Ser Pro Ser Tyr Asn Ala Glu Thr Glu Asp Cys Asp Leu Asn Tyr Ser
 1090 1095 1100
 Asp Val Met Ser Lys Leu Gly Phe Leu Ser Glu Arg Ser Thr Ser Pro
 1105 1110 1115 1120
 Ile Asn Ser Ser Pro Pro Arg Cys Trp Ser Pro Thr Asp Pro Arg Ala
 1125 1130 1135
 Glu Glu Ile Met Ala Ala Ala Glu Lys Glu Ala Met Leu Phe Lys Gly
 1140 1145 1150
 Pro Asn Val Tyr Lys Lys Thr Val Asn Ser Arg Ile Gly Lys Thr Ser
 1155 1160 1165
 Arg Ala Arg Ala Gln Ile Lys Lys Ser Lys Ala Lys Leu Ala Asn Pro
 1170 1175 1180
 Ser Ile Val Thr Lys Lys Arg Asn Lys Arg Asn Gln Thr Asn Lys Leu
 1185 1190 1195 1200
 Val Asp Asp Gly Lys Lys Lys Pro Arg Ala Lys Gln Lys Thr Asn Glu
 1205 1210 1215
 Lys Gly Thr Ser Arg Lys His Thr Thr Leu Lys Asp Glu Lys Ile Lys
 1220 1225 1230
 Ser Gln Ser Gly Ala Glu Val Lys Phe Val Leu Lys His Gln Asn Val
 1235 1240 1245
 Ser Glu Phe Ala Ser Ser Ser Gly Gly Ser Gln Leu Leu Phe Lys Gln
 1250 1255 1260
 Lys Asp Met Pro Leu Met Gly Ser Ala Val Asp His Pro Leu Ser Ala
 1265 1270 1275 1280
 Ser Leu Pro Thr Gly Ile Asn Ala Gln Gln Lys Leu Ser Gly Cys Phe
 1285 1290 1295
 Ser Ser Phe Leu Glu Ser Lys Lys Ser Val Asp Leu Gln Thr Phe Pro
 1300 1305 1310
 Ser Ser Arg Asp Asp Leu His Pro Ser Val Val Cys Asn Ser Ile Gly
 1315 1320 1325
 Pro Gly Val Ser Lys Ile Asn Val Gln Arg Pro His Asn Gln Ser Ala
 1330 1335 1340
 Met Phe Thr Leu Lys Glu Ser Thr Leu Ile Gln Lys Asn Ile Phe Asp
 1345 1350 1355 1360
 Leu Ser Asn His Leu Ser Gln Val Ala Gln Asn Thr Gln Ile Ser Ser
 1365 1370 1375
 Gly Met Ser Ser Lys Ile Glu Asp Asn Ala Asn Asn Ile Gln Arg Asn
 1380 1385 1390
 Tyr Leu Ser Ser Ile Gly Lys Leu Ser Glu Tyr Arg Asn Ser Leu Glu
 1395 1400 1405

Ser Lys Leu Asp Gln Ala Tyr Thr Pro Asn Phe Leu His Cys Lys Asp
 1410 1415 1420
 Ser Gln Gln Gln Ile Val Cys Ile Ala Glu Gln Ser Lys His Ser Glu
 1425 1430 1435 1440
 Thr Cys Ser Pro Gly Asn Thr Ala Ser Glu Glu Ser Gln Met Pro Asn
 1445 1450 1455
 Asn Cys Phe Val Thr Ser Leu Arg Ser Pro Ile Lys Gln Ile Ala Trp
 1460 1465 1470
 Glu Gln Lys Gln Arg Gly Phe Ile Leu Asp Met Ser Asn Phe Lys Pro
 1475 1480 1485
 Glu Arg Val Lys Pro Arg Ser Leu Ser Glu Ala Ile Ser Gln Thr Lys
 1490 1495 1500
 Ala Leu Ser Gln Cys Lys Asn Arg Asn Val Ser Thr Pro Ser Ala Phe
 1505 1510 1515 1520
 Gly Glu Gly Gln Ser Gly Leu Ala Val Leu Lys Glu Leu Leu Gln Lys
 1525 1530 1535
 Arg Gln Gln Lys Ala Gln Asn Ala Asn Thr Thr Gln Asp Pro Leu Ser
 1540 1545 1550
 Asn Lys His Gln Pro Asn Lys Asn Ile Ser Gly Ser Leu Glu His Asn
 1555 1560 1565
 Lys Ala Asn Lys Arg Thr Arg Ser Val Thr Ser Pro Arg Lys Pro Arg
 1570 1575 1580
 Thr Pro Arg Ser Thr Lys Gln Lys Glu Lys Ile Pro Lys Leu Leu Lys
 1585 1590 1595 1600
 Val Asp Ser Leu Asn Leu Gln Asn Ser Ser Gln Leu Asp Asn Ser Val
 1605 1610 1615
 Ser Asp Asp Ser Pro Ile Phe Phe Ser Asp Pro Gly Phe Glu Ser Cys
 1620 1625 1630
 Tyr Ser Leu Glu Asp Ser Leu Ser Pro Glu His Asn Tyr Asn Phe Asp
 1635 1640 1645
 Ile Asn Thr Ile Gly Gln Thr Gly Phe Cys Ser Phe Tyr Ser Gly Ser
 1650 1655 1660
 Gln Phe Val Pro Ala Asp Gln Asn Leu Pro Gln Lys Phe Leu Ser Asp
 1665 1670 1675 1680
 Ala Val Gln Asp Leu Phe Pro Gly Gln Ala Ile Glu Lys Asn Glu Phe
 1685 1690 1695
 Leu Ser His Asp Asn Gln Lys Cys Asp Glu Asp Lys His His Thr Thr
 1700 1705 1710
 Asp Ser Ala Ser Trp Ile Arg Ser Gly Thr Leu Ser Pro Glu Ile Phe
 1715 1720 1725
 Glu Lys Ser Thr Ile Asp Ser Asn Glu Asn Arg Arg His Asn Gln Trp
 1730 1735 1740
 Lys Asn Ser Phe His Pro Leu Thr Thr Arg Ser Asn Ser Ile Met Asp
 1745 1750 1755 1760
 Ser Phe Cys Val Gln Gln Ala Glu Asp Cys Leu Ser Glu Lys Ser Arg
 1765 1770 1775
 Leu Asn Arg Ser Ser Val Ser Lys Glu Val Phe Leu Ser Leu Pro Gln
 1780 1785 1790
 Pro Asn Asn Ser Asp Trp Ile Gln Gly His Thr Arg Lys Glu Met Gly
 1795 1800 1805
 Gln Ser Leu Asp Ser Ala Asn Thr Ser Phe Thr Ala Ile Leu Ser Ser
 1810 1815 1820
 Pro Asp Gly Glu Leu Val Asp Val Ala Cys Glu Asp Leu Glu Leu Tyr
 1825 1830 1835 1840

Val	Ser	Arg	Asn	Asn	Asp	Met	Leu	Thr	Pro	Thr	Pro	Asp	Ser	Ser	Pro	1845	1850	1855
Arg	Ser	Thr	Ser	Ser	Pro	Ser	Gln	Ser	Lys	Asn	Gly	Ser	Phe	Thr	Pro	1860	1865	1870
Arg	Thr	Ala	Asn	Ile	Leu	Lys	Pro	Leu	Met	Ser	Pro	Pro	Ser	Arg	Glu	1875	1880	1885
Glu	Ile	Met	Ala	Thr	Leu	Leu	Asp	His	Asp	Leu	Ser	Glu	Thr	Ile	Tyr	1890	1895	1900
Gln	Glu	Pro	Phe	Cys	Ser	Asn	Pro	Ser	Asp	Val	Pro	Glu	Lys	Pro	Arg	1905	1910	1915
Glu	Ile	Gly	Gly	Arg	Leu	Leu	Met	Val	Glu	Thr	Arg	Leu	Ala	Asn	Asp	1925	1930	1935
Leu	Ala	Glu	Phe	Glu	Gly	Asp	Phe	Ser	Leu	Glu	Gly	Leu	Arg	Leu	Trp	1940	1945	1950
Lys	Thr	Ala	Phe	Ser	Ala	Met	Thr	Gln	Asn	Pro	Arg	Pro	Gly	Ser	Pro	1955	1960	1965
Leu	Arg	Ser	Gly	Gln	Gly	Val	Val	Asn	Lys	Gly	Ser	Ser	Asn	Ser	Pro	1970	1975	1980
Lys	Met	Val	Glu	Asp	Lys	Lys	Ile	Val	Ile	Met	Pro	Cys	Lys	Cys	Ala	1985	1990	1995
Pro	Ser	Arg	Gln	Leu	Val	Gln	Val	Trp	Leu	Gln	Ala	Lys	Glu	Glu	Tyr	2005	2010	2015
Glu	Arg	Ser	Lys	Lys	Leu	Pro	Lys	Thr	Lys	Pro	Thr	Gly	Val	Val	Lys	2020	2025	2030
Ser	Ala	Glu	Asn	Phe	Ser	Ser	Ser	Val	Asn	Pro	Asp	Asp	Lys	Pro	Val	2035	2040	2045
Val	Pro	Pro	Lys	Met	Asp	Val	Ser	Pro	Cys	Ile	Leu	Pro	Thr	Thr	Ala	2050	2055	2060
His	Thr	Lys	Glu	Asp	Val	Asp	Asn	Ser	Gln	Ile	Ala	Leu	Gln	Ala	Pro	2065	2070	2075
Thr	Thr	Gly	Cys	Ser	Gln	Thr	Ala	Ser	Glu	Ser	Gln	Met	Leu	Pro	Pro	2085	2090	2095
Val	Ala	Ser	Ala	Ser	Asp	Pro	Glu	Lys	Asp	Glu	Asp	Asp	Asp	Asp	Asn	2100	2105	2110
Tyr	Tyr	Ile	Ser	Tyr	Ser	Ser	Pro	Asp	Ser	Pro	Val	Ile	Pro	Pro	Trp	2115	2120	2125
Gln	Gln	Pro	Ile	Ser	Pro	Asp	Ser	Lys	Ala	Leu	Asn	Gly	Asp	Asp	Arg	2130	2135	2140
Pro	Ser	Ser	Pro	Val	Glu	Glu	Leu	Pro	Ser	Leu	Ala	Phe	Glu	Asn	Phe	2145	2150	2155
Leu	Lys	Pro	Ile	Lys	Asp	Gly	Ile	Gln	Lys	Ser	Pro	Cys	Ser	Glu	Pro	2165	2170	2175
Gln	Glu	Pro	Leu	Val	Ile	Ser	Pro	Ile	Asn	Thr	Arg	Ala	Arg	Thr	Gly	2180	2185	2190
Lys	Cys	Glu	Ser	Leu	Cys	Phe	His	Ser	Thr	Pro	Ile	Ile	Gln	Arg	Lys	2195	2200	2205
Leu	Leu	Glu	Arg	Leu	Pro	Glu	Ala	Pro	Gly	Leu	Ser	Pro	Leu	Ser	Thr	2210	2215	2220
Glu	Pro	Lys	Thr	Gln	Lys	Leu	Ser	Asn	Lys	Lys	Gly	Ser	Asn	Thr	Asp	2225	2230	2235
Thr	Leu	Arg	Arg	Val	Leu	Leu	Thr	Gln	Ala	Lys	Asn	Gln	Phe	Ala	Ala	2245	2250	2255
Val	Asn	Thr	Pro	Gln	Lys	Glu	Thr	Ser	Gln	Ile	Asp	Gly	Pro	Ser	Leu	2260	2265	2270

Asn Asn Thr Tyr Gly Phe Lys Val Ser Ile Gln Asn Leu Gln Glu Ala
 2275 2280 2285
 Lys Ala Leu His Glu Ile Gln Asn Leu Thr Leu Ile Ser Val Glu Leu
 2290 2295 2300
 His Ala Arg Thr Arg Arg Asp Leu Glu Pro Asp Pro Glu Phe Asp Pro
 2305 2310 2315 2320
 Ile Cys Ala Leu Phe Tyr Cys Ile Ser Ser Asp Thr Pro Leu Pro Asp
 2325 2330 2335
 Thr Glu Lys Thr Glu Leu Thr Gly Val Ile Val Ile Asp Lys Asp Lys
 2340 2345 2350
 Thr Val Phe Ser Gln Asp Ile Arg Tyr Gln Thr Pro Leu Leu Ile Arg
 2355 2360 2365
 Ser Gly Ile Thr Gly Leu Glu Val Thr Tyr Ala Ala Asp Glu Lys Ala
 2370 2375 2380
 Leu Phe His Glu Ile Ala Asn Ile Ile Lys Arg Tyr Asp Pro Asp Ile
 2385 2390 2395 2400
 Leu Leu Gly Tyr Glu Ile Gln Met His Ser Trp Gly Tyr Leu Leu Gln
 2405 2410 2415
 Arg Ala Ala Ala Leu Ser Ile Asp Leu Cys Arg Met Ile Ser Arg Val
 2420 2425 2430
 Pro Asp Asp Lys Ile Glu Asn Arg Phe Ala Ala Glu Arg Asp Glu Tyr
 2435 2440 2445
 Gly Ser Tyr Thr Met Ser Glu Ile Asn Ile Val Gly Arg Ile Thr Leu
 2450 2455 2460
 Asn Leu Trp Arg Ile Met Arg Asn Glu Val Ala Leu Thr Asn Tyr Thr
 2465 2470 2475 2480
 Phe Glu Asn Val Ser Phe His Val Leu His Gln Arg Phe Pro Leu Phe
 2485 2490 2495
 Thr Phe Arg Val Leu Ser Asp Trp Phe Asp Asn Lys Thr Asp Leu Tyr
 2500 2505 2510
 Arg Trp Lys Met Val Asp His Tyr Val Ser Arg Val Arg Gly Asn Leu
 2515 2520 2525
 Gln Met Leu Glu Gln Leu Asp Leu Ile Gly Lys Thr Ser Glu Met Ala
 2530 2535 2540
 Arg Leu Phe Gly Ile Gln Phe Leu His Val Leu Thr Arg Gly Ser Gln
 2545 2550 2555 2560
 Tyr Arg Val Glu Ser Met Met Leu Arg Ile Ala Lys Pro Met Asn Tyr
 2565 2570 2575
 Ile Pro Val Thr Pro Ser Val Gln Gln Arg Ser Gln Met Arg Ala Pro
 2580 2585 2590
 Gln Cys Val Pro Leu Ile Met Glu Pro Glu Ser Arg Phe Tyr Ser Asn
 2595 2600 2605
 Ser Val Leu Val Leu Asp Phe Gln Ser Leu Tyr Pro Ser Ile Val Ile
 2610 2615 2620
 Ala Tyr Asn Tyr Cys Phe Ser Thr Cys Leu Gly His Val Glu Asn Leu
 2625 2630 2635 2640
 Gly Lys Tyr Asp Glu Phe Lys Phe Gly Cys Thr Ser Leu Arg Val Pro
 2645 2650 2655
 Pro Asp Leu Leu Tyr Gln Val Arg His Asp Ile Thr Val Ser Pro Asn
 2660 2665 2670
 Gly Val Ala Phe Val Lys Pro Ser Val Arg Lys Gly Val Leu Pro Arg
 2675 2680 2685
 Met Leu Glu Glu Ile Leu Lys Thr Arg Phe Met Val Lys Gln Ser Met
 2690 2695 2700

Lys Ala Tyr Lys Gln Asp Arg Ala Leu Ser Arg Met Leu Asp Ala Arg
 2705 2710 2715 2720
 Gln Leu Gly Leu Lys Leu Ile Ala Asn Val Thr Phe Gly Tyr Thr Ser
 2725 2730 2735
 Ala Asn Phe Ser Gly Arg Met Pro Cys Ile Glu Val Gly Asp Ser Ile
 2740 2745 2750
 Val His Lys Ala Arg Glu Thr Leu Glu Arg Ala Ile Lys Leu Val Asn
 2755 2760 2765
 Asp Thr Lys Lys Trp Gly Ala Arg Val Val Tyr Gly Asp Thr Asp Ser
 2770 2775 2780
 Met Phe Val Leu Leu Lys Gly Ala Thr Lys Glu Gln Ser Phe Lys Ile
 2785 2790 2795 2800
 Gly Gln Glu Ile Ala Glu Ala Val Thr Ala Thr Asn Pro Lys Pro Val
 2805 2810 2815
 Lys Leu Lys Phe Glu Lys Val Tyr Leu Pro Cys Val Leu Gln Thr Lys
 2820 2825 2830
 Lys Arg Tyr Val Gly Tyr Met Tyr Glu Thr Leu Asp Gln Lys Asp Pro
 2835 2840 2845
 Val Phe Asp Ala Lys Gly Ile Glu Thr Val Arg Arg Asp Ser Cys Pro
 2850 2855 2860
 Ala Val Ser Lys Ile Leu Glu Arg Ser Leu Lys Leu Leu Phe Glu Thr
 2865 2870 2875 2880
 Arg Asp Ile Ser Leu Ile Lys Gln Tyr Val Gln Arg Gln Cys Met Lys
 2885 2890 2895
 Leu Leu Glu Gly Lys Ala Ser Ile Gln Asp Phe Ile Phe Ala Lys Glu
 2900 2905 2910
 Tyr Arg Gly Ser Phe Ser Tyr Lys Pro Gly Ala Cys Val Pro Ala Leu
 2915 2920 2925
 Glu Leu Thr Arg Lys Met Leu Thr Tyr Asp Arg Arg Ser Glu Pro Gln
 2930 2935 2940
 Val Gly Glu Arg Val Pro Tyr Val Ile Ile Tyr Gly Thr Pro Gly Val
 2945 2950 2955 2960
 Pro Leu Ile Gln Leu Val Arg Arg Pro Val Glu Val Leu Gln Asp Pro
 2965 2970 2975
 Thr Leu Arg Leu Asn Ala Thr Tyr Tyr Ile Thr Lys Gln Ile Leu Pro
 2980 2985 2990
 Pro Leu Ala Arg Ile Phe Ser Leu Ile Gly Ile Asp Val Phe Ser Trp
 2995 3000 3005
 Tyr His Glu Leu Pro Arg Ile His Lys Ala Thr Ser Ser Ser Arg Ser
 3010 3015 3020
 Glu Pro Glu Gly Arg Lys Gly Thr Ile Ser Gln Tyr Phe Thr Thr Leu
 3025 3030 3035 3040
 His Cys Pro Val Cys Asp Asp Leu Thr Gln His Gly Ile Cys Ser Lys
 3045 3050 3055
 Cys Arg Ser Gln Pro Gln His Val Ala Val Ile Leu Asn Gln Glu Ile
 3060 3065 3070
 Arg Glu Leu Glu Arg Gln Gln Glu Gln Leu Val Lys Ile Cys Lys Asn
 3075 3080 3085
 Cys Thr Gly Cys Phe Asp Arg His Ile Pro Cys Val Ser Leu Asn Cys
 3090 3095 3100
 Pro Val Leu Phe Lys Leu Ser Arg Val Asn Arg Glu Leu Ser Lys Ala
 3105 3110 3115 3120
 Pro Tyr Leu Arg Gln Leu Leu Asp Gln Phe
 3125 3130

<210> 200
 <211> 616
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >RPA1|ENSG00000132383|ENST00000254719|1851

<400> 200
 Met Val Gly Gln Leu Ser Glu Gly Ala Ile Ala Ala Ile Met Gln Lys
 1 5 10 15
 Gly Asp Thr Asn Ile Lys Pro Ile Leu Gln Val Ile Asn Ile Arg Pro
 20 25 30
 Ile Thr Thr Gly Asn Ser Pro Pro Arg Tyr Arg Leu Leu Met Ser Asp
 35 40 45
 Gly Leu Asn Thr Leu Ser Ser Phe Met Leu Ala Thr Gln Leu Asn Pro
 50 55 60
 Leu Val Glu Glu Glu Gln Leu Ser Ser Asn Cys Val Cys Gln Ile His
 65 70 75 80
 Arg Phe Ile Val Asn Thr Leu Lys Asp Gly Arg Arg Val Val Ile Leu
 85 90 95
 Met Glu Leu Glu Val Leu Lys Ser Ala Glu Ala Val Gly Val Lys Ile
 100 105 110
 Gly Asn Pro Val Pro Tyr Asn Glu Gly Leu Gly Gln Pro Gln Val Ala
 115 120 125
 Pro Pro Ala Pro Ala Ala Ser Pro Ala Ala Ser Ser Arg Pro Gln Pro
 130 135 140
 Gln Asn Gly Ser Ser Gly Met Gly Ser Thr Val Ser Lys Ala Tyr Gly
 145 150 155 160
 Ala Ser Lys Thr Phe Gly Lys Ala Ala Gly Pro Ser Leu Ser His Thr
 165 170 175
 Ser Gly Gly Thr Gln Ser Lys Val Val Pro Ile Ala Ser Leu Thr Pro
 180 185 190
 Tyr Gln Ser Lys Trp Thr Ile Cys Ala Arg Val Thr Asn Lys Ser Gln
 195 200 205
 Ile Arg Thr Trp Ser Asn Ser Arg Gly Glu Gly Lys Leu Phe Ser Leu
 210 215 220
 Glu Leu Val Asp Glu Ser Gly Glu Ile Arg Ala Thr Ala Phe Asn Glu
 225 230 235 240
 Gln Val Asp Lys Phe Phe Pro Leu Ile Glu Val Asn Lys Val Tyr Tyr
 245 250 255
 Phe Ser Lys Gly Thr Leu Lys Ile Ala Asn Lys Gln Phe Thr Ala Val
 260 265 270
 Lys Asn Asp Tyr Glu Met Thr Phe Asn Asn Glu Thr Ser Val Met Pro
 275 280 285
 Cys Glu Asp Asp His His Leu Pro Thr Val Gln Phe Asp Phe Thr Gly
 290 295 300
 Ile Asp Asp Leu Glu Asn Lys Ser Lys Asp Ser Leu Val Asp Ile Ile
 305 310 315 320
 Gly Ile Cys Lys Ser Tyr Glu Asp Ala Thr Lys Ile Thr Val Arg Ser
 325 330 335
 Asn Asn Arg Glu Val Ala Lys Arg Asn Ile Tyr Leu Met Asp Thr Ser

Arg	Ile	Gly	Asn	Val	Glu	Ile	Ser	Gln	Val	Thr	Ile	Val	Gly	Ile	Ile
65					70					75					80
Arg	His	Ala	Glu	Lys	Ala	Pro	Thr	Asn	Ile	Val	Tyr	Lys	Ile	Asp	Asp
				85					90					95	
Met	Thr	Ala	Ala	Pro	Met	Asp	Val	Arg	Gln	Trp	Val	Asp	Thr	Asp	Asp
			100					105					110		
Thr	Ser	Ser	Glu	Asn	Thr	Val	Val	Pro	Pro	Glu	Thr	Tyr	Val	Lys	Val
		115					120					125			
Ala	Gly	His	Leu	Arg	Ser	Phe	Gln	Asn	Lys	Lys	Ser	Leu	Val	Ala	Phe
	130					135					140				
Lys	Ile	Met	Pro	Leu	Glu	Asp	Met	Asn	Glu	Phe	Thr	Thr	His	Ile	Leu
145					150					155					160
Glu	Val	Ile	Asn	Ala	His	Met	Val	Leu	Ser	Lys	Ala	Asn	Ser	Gln	Pro
				165					170					175	
Ser	Ala	Gly	Arg	Ala	Pro	Ile	Ser	Asn	Pro	Gly	Met	Ser	Glu	Ala	Gly
			180					185					190		
Asn	Phe	Gly	Gly	Asn	Ser	Phe	Met	Pro	Ala	Asn	Gly	Leu	Thr	Val	Ala
		195					200					205			
Gln	Asn	Gln	Val	Leu	Asn	Leu	Ile	Lys	Ala	Cys	Pro	Arg	Pro	Glu	Gly
	210					215					220				
Leu	Asn	Phe	Gln	Asp	Leu	Lys	Asn	Gln	Leu	Lys	His	Met	Ser	Val	Ser
225					230					235					240
Ser	Ile	Lys	Gln	Ala	Val	Asp	Phe	Leu	Ser	Asn	Glu	Gly	His	Ile	Tyr
			245					250						255	
Ser	Thr	Val	Asp	Asp	Asp	His	Phe	Lys	Ser	Thr	Asp	Ala	Glu		
			260					265					270		

<210> 202

<211> 1834

<212> PRT

<213> Homo sapiens

<220>

<223> >SLX4|ENSG00000188827|ENST00000294008|5505

<400> 202

Met	Lys	Leu	Ser	Val	Asn	Glu	Ala	Gln	Leu	Gly	Phe	Tyr	Leu	Gly	Ser
1				5					10					15	
Leu	Ser	His	Leu	Ser	Ala	Cys	Pro	Gly	Ile	Asp	Pro	Arg	Ser	Ser	Glu
		20						25					30		
Asp	Gln	Pro	Glu	Ser	Leu	Lys	Thr	Gly	Gln	Met	Met	Asp	Glu	Ser	Asp
		35					40					45			
Glu	Asp	Phe	Lys	Glu	Leu	Cys	Ala	Ser	Phe	Phe	Gln	Arg	Val	Lys	Lys
	50					55					60				
His	Gly	Ile	Lys	Glu	Val	Ser	Gly	Glu	Arg	Lys	Thr	Gln	Lys	Ala	Ala
65				70					75						80
Ser	Asn	Gly	Thr	Gln	Ile	Arg	Ser	Lys	Leu	Lys	Arg	Thr	Lys	Gln	Thr
			85					90						95	
Ala	Thr	Lys	Thr	Lys	Thr	Leu	Gln	Gly	Pro	Ala	Glu	Lys	Lys	Pro	Pro
		100						105					110		
Ser	Gly	Ser	Gln	Ala	Pro	Arg	Thr	Lys	Lys	Gln	Arg	Val	Thr	Lys	Trp
		115					120					125			
Gln	Ala	Ser	Glu	Pro	Ala	His	Ser	Val	Asn	Gly	Glu	Gly	Gly	Val	Leu

130	135	140																	
Ala	Ser	Ala	Pro	Asp	Pro	Pro	Val	Leu	Arg	Glu	Thr	Ala	Gln	Asn	Thr				
145					150					155					160				
Gln	Thr	Gly	Asn	Gln	Gln	Glu	Pro	Ser	Pro	Asn	Leu	Ser	Arg	Glu	Lys				
				165					170						175				
Thr	Arg	Glu	Asn	Val	Pro	Asn	Ser	Asp	Ser	Gln	Pro	Pro	Pro	Ser	Cys				
				180					185						190				
Leu	Thr	Thr	Ala	Val	Pro	Ser	Pro	Ser	Lys	Pro	Arg	Thr	Ala	Gln	Leu				
				195			200						205						
Val	Leu	Gln	Arg	Met	Gln	Gln	Phe	Lys	Arg	Ala	Asp	Pro	Glu	Arg	Leu				
						215					220								
Arg	His	Ala	Ser	Glu	Glu	Cys	Ser	Leu	Glu	Ala	Ala	Arg	Glu	Glu	Asn				
225					230					235					240				
Val	Pro	Lys	Asp	Pro	Gln	Glu	Glu	Met	Met	Ala	Gly	Asn	Val	Tyr	Gly				
				245					250						255				
Leu	Gly	Pro	Pro	Ala	Pro	Glu	Ser	Asp	Ala	Ala	Val	Ala	Leu	Thr	Leu				
				260				265						270					
Gln	Gln	Glu	Phe	Ala	Arg	Val	Gly	Ala	Ser	Ala	His	Asp	Asp	Ser	Leu				
				275			280					285							
Glu	Glu	Lys	Gly	Leu	Phe	Phe	Cys	Gln	Ile	Cys	Gln	Lys	Asn	Leu	Ser				
				290		295					300								
Ala	Met	Asn	Val	Thr	Arg	Arg	Glu	Gln	His	Val	Asn	Arg	Cys	Leu	Asp				
305					310					315					320				
Glu	Ala	Glu	Lys	Thr	Leu	Arg	Pro	Ser	Val	Pro	Gln	Ile	Pro	Glu	Cys				
				325				330						335					
Pro	Ile	Cys	Gly	Lys	Pro	Phe	Leu	Thr	Leu	Lys	Ser	Arg	Thr	Ser	His				
				340				345						350					
Leu	Lys	Gln	Cys	Ala	Val	Lys	Met	Glu	Val	Gly	Pro	Gln	Leu	Leu	Leu				
				355			360					365							
Gln	Ala	Val	Arg	Leu	Gln	Thr	Ala	Gln	Pro	Glu	Gly	Ser	Ser	Ser	Pro				
				370		375					380								
Pro	Met	Phe	Ser	Phe	Ser	Asp	His	Ser	Arg	Gly	Leu	Lys	Arg	Arg	Gly				
385					390					395					400				
Pro	Thr	Ser	Lys	Lys	Glu	Pro	Arg	Lys	Arg	Arg	Lys	Val	Asp	Glu	Ala				
				405				410						415					
Pro	Ser	Glu	Asp	Leu	Leu	Val	Ala	Met	Ala	Leu	Ser	Arg	Ser	Glu	Met				
				420			425							430					
Glu	Pro	Gly	Ala	Ala	Val	Pro	Ala	Leu	Arg	Leu	Glu	Ser	Ala	Phe	Ser				
				435			440					445							
Glu	Arg	Ile	Arg	Pro	Glu	Ala	Glu	Asn	Lys	Ser	Arg	Lys	Lys	Lys	Pro				
						455				460									
Pro	Val	Ser	Pro	Pro	Leu	Leu	Leu	Val	Gln	Asp	Ser	Glu	Thr	Thr	Gly				
465					470					475					480				
Arg	Gln	Ile	Glu	Asp	Arg	Val	Ala	Leu	Leu	Leu	Ser	Glu	Glu	Val	Glu				
				485				490						495					
Leu	Ser	Ser	Thr	Pro	Pro	Leu	Pro	Ala	Ser	Arg	Ile	Leu	Lys	Glu	Gly				
				500			505							510					
Trp	Glu	Arg	Ala	Gly	Gln	Cys	Pro	Pro	Pro	Pro	Glu	Arg	Lys	Gln	Ser				
				515			520					525							
Phe	Leu	Trp	Glu	Gly	Ser	Ala	Leu	Thr	Gly	Ala	Trp	Ala	Met	Glu	Asp				
						535				540									
Phe	Tyr	Thr	Ala	Arg	Leu	Val	Pro	Pro	Leu	Val	Pro	Gln	Arg	Pro	Ala				
545					550					555					560				
Gln	Gly	Leu	Met	Gln	Glu	Pro	Val	Pro	Pro	Leu	Val	Pro	Pro	Glu	His				

995					1000					1005						
Gly	Ala	Val	Arg	Glu	Arg	Gly	Leu	Glu	Val	Ser	His	Arg	Leu	Ala	Pro	
1010					1015					1020						
Trp	Gln	Ala	Ser	Pro	Pro	His	Pro	Cys	Arg	Phe	Leu	Leu	Gly	Pro	Pro	
1025					1030					1035					1040	
Gln	Gly	Gly	Ser	Pro	Arg	Gly	Ser	His	His	Thr	Ser	Gly	Ser	Ser	Leu	
1045					1050					1055						
Ser	Thr	Pro	Arg	Ser	Arg	Gly	Gly	Thr	Ser	Gln	Val	Gly	Ser	Pro	Thr	
1060					1065					1070						
Leu	Leu	Ser	Pro	Ala	Val	Pro	Ser	Lys	Gln	Lys	Arg	Asp	Arg	Ser	Ile	
1075					1080					1085						
Leu	Thr	Leu	Ser	Lys	Glu	Pro	Gly	His	Gln	Lys	Gly	Lys	Glu	Arg	Arg	
1090					1095					1100						
Ser	Val	Leu	Glu	Cys	Arg	Asn	Lys	Gly	Val	Leu	Met	Phe	Pro	Glu	Lys	
1105					1110					1115					1120	
Ser	Pro	Ser	Ile	Asp	Leu	Thr	Gln	Ser	Asn	Pro	Asp	His	Ser	Ser	Ser	
1125					1130					1135						
Arg	Ser	Gln	Lys	Ser	Ser	Ser	Lys	Leu	Asn	Glu	Glu	Asp	Glu	Val	Ile	
1140					1145					1150						
Leu	Leu	Leu	Asp	Ser	Asp	Glu	Glu	Leu	Glu	Leu	Glu	Gln	Thr	Lys	Met	
1155					1160					1165						
Lys	Ser	Ile	Ser	Ser	Asp	Pro	Leu	Glu	Glu	Lys	Lys	Ala	Leu	Glu	Ile	
1170					1175					1180						
Ser	Pro	Arg	Ser	Cys	Glu	Leu	Phe	Ser	Ile	Ile	Asp	Val	Asp	Ala	Asp	
1185					1190					1195					1200	
Gln	Glu	Pro	Ser	Gln	Ser	Pro	Pro	Arg	Ser	Glu	Ala	Val	Leu	Gln	Gln	
1205					1210					1215						
Glu	Asp	Glu	Gly	Ala	Leu	Pro	Glu	Asn	Arg	Gly	Ser	Leu	Gly	Arg	Arg	
1220					1225					1230						
Gly	Ala	Pro	Trp	Leu	Phe	Cys	Asp	Arg	Glu	Ser	Ser	Pro	Ser	Glu	Ala	
1235					1240					1245						
Ser	Thr	Thr	Asp	Thr	Ser	Trp	Leu	Val	Pro	Ala	Thr	Pro	Leu	Ala	Ser	
1250					1255					1260						
Arg	Ser	Arg	Asp	Cys	Ser	Ser	Gln	Thr	Gln	Ile	Ser	Ser	Leu	Arg	Ser	
1265					1270					1275					1280	
Gly	Leu	Ala	Val	Gln	Ala	Val	Thr	Gln	His	Thr	Pro	Arg	Ala	Ser	Val	
1285					1290					1295						
Gly	Asn	Arg	Glu	Gly	Asn	Glu	Val	Ala	Gln	Lys	Phe	Ser	Val	Ile	Arg	
1300					1305					1310						
Pro	Gln	Thr	Pro	Pro	Pro	Gln	Thr	Pro	Ser	Ser	Cys	Leu	Thr	Pro	Val	
1315					1320					1325						
Ser	Pro	Gly	Thr	Ser	Asp	Gly	Arg	Arg	Gln	Gly	His	Arg	Ser	Pro	Ser	
1330					1335					1340						
Arg	Pro	His	Pro	Gly	Gly	His	Pro	His	Ser	Ser	Pro	Leu	Ala	Pro	His	
1345					1350					1355					1360	
Pro	Ile	Ser	Gly	Asp	Arg	Ala	His	Phe	Ser	Arg	Arg	Phe	Leu	Lys	His	
1365					1370					1375						
Ser	Pro	Pro	Gly	Pro	Ser	Phe	Leu	Asn	Gln	Thr	Pro	Ala	Gly	Glu	Val	
1380					1385					1390						
Val	Glu	Val	Gly	Asp	Ser	Asp	Asp	Glu	Gln	Glu	Val	Ala	Ser	His	Gln	
1395					1400					1405						
Ala	Asn	Arg	Ser	Pro	Pro	Leu	Asp	Ser	Asp	Pro	Pro	Ile	Pro	Ile	Asp	
1410					1415					1420						
Asp	Cys	Cys	Trp	His	Met	Glu	Pro	Leu	Ser	Pro	Ile	Pro	Ile	Asp	His	

1425					1430					1435				1440
Trp	Asn	Leu	Glu	Arg	Thr	Gly	Pro	Leu	Ser	Thr	Ser	Ser	Pro	Ser
					1445					1450				1455
Arg	Met	Asn	Glu	Ala	Ala	Asp	Ser	Arg	Asp	Cys	Arg	Ser	Pro	Gly
			1460						1465					1470
Leu	Asp	Thr	Thr	Pro	Ile	Arg	Gly	Ser	Cys	Thr	Thr	Gln	Arg	Lys
		1475					1480					1485		
Gln	Glu	Lys	Ser	Ser	Gly	Ala	Gly	Ser	Leu	Gly	Asn	Ser	Arg	Pro
	1490				1495						1500			
Phe	Leu	Asn	Ser	Ala	Leu	Trp	Asp	Val	Trp	Asp	Gly	Glu	Glu	Gln
1505					1510					1515				1520
Pro	Pro	Glu	Thr	Pro	Pro	Pro	Ala	Gln	Met	Pro	Ser	Ala	Gly	Gly
				1525					1530					1535
Gln	Lys	Pro	Glu	Gly	Leu	Glu	Thr	Pro	Lys	Gly	Ala	Asn	Arg	Lys
			1540					1545					1550	
Asn	Leu	Pro	Pro	Lys	Val	Pro	Ile	Thr	Pro	Met	Pro	Gln	Tyr	Ser
	1555						1560					1565		
Met	Glu	Thr	Pro	Val	Leu	Lys	Lys	Glu	Leu	Asp	Arg	Phe	Gly	Val
	1570					1575				1580				
Pro	Leu	Pro	Lys	Arg	Gln	Met	Val	Leu	Lys	Leu	Lys	Glu	Ile	Phe
1585				1590					1595					1600
Tyr	Thr	His	Gln	Thr	Leu	Asp	Ser	Asp	Ser	Glu	Asp	Glu	Ser	Gln
			1605					1610						1615
Ser	Gln	Pro	Leu	Leu	Gln	Ala	Pro	His	Cys	Gln	Thr	Leu	Ala	Ser
		1620				1625				1630				
Thr	Tyr	Lys	Pro	Ser	Arg	Ala	Gly	Val	His	Ala	Gln	Gln	Glu	Ala
	1635					1640				1645				
Thr	Gly	Pro	Gly	Ala	His	Arg	Pro	Lys	Gly	Pro	Ala	Lys	Thr	Lys
	1650				1655				1660					
Pro	Arg	His	Gln	Arg	Lys	His	His	Glu	Ser	Ile	Thr	Pro	Pro	Ser
1665				1670					1675					1680
Ser	Pro	Thr	Lys	Glu	Ala	Pro	Pro	Gly	Leu	Asn	Asp	Asp	Ala	Gln
			1685			1690								1695
Pro	Ala	Ser	Gln	Glu	Ser	Val	Ala	Thr	Ser	Val	Asp	Gly	Ser	Asp
		1700				1705						1710		
Ser	Leu	Ser	Ser	Gln	Ser	Ser	Ser	Ser	Cys	Glu	Phe	Gly	Ala	Ala
	1715					1720					1725			
Glu	Ser	Ala	Gly	Glu	Glu	Glu	Gly	Glu	Gly	Glu	Val	Ser	Ala	Ser
	1730			1735				1740						
Ala	Ala	Val	Gln	Ala	Ala	Asp	Thr	Asp	Glu	Ala	Leu	Arg	Cys	Tyr
1745				1750				1755						1760
Arg	Ser	Lys	Pro	Ala	Leu	Tyr	Gln	Lys	Val	Leu	Leu	Tyr	Gln	Pro
			1765			1770							1775	
Glu	Leu	Arg	Glu	Leu	Gln	Ala	Glu	Leu	Arg	Gln	Asn	Gly	Leu	Arg
	1780					1785					1790			
Ser	Ser	Arg	Arg	Leu	Leu	Asp	Phe	Leu	Asp	Thr	His	Cys	Ile	Thr
	1795					1800					1805			
Thr	Thr	Ala	Ala	Thr	Arg	Arg	Glu	Lys	Leu	Gln	Gly	Arg	Arg	Arg
	1810				1815					1820				
Pro	Arg	Gly	Lys	Lys	Lys	Val	Glu	Arg	Asn					
1825				1830										

<210> 203

<211> 608

<212> PRT
 <213> Homo sapiens

<220>
 <223> >TDP1|ENSG00000042088|ENST00000335725|1827

<400> 203

Met	Ser	Gln	Glu	Gly	Asp	Tyr	Gly	Arg	Trp	Thr	Ile	Ser	Ser	Ser	Asp	1	5	10	15
Glu	Ser	Glu	Glu	Glu	Lys	Pro	Lys	Pro	Asp	Lys	Pro	Ser	Thr	Ser	Ser	20	25	30	
Leu	Leu	Cys	Ala	Arg	Gln	Gly	Ala	Ala	Asn	Glu	Pro	Arg	Tyr	Thr	Cys	35	40	45	
Ser	Glu	Ala	Gln	Lys	Ala	Ala	His	Lys	Arg	Lys	Ile	Ser	Pro	Val	Lys	50	55	60	
Phe	Ser	Asn	Thr	Asp	Ser	Val	Leu	Pro	Pro	Lys	Arg	Gln	Lys	Ser	Gly	65	70	75	80
Ser	Gln	Glu	Asp	Leu	Gly	Trp	Cys	Leu	Ser	Ser	Ser	Asp	Asp	Glu	Leu	85	90	95	
Gln	Pro	Glu	Met	Pro	Gln	Lys	Gln	Ala	Glu	Lys	Val	Val	Ile	Lys	Lys	100	105	110	
Glu	Lys	Asp	Ile	Ser	Ala	Pro	Asn	Asp	Gly	Thr	Ala	Gln	Arg	Thr	Glu	115	120	125	
Asn	His	Gly	Ala	Pro	Ala	Cys	His	Arg	Leu	Lys	Glu	Glu	Glu	Asp	Glu	130	135	140	
Tyr	Glu	Thr	Ser	Gly	Glu	Gly	Gln	Asp	Ile	Trp	Asp	Met	Leu	Asp	Lys	145	150	155	160
Gly	Asn	Pro	Phe	Gln	Phe	Tyr	Leu	Thr	Arg	Val	Ser	Gly	Val	Lys	Pro	165	170	175	
Lys	Tyr	Asn	Ser	Gly	Ala	Leu	His	Ile	Lys	Asp	Ile	Leu	Ser	Pro	Leu	180	185	190	
Phe	Gly	Thr	Leu	Val	Ser	Ser	Ala	Gln	Phe	Asn	Tyr	Cys	Phe	Asp	Val	195	200	205	
Asp	Trp	Leu	Val	Lys	Gln	Tyr	Pro	Pro	Glu	Phe	Arg	Lys	Lys	Pro	Ile	210	215	220	
Leu	Leu	Val	His	Gly	Asp	Lys	Arg	Glu	Ala	Lys	Ala	His	Leu	His	Ala	225	230	235	240
Gln	Ala	Lys	Pro	Tyr	Glu	Asn	Ile	Ser	Leu	Cys	Gln	Ala	Lys	Leu	Asp	245	250	255	
Ile	Ala	Phe	Gly	Thr	His	His	Thr	Lys	Met	Met	Leu	Leu	Leu	Tyr	Glu	260	265	270	
Glu	Gly	Leu	Arg	Val	Val	Ile	His	Thr	Ser	Asn	Leu	Ile	His	Ala	Asp	275	280	285	
Trp	His	Gln	Lys	Thr	Gln	Gly	Ile	Trp	Leu	Ser	Pro	Leu	Tyr	Pro	Arg	290	295	300	
Ile	Ala	Asp	Gly	Thr	His	Lys	Ser	Gly	Glu	Ser	Pro	Thr	His	Phe	Lys	305	310	315	320
Ala	Asp	Leu	Ile	Ser	Tyr	Leu	Met	Ala	Tyr	Asn	Ala	Pro	Ser	Leu	Lys	325	330	335	
Glu	Trp	Ile	Asp	Val	Ile	His	Lys	His	Asp	Leu	Ser	Glu	Thr	Asn	Val	340	345	350	
Tyr	Leu	Ile	Gly	Ser	Thr	Pro	Gly	Arg	Phe	Gln	Gly	Ser	Gln	Lys	Asp	355	360	365	

Asn	Trp	Gly	His	Phe	Arg	Leu	Lys	Lys	Leu	Leu	Lys	Asp	His	Ala	Ser
370						375					380				
Ser	Met	Pro	Asn	Ala	Glu	Ser	Trp	Pro	Val	Val	Gly	Gln	Phe	Ser	Ser
385					390					395					400
Val	Gly	Ser	Leu	Gly	Ala	Asp	Glu	Ser	Lys	Trp	Leu	Cys	Ser	Glu	Phe
				405					410					415	
Lys	Glu	Ser	Met	Leu	Thr	Leu	Gly	Lys	Glu	Ser	Lys	Thr	Pro	Gly	Lys
			420					425					430		
Ser	Ser	Val	Pro	Leu	Tyr	Leu	Ile	Tyr	Pro	Ser	Val	Glu	Asn	Val	Arg
		435				440					445				
Thr	Ser	Leu	Glu	Gly	Tyr	Pro	Ala	Gly	Gly	Ser	Leu	Pro	Tyr	Ser	Ile
450					455						460				
Gln	Thr	Ala	Glu	Lys	Gln	Asn	Trp	Leu	His	Ser	Tyr	Phe	His	Lys	Trp
465					470					475					480
Ser	Ala	Glu	Thr	Ser	Gly	Arg	Ser	Asn	Ala	Met	Pro	His	Ile	Lys	Thr
				485				490						495	
Tyr	Met	Arg	Pro	Ser	Pro	Asp	Phe	Ser	Lys	Ile	Ala	Trp	Phe	Leu	Val
		500						505					510		
Thr	Ser	Ala	Asn	Leu	Ser	Lys	Ala	Ala	Trp	Gly	Ala	Leu	Glu	Lys	Asn
		515				520						525			
Gly	Thr	Gln	Leu	Met	Ile	Arg	Ser	Tyr	Glu	Leu	Gly	Val	Leu	Phe	Leu
530					535						540				
Pro	Ser	Ala	Phe	Gly	Leu	Asp	Ser	Phe	Lys	Val	Lys	Gln	Lys	Phe	Phe
545				550				555							560
Ala	Gly	Ser	Gln	Glu	Pro	Met	Ala	Thr	Phe	Pro	Val	Pro	Tyr	Asp	Leu
			565					570						575	
Pro	Pro	Glu	Leu	Tyr	Gly	Ser	Lys	Asp	Arg	Pro	Trp	Ile	Trp	Asn	Ile
		580				585							590		
Pro	Tyr	Val	Lys	Ala	Pro	Asp	Thr	His	Gly	Asn	Met	Trp	Val	Pro	Ser
		595				600						605			

<210> 204

<211> 362

<212> PRT

<213> Homo sapiens

<220>

<223> >TDP2|ENSG00000111802|ENST00000378198|1089

<400> 204

Met	Glu	Leu	Gly	Ser	Cys	Leu	Glu	Gly	Gly	Arg	Glu	Ala	Ala	Glu	Glu
1				5				10						15	
Glu	Gly	Glu	Pro	Glu	Val	Lys	Lys	Arg	Arg	Leu	Leu	Cys	Val	Glu	Phe
			20					25					30		
Ala	Ser	Val	Ala	Ser	Cys	Asp	Ala	Ala	Val	Ala	Gln	Cys	Phe	Leu	Ala
		35				40						45			
Glu	Asn	Asp	Trp	Glu	Met	Glu	Arg	Ala	Leu	Asn	Ser	Tyr	Phe	Glu	Pro
	50				55					60					
Pro	Val	Glu	Glu	Ser	Ala	Leu	Glu	Arg	Arg	Pro	Glu	Thr	Ile	Ser	Glu
65				70				75						80	
Pro	Lys	Thr	Tyr	Val	Asp	Leu	Thr	Asn	Glu	Glu	Thr	Thr	Asp	Ser	Thr

				85					90					95			
Thr	Ser	Lys	Ile	Ser	Pro	Ser	Glu	Asp	Thr	Gln	Gln	Glu	Asn	Gly	Ser		
			100					105					110				
Met	Phe	Ser	Leu	Ile	Thr	Trp	Asn	Ile	Asp	Gly	Leu	Asp	Leu	Asn	Asn		
		115					120					125					
Leu	Ser	Glu	Arg	Ala	Arg	Gly	Val	Cys	Ser	Tyr	Leu	Ala	Leu	Tyr	Ser		
		130				135					140						
Pro	Asp	Val	Ile	Phe	Leu	Gln	Glu	Val	Ile	Pro	Pro	Tyr	Tyr	Ser	Tyr		
145					150					155					160		
Leu	Lys	Lys	Arg	Ser	Ser	Asn	Tyr	Glu	Ile	Ile	Thr	Gly	His	Glu	Glu		
			165					170						175			
Gly	Tyr	Phe	Thr	Ala	Ile	Met	Leu	Lys	Lys	Ser	Arg	Val	Lys	Leu	Lys		
			180					185					190				
Ser	Gln	Glu	Ile	Ile	Pro	Phe	Pro	Ser	Thr	Lys	Met	Met	Arg	Asn	Leu		
		195					200					205					
Leu	Cys	Val	His	Val	Asn	Val	Ser	Gly	Asn	Glu	Leu	Cys	Leu	Met	Thr		
	210					215					220						
Ser	His	Leu	Glu	Ser	Thr	Arg	Gly	His	Ala	Ala	Glu	Arg	Met	Asn	Gln		
225					230					235					240		
Leu	Lys	Met	Val	Leu	Lys	Lys	Met	Gln	Glu	Ala	Pro	Glu	Ser	Ala	Thr		
			245					250						255			
Val	Ile	Phe	Ala	Gly	Asp	Thr	Asn	Leu	Arg	Asp	Arg	Glu	Val	Thr	Arg		
		260					265						270				
Cys	Gly	Gly	Leu	Pro	Asn	Asn	Ile	Val	Asp	Val	Trp	Glu	Phe	Leu	Gly		
	275					280					285						
Lys	Pro	Lys	His	Cys	Gln	Tyr	Thr	Trp	Asp	Thr	Gln	Met	Asn	Ser	Asn		
	290				295						300						
Leu	Gly	Ile	Thr	Ala	Ala	Cys	Lys	Leu	Arg	Phe	Asp	Arg	Ile	Phe	Phe		
					310					315					320		
Arg	Ala	Ala	Ala	Glu	Glu	Gly	His	Ile	Ile	Pro	Arg	Ser	Leu	Asp	Leu		
				325				330						335			
Leu	Gly	Leu	Glu	Lys	Leu	Asp	Cys	Gly	Arg	Phe	Pro	Ser	Asp	His	Trp		
		340					345						350				
Gly	Leu	Leu	Cys	Asn	Leu	Asp	Ile	Ile	Leu								
		355					360										

<210> 205

<211> 529

<212> PRT

<213> Homo sapiens

<220>

<223> >TMPRSS2|ENSG00000184012|ENST00000398585|1590

<400> 205

Met	Pro	Pro	Ala	Pro	Pro	Gly	Gly	Glu	Ser	Gly	Cys	Glu	Glu	Arg	Gly		
1				5				10						15			
Ala	Ala	Gly	His	Ile	Glu	His	Ser	Arg	Tyr	Leu	Ser	Leu	Leu	Asp	Ala		
			20					25					30				
Val	Asp	Asn	Ser	Lys	Met	Ala	Leu	Asn	Ser	Gly	Ser	Pro	Pro	Ala	Ile		
		35					40					45					
Gly	Pro	Tyr	Tyr	Glu	Asn	His	Gly	Tyr	Gln	Pro	Glu	Asn	Pro	Tyr	Pro		
	50					55					60						

Ala	Gln	Pro	Thr	Val	Val	Pro	Thr	Val	Tyr	Glu	Val	His	Pro	Ala	Gln
65					70					75					80
Tyr	Tyr	Pro	Ser	Pro	Val	Pro	Gln	Tyr	Ala	Pro	Arg	Val	Leu	Thr	Gln
				85					90					95	
Ala	Ser	Asn	Pro	Val	Val	Cys	Thr	Gln	Pro	Lys	Ser	Pro	Ser	Gly	Thr
			100					105						110	
Val	Cys	Thr	Ser	Lys	Thr	Lys	Lys	Ala	Leu	Cys	Ile	Thr	Leu	Thr	Leu
		115					120					125			
Gly	Thr	Phe	Leu	Val	Gly	Ala	Ala	Leu	Ala	Ala	Gly	Leu	Leu	Trp	Lys
		130				135					140				
Phe	Met	Gly	Ser	Lys	Cys	Ser	Asn	Ser	Gly	Ile	Glu	Cys	Asp	Ser	Ser
145					150					155					160
Gly	Thr	Cys	Ile	Asn	Pro	Ser	Asn	Trp	Cys	Asp	Gly	Val	Ser	His	Cys
				165					170					175	
Pro	Gly	Gly	Glu	Asp	Glu	Asn	Arg	Cys	Val	Arg	Leu	Tyr	Gly	Pro	Asn
			180					185					190		
Phe	Ile	Leu	Gln	Val	Tyr	Ser	Ser	Gln	Arg	Lys	Ser	Trp	His	Pro	Val
		195					200					205			
Cys	Gln	Asp	Asp	Trp	Asn	Glu	Asn	Tyr	Gly	Arg	Ala	Ala	Cys	Arg	Asp
		210				215					220				
Met	Gly	Tyr	Lys	Asn	Asn	Phe	Tyr	Ser	Ser	Gln	Gly	Ile	Val	Asp	Asp
225					230					235					240
Ser	Gly	Ser	Thr	Ser	Phe	Met	Lys	Leu	Asn	Thr	Ser	Ala	Gly	Asn	Val
				245					250					255	
Asp	Ile	Tyr	Lys	Lys	Leu	Tyr	His	Ser	Asp	Ala	Cys	Ser	Ser	Lys	Ala
			260					265					270		
Val	Val	Ser	Leu	Arg	Cys	Ile	Ala	Cys	Gly	Val	Asn	Leu	Asn	Ser	Ser
		275					280					285			
Arg	Gln	Ser	Arg	Ile	Val	Gly	Gly	Glu	Ser	Ala	Leu	Pro	Gly	Ala	Trp
		290				295					300				
Pro	Trp	Gln	Val	Ser	Leu	His	Val	Gln	Asn	Val	His	Val	Cys	Gly	Gly
305					310					315					320
Ser	Ile	Ile	Thr	Pro	Glu	Trp	Ile	Val	Thr	Ala	Ala	His	Cys	Val	Glu
				325					330					335	
Lys	Pro	Leu	Asn	Asn	Pro	Trp	His	Trp	Thr	Ala	Phe	Ala	Gly	Ile	Leu
			340					345					350		
Arg	Gln	Ser	Phe	Met	Phe	Tyr	Gly	Ala	Gly	Tyr	Gln	Val	Glu	Lys	Val
		355					360					365			
Ile	Ser	His	Pro	Asn	Tyr	Asp	Ser	Lys	Thr	Lys	Asn	Asn	Asp	Ile	Ala
		370				375					380				
Leu	Met	Lys	Leu	Gln	Lys	Pro	Leu	Thr	Phe	Asn	Asp	Leu	Val	Lys	Pro
385					390					395					400
Val	Cys	Leu	Pro	Asn	Pro	Gly	Met	Met	Leu	Gln	Pro	Glu	Gln	Leu	Cys
				405					410					415	
Trp	Ile	Ser	Gly	Trp	Gly	Ala	Thr	Glu	Glu	Lys	Gly	Lys	Thr	Ser	Glu
			420					425					430		
Val	Leu	Asn	Ala	Ala	Lys	Val	Leu	Leu	Ile	Glu	Thr	Gln	Arg	Cys	Asn
		435					440					445			
Ser	Arg	Tyr	Val	Tyr	Asp	Asn	Leu	Ile	Thr	Pro	Ala	Met	Ile	Cys	Ala
						455					460				
Gly	Phe	Leu	Gln	Gly	Asn	Val	Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly
465					470					475					480
Pro	Leu	Val	Thr	Ser	Lys	Asn	Asn	Ile	Trp	Trp	Leu	Ile	Gly	Asp	Thr
				485					490					495	

Ser Trp Gly Ser Gly Cys Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly
500 505 510
Asn Val Met Val Phe Thr Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp
515 520 525
Gly

<210> 206
<211> 1531
<212> PRT
<213> Homo sapiens

<220>
<223> >TOP2A|ENSG00000131747|ENST00000423485|4596

<400> 206
Met Glu Val Ser Pro Leu Gln Pro Val Asn Glu Asn Met Gln Val Asn
1 5 10 15
Lys Ile Lys Lys Asn Glu Asp Ala Lys Lys Arg Leu Ser Val Glu Arg
20 25 30
Ile Tyr Gln Lys Lys Thr Gln Leu Glu His Ile Leu Leu Arg Pro Asp
35 40 45
Thr Tyr Ile Gly Ser Val Glu Leu Val Thr Gln Gln Met Trp Val Tyr
50 55 60
Asp Glu Asp Val Gly Ile Asn Tyr Arg Glu Val Thr Phe Val Pro Gly
65 70 75 80
Leu Tyr Lys Ile Phe Asp Glu Ile Leu Val Asn Ala Ala Asp Asn Lys
85 90 95
Gln Arg Asp Pro Lys Met Ser Cys Ile Arg Val Thr Ile Asp Pro Glu
100 105 110
Asn Asn Leu Ile Ser Ile Trp Asn Asn Gly Lys Gly Ile Pro Val Val
115 120 125
Glu His Lys Val Glu Lys Met Tyr Val Pro Ala Leu Ile Phe Gly Gln
130 135 140
Leu Leu Thr Ser Ser Asn Tyr Asp Asp Asp Glu Lys Lys Val Thr Gly
145 150 155 160
Gly Arg Asn Gly Tyr Gly Ala Lys Leu Cys Asn Ile Phe Ser Thr Lys
165 170 175
Phe Thr Val Glu Thr Ala Ser Arg Glu Tyr Lys Lys Met Phe Lys Gln
180 185 190
Thr Trp Met Asp Asn Met Gly Arg Ala Gly Glu Met Glu Leu Lys Pro
195 200 205
Phe Asn Gly Glu Asp Tyr Thr Cys Ile Thr Phe Gln Pro Asp Leu Ser
210 215 220
Lys Phe Lys Met Gln Ser Leu Asp Lys Asp Ile Val Ala Leu Met Val
225 230 235 240
Arg Arg Ala Tyr Asp Ile Ala Gly Ser Thr Lys Asp Val Lys Val Phe
245 250 255
Leu Asn Gly Asn Lys Leu Pro Val Lys Gly Phe Arg Ser Tyr Val Asp
260 265 270
Met Tyr Leu Lys Asp Lys Leu Asp Glu Thr Gly Asn Ser Leu Lys Val
275 280 285
Ile His Glu Gln Val Asn His Arg Trp Glu Val Cys Leu Thr Met Ser

290		295		300
Glu Lys Gly Phe Gln	Gln Ile Ser Phe Val	Asn Ser Ile Ala Thr Ser		
305	310	315	320	
Lys Gly Gly Arg His	Val Asp Tyr Val Ala	Asp Gln Ile Val Thr Lys		
	325	330	335	
Leu Val Asp Val Val	Lys Lys Lys Asn Lys	Gly Gly Val Ala Val Lys		
	340	345	350	
Ala His Gln Val Lys	Asn His Met Trp Ile Phe	Val Asn Ala Leu Ile		
	355	360	365	
Glu Asn Pro Thr Phe	Asp Ser Gln Thr Lys	Glu Asn Met Thr Leu Gln		
	370	375	380	
Pro Lys Ser Phe Gly	Ser Thr Cys Gln Leu Ser	Glu Lys Phe Ile Lys		
385	390	395	400	
Ala Ala Ile Gly Cys	Gly Ile Val Glu Ser	Ile Leu Asn Trp Val Lys		
	405	410	415	
Phe Lys Ala Gln Val	Gln Leu Asn Lys Lys	Cys Ser Ala Val Lys His		
	420	425	430	
Asn Arg Ile Lys Gly	Ile Pro Lys Leu Asp	Asp Ala Asn Asp Ala Gly		
	435	440	445	
Gly Arg Asn Ser Thr	Glu Cys Thr Leu Ile	Leu Thr Glu Gly Asp Ser		
	450	455	460	
Ala Lys Thr Leu Ala	Val Ser Gly Leu Gly	Val Val Gly Arg Asp Lys		
465	470	475	480	
Tyr Gly Val Phe Pro	Leu Arg Gly Lys Ile	Leu Asn Val Arg Glu Ala		
	485	490	495	
Ser His Lys Gln Ile	Met Glu Asn Ala Glu	Ile Asn Asn Ile Ile Lys		
	500	505	510	
Ile Val Gly Leu Gln	Tyr Lys Lys Asn Tyr	Glu Asp Glu Asp Ser Leu		
	515	520	525	
Lys Thr Leu Arg Tyr	Gly Lys Ile Met Ile	Met Thr Asp Gln Asp Gln		
	530	535	540	
Asp Gly Ser His Ile	Lys Gly Leu Leu Ile	Asn Phe Ile His His Asn		
545	550	555	560	
Trp Pro Ser Leu Leu	Arg His Arg Phe Leu	Glu Glu Phe Ile Thr Pro		
	565	570	575	
Ile Val Lys Val Ser	Lys Asn Lys Gln Glu	Met Ala Phe Tyr Ser Leu		
	580	585	590	
Pro Glu Phe Glu Glu	Trp Lys Ser Ser Thr	Pro Asn His Lys Lys Trp		
	595	600	605	
Lys Val Lys Tyr Tyr	Lys Gly Leu Gly Thr	Ser Thr Ser Lys Glu Ala		
	610	615	620	
Lys Glu Tyr Phe Ala	Asp Met Lys Arg His	Arg Ile Gln Phe Lys Tyr		
625	630	635	640	
Ser Gly Pro Glu Asp	Asp Ala Ala Ile Ser	Leu Ala Phe Ser Lys Lys		
	645	650	655	
Gln Ile Asp Asp Arg	Lys Glu Trp Leu Thr	Asn Phe Met Glu Asp Arg		
	660	665	670	
Arg Gln Arg Lys Leu	Leu Gly Leu Pro Glu	Asp Tyr Leu Tyr Gly Gln		
	675	680	685	
Thr Thr Thr Tyr Leu	Thr Tyr Asn Asp Phe	Ile Asn Lys Glu Leu Ile		
	690	695	700	
Leu Phe Ser Asn Ser	Asp Asn Glu Arg Ser	Ile Pro Ser Met Val Asp		
705	710	715	720	
Gly Leu Lys Pro Gly	Gln Arg Lys Val Leu	Phe Thr Cys Phe Lys Arg		

				725				730				735			
Asn	Asp	Lys	Arg	Glu	Val	Lys	Val	Ala	Gln	Leu	Ala	Gly	Ser	Val	Ala
			740					745					750		
Glu	Met	Ser	Ser	Tyr	His	His	Gly	Glu	Met	Ser	Leu	Met	Met	Thr	Ile
		755					760					765			
Ile	Asn	Leu	Ala	Gln	Asn	Phe	Val	Gly	Ser	Asn	Asn	Leu	Asn	Leu	Leu
	770					775					780				
Gln	Pro	Ile	Gly	Gln	Phe	Gly	Thr	Arg	Leu	His	Gly	Gly	Lys	Asp	Ser
785					790					795					800
Ala	Ser	Pro	Arg	Tyr	Ile	Phe	Thr	Met	Leu	Ser	Ser	Leu	Ala	Arg	Leu
			805					810						815	
Leu	Phe	Pro	Pro	Lys	Asp	Asp	His	Thr	Leu	Lys	Phe	Leu	Tyr	Asp	Asp
			820					825					830		
Asn	Gln	Arg	Val	Glu	Pro	Glu	Trp	Tyr	Ile	Pro	Ile	Ile	Pro	Met	Val
		835					840					845			
Leu	Ile	Asn	Gly	Ala	Glu	Gly	Ile	Gly	Thr	Gly	Trp	Ser	Cys	Lys	Ile
	850					855					860				
Pro	Asn	Phe	Asp	Val	Arg	Glu	Ile	Val	Asn	Asn	Ile	Arg	Arg	Leu	Met
865					870					875					880
Asp	Gly	Glu	Glu	Pro	Leu	Pro	Met	Leu	Pro	Ser	Tyr	Lys	Asn	Phe	Lys
				885					890					895	
Gly	Thr	Ile	Glu	Glu	Leu	Ala	Pro	Asn	Gln	Tyr	Val	Ile	Ser	Gly	Glu
			900					905						910	
Val	Ala	Ile	Leu	Asn	Ser	Thr	Thr	Ile	Glu	Ile	Ser	Glu	Leu	Pro	Val
		915					920					925			
Arg	Thr	Trp	Thr	Gln	Thr	Tyr	Lys	Glu	Gln	Val	Leu	Glu	Pro	Met	Leu
						935					940				
Asn	Gly	Thr	Glu	Lys	Thr	Pro	Pro	Leu	Ile	Thr	Asp	Tyr	Arg	Glu	Tyr
945					950					955					960
His	Thr	Asp	Thr	Thr	Val	Lys	Phe	Val	Val	Lys	Met	Thr	Glu	Glu	Lys
				965					970					975	
Leu	Ala	Glu	Ala	Glu	Arg	Val	Gly	Leu	His	Lys	Val	Phe	Lys	Leu	Gln
			980					985					990		
Thr	Ser	Leu	Thr	Cys	Asn	Ser	Met	Val	Leu	Phe	Asp	His	Val	Gly	Cys
		995					1000					1005			
Leu	Lys	Lys	Tyr	Asp	Thr	Val	Leu	Asp	Ile	Leu	Arg	Asp	Phe	Phe	Glu
	1010					1015					1020				
Leu	Arg	Leu	Lys	Tyr	Tyr	Gly	Leu	Arg	Lys	Glu	Trp	Leu	Leu	Gly	Met
1025					1030					1035					1040
Leu	Gly	Ala	Glu	Ser	Ala	Lys	Leu	Asn	Asn	Gln	Ala	Arg	Phe	Ile	Leu
				1045					1050					1055	
Glu	Lys	Ile	Asp	Gly	Lys	Ile	Ile	Ile	Glu	Asn	Lys	Pro	Lys	Lys	Glu
			1060					1065						1070	
Leu	Ile	Lys	Val	Leu	Ile	Gln	Arg	Gly	Tyr	Asp	Ser	Asp	Pro	Val	Lys
	1075						1080					1085			
Ala	Trp	Lys	Glu	Ala	Gln	Gln	Lys	Val	Pro	Asp	Glu	Glu	Glu	Asn	Glu
	1090					1095					1100				
Glu	Ser	Asp	Asn	Glu	Lys	Glu	Thr	Glu	Lys	Ser	Asp	Ser	Val	Thr	Asp
1105					1110					1115					1120
Ser	Gly	Pro	Thr	Phe	Asn	Tyr	Leu	Leu	Asp	Met	Pro	Leu	Trp	Tyr	Leu
				1125					1130					1135	
Thr	Lys	Glu	Lys	Lys	Asp	Glu	Leu	Cys	Arg	Leu	Arg	Asn	Glu	Lys	Glu
			1140					1145					1150		
Gln	Glu	Leu	Asp	Thr	Leu	Lys	Arg	Lys	Ser	Pro	Ser	Asp	Leu	Trp	Lys

1155					1160					1165					
Glu	Asp	Leu	Ala	Thr	Phe	Ile	Glu	Glu	Leu	Glu	Ala	Val	Glu	Ala	Lys
1170					1175					1180					
Glu	Lys	Gln	Asp	Glu	Gln	Val	Gly	Leu	Pro	Gly	Lys	Gly	Gly	Lys	Ala
1185					1190					1195					1200
Lys	Gly	Lys	Lys	Thr	Gln	Met	Ala	Glu	Val	Leu	Pro	Ser	Pro	Arg	Gly
1205					1210					1215					
Gln	Arg	Val	Ile	Pro	Arg	Ile	Thr	Ile	Glu	Met	Lys	Ala	Glu	Ala	Glu
1220					1225					1230					
Lys	Lys	Asn	Lys	Lys	Lys	Ile	Lys	Asn	Glu	Asn	Thr	Glu	Gly	Ser	Pro
1235					1240					1245					
Gln	Glu	Asp	Gly	Val	Glu	Leu	Glu	Gly	Leu	Lys	Gln	Arg	Leu	Glu	Lys
1250					1255					1260					
Lys	Gln	Lys	Arg	Glu	Pro	Gly	Thr	Lys	Thr	Lys	Lys	Gln	Thr	Thr	Leu
1265					1270					1275					1280
Ala	Phe	Lys	Pro	Ile	Lys	Lys	Gly	Lys	Lys	Arg	Asn	Pro	Trp	Ser	Asp
1285					1290					1295					
Ser	Glu	Ser	Asp	Arg	Ser	Ser	Asp	Glu	Ser	Asn	Phe	Asp	Val	Pro	Pro
1300					1305					1310					
Arg	Glu	Thr	Glu	Pro	Arg	Arg	Ala	Ala	Thr	Lys	Thr	Lys	Phe	Thr	Met
1315					1320					1325					
Asp	Leu	Asp	Ser	Asp	Glu	Asp	Phe	Ser	Asp	Phe	Asp	Glu	Lys	Thr	Asp
1330					1335					1340					
Asp	Glu	Asp	Phe	Val	Pro	Ser	Asp	Ala	Ser	Pro	Pro	Lys	Thr	Lys	Thr
1345					1350					1355					1360
Ser	Pro	Lys	Leu	Ser	Asn	Lys	Glu	Leu	Lys	Pro	Gln	Lys	Ser	Val	Val
1365					1370					1375					
Ser	Asp	Leu	Glu	Ala	Asp	Asp	Val	Lys	Gly	Ser	Val	Pro	Leu	Ser	Ser
1380					1385					1390					
Ser	Pro	Pro	Ala	Thr	His	Phe	Pro	Asp	Glu	Thr	Glu	Ile	Thr	Asn	Pro
1395					1400					1405					
Val	Pro	Lys	Lys	Asn	Val	Thr	Val	Lys	Lys	Thr	Ala	Ala	Lys	Ser	Gln
1410					1415					1420					
Ser	Ser	Thr	Ser	Thr	Thr	Gly	Ala	Lys	Lys	Arg	Ala	Ala	Pro	Lys	Gly
1425					1430					1435					1440
Thr	Lys	Arg	Asp	Pro	Ala	Leu	Asn	Ser	Gly	Val	Ser	Gln	Lys	Pro	Asp
1445					1450					1455					
Pro	Ala	Lys	Thr	Lys	Asn	Arg	Arg	Lys	Arg	Lys	Pro	Ser	Thr	Ser	Asp
1460					1465					1470					
Asp	Ser	Asp	Ser	Asn	Phe	Glu	Lys	Ile	Val	Ser	Lys	Ala	Val	Thr	Ser
1475					1480					1485					
Lys	Lys	Ser	Lys	Gly	Glu	Ser	Asp	Asp	Phe	His	Met	Asp	Phe	Asp	Ser
1490					1495					1500					
Ala	Val	Ala	Pro	Arg	Ala	Lys	Ser	Val	Arg	Ala	Lys	Lys	Pro	Ile	Lys
1505					1510					1515					1520
Tyr	Leu	Glu	Glu	Ser	Asp	Glu	Asp	Asp	Leu	Phe					
1525					1530										

<210> 207

<211> 1621

<212> PRT

<213> Homo sapiens

<220>

<223> >TOP2B|ENSG00000077097|ENST00000435706|4866

<400> 207

Met	Ala	Lys	Ser	Gly	Gly	Cys	Gly	Ala	Gly	Ala	Gly	Val	Gly	Gly	Gly
1				5					10					15	
Asn	Gly	Ala	Leu	Thr	Trp	Val	Asn	Asn	Ala	Ala	Lys	Lys	Glu	Glu	Ser
			20					25					30		
Glu	Thr	Ala	Asn	Lys	Asn	Asp	Ser	Ser	Lys	Lys	Leu	Ser	Val	Glu	Arg
		35				40						45			
Val	Tyr	Gln	Lys	Lys	Thr	Gln	Leu	Glu	His	Ile	Leu	Leu	Arg	Pro	Asp
	50					55					60				
Thr	Tyr	Ile	Gly	Ser	Val	Glu	Pro	Leu	Thr	Gln	Phe	Met	Trp	Val	Tyr
65					70					75					80
Asp	Glu	Asp	Val	Gly	Met	Asn	Cys	Arg	Glu	Val	Thr	Phe	Val	Pro	Gly
				85					90					95	
Leu	Tyr	Lys	Ile	Phe	Asp	Glu	Ile	Leu	Val	Asn	Ala	Ala	Asp	Asn	Lys
			100					105					110		
Gln	Arg	Asp	Lys	Asn	Met	Thr	Cys	Ile	Lys	Val	Ser	Ile	Asp	Pro	Glu
		115					120					125			
Ser	Asn	Ile	Ile	Ser	Ile	Trp	Asn	Asn	Gly	Lys	Gly	Ile	Pro	Val	Val
	130					135					140				
Glu	His	Lys	Val	Glu	Lys	Val	Tyr	Val	Pro	Ala	Leu	Ile	Phe	Gly	Gln
145					150					155					160
Leu	Leu	Thr	Ser	Ser	Asn	Tyr	Asp	Asp	Asp	Glu	Lys	Lys	Val	Thr	Gly
				165					170					175	
Gly	Arg	Asn	Gly	Tyr	Gly	Ala	Lys	Leu	Cys	Asn	Ile	Phe	Ser	Thr	Lys
			180					185					190		
Phe	Thr	Val	Glu	Thr	Ala	Cys	Lys	Glu	Tyr	Lys	His	Ser	Phe	Lys	Gln
		195					200					205			
Thr	Trp	Met	Asn	Asn	Met	Met	Lys	Thr	Ser	Glu	Ala	Lys	Ile	Lys	His
	210					215					220				
Phe	Asp	Gly	Glu	Asp	Tyr	Thr	Cys	Ile	Thr	Phe	Gln	Pro	Asp	Leu	Ser
225					230					235					240
Lys	Phe	Lys	Met	Glu	Lys	Leu	Asp	Lys	Asp	Ile	Val	Ala	Leu	Met	Thr
			245						250					255	
Arg	Arg	Ala	Tyr	Asp	Leu	Ala	Gly	Ser	Cys	Arg	Gly	Val	Lys	Val	Met
			260					265					270		
Phe	Asn	Gly	Lys	Lys	Leu	Pro	Val	Asn	Gly	Phe	Arg	Ser	Tyr	Val	Asp
		275					280					285			
Leu	Tyr	Val	Lys	Asp	Lys	Leu	Asp	Glu	Thr	Gly	Val	Ala	Leu	Lys	Val
	290					295					300				
Ile	His	Glu	Leu	Ala	Asn	Glu	Arg	Trp	Asp	Val	Cys	Leu	Thr	Leu	Ser
305					310					315					320
Glu	Lys	Gly	Phe	Gln	Gln	Ile	Ser	Phe	Val	Asn	Ser	Ile	Ala	Thr	Thr
				325					330					335	
Lys	Gly	Gly	Arg	His	Val	Asp	Tyr	Val	Val	Asp	Gln	Val	Val	Gly	Lys
			340					345					350		
Leu	Ile	Glu	Val	Val	Lys	Lys	Lys	Asn	Lys	Ala	Gly	Val	Ser	Val	Lys
		355					360					365			
Pro	Phe	Gln	Val	Lys	Asn	His	Ile	Trp	Val	Phe	Ile	Asn	Cys	Leu	Ile
	370					375					380				
Glu	Asn	Pro	Thr	Phe	Asp	Ser	Gln	Thr	Lys	Glu	Asn	Met	Thr	Leu	Gln
385					390					395					400

Pro	Lys	Ser	Phe	Gly	Ser	Lys	Cys	Gln	Leu	Ser	Glu	Lys	Phe	Phe	Lys			
				405					410					415				
Ala	Ala	Ser	Asn	Cys	Gly	Ile	Val	Glu	Ser	Ile	Leu	Asn	Trp	Val	Lys			
			420					425					430					
Phe	Lys	Ala	Gln	Thr	Gln	Leu	Asn	Lys	Lys	Cys	Ser	Ser	Val	Lys	Tyr			
		435					440					445						
Ser	Lys	Ile	Lys	Gly	Ile	Pro	Lys	Leu	Asp	Asp	Ala	Asn	Asp	Ala	Gly			
	450				455					460								
Gly	Lys	His	Ser	Leu	Glu	Cys	Thr	Leu	Ile	Leu	Thr	Glu	Gly	Asp	Ser			
465				470					475						480			
Ala	Lys	Ser	Leu	Ala	Val	Ser	Gly	Leu	Gly	Val	Ile	Gly	Arg	Asp	Arg			
			485					490						495				
Tyr	Gly	Val	Phe	Pro	Leu	Arg	Gly	Lys	Ile	Leu	Asn	Val	Arg	Glu	Ala			
			500				505						510					
Ser	His	Lys	Gln	Ile	Met	Glu	Asn	Ala	Glu	Ile	Asn	Asn	Ile	Ile	Lys			
	515						520					525						
Ile	Val	Gly	Leu	Gln	Tyr	Lys	Lys	Ser	Tyr	Asp	Asp	Ala	Glu	Ser	Leu			
	530				535					540								
Lys	Thr	Leu	Arg	Tyr	Gly	Lys	Ile	Met	Ile	Met	Thr	Asp	Gln	Asp	Gln			
545				550					555						560			
Asp	Gly	Ser	His	Ile	Lys	Gly	Leu	Leu	Ile	Asn	Phe	Ile	His	His	Asn			
			565					570						575				
Trp	Pro	Ser	Leu	Leu	Lys	His	Gly	Phe	Leu	Glu	Glu	Phe	Ile	Thr	Pro			
			580				585						590					
Ile	Val	Lys	Ala	Ser	Lys	Asn	Lys	Gln	Glu	Leu	Ser	Phe	Tyr	Ser	Ile			
	595					600					605							
Pro	Glu	Phe	Asp	Glu	Trp	Lys	Lys	His	Ile	Glu	Asn	Gln	Lys	Ala	Trp			
	610					615				620								
Lys	Ile	Lys	Tyr	Tyr	Lys	Gly	Leu	Gly	Thr	Ser	Thr	Ala	Lys	Glu	Ala			
625				630					635						640			
Lys	Glu	Tyr	Phe	Ala	Asp	Met	Glu	Arg	His	Arg	Ile	Leu	Phe	Arg	Tyr			
			645					650						655				
Ala	Gly	Pro	Glu	Asp	Asp	Ala	Ala	Ile	Thr	Leu	Ala	Phe	Ser	Lys	Lys			
			660				665						670					
Lys	Ile	Asp	Asp	Arg	Lys	Glu	Trp	Leu	Thr	Asn	Phe	Met	Glu	Asp	Arg			
	675					680						685						
Arg	Gln	Arg	Arg	Leu	His	Gly	Leu	Pro	Glu	Gln	Phe	Leu	Tyr	Gly	Thr			
	690				695				700									
Ala	Thr	Lys	His	Leu	Thr	Tyr	Asn	Asp	Phe	Ile	Asn	Lys	Glu	Leu	Ile			
705				710					715						720			
Leu	Phe	Ser	Asn	Ser	Asp	Asn	Glu	Arg	Ser	Ile	Pro	Ser	Leu	Val	Asp			
			725					730						735				
Gly	Phe	Lys	Pro	Gly	Gln	Arg	Lys	Val	Leu	Phe	Thr	Cys	Phe	Lys	Arg			
			740				745						750					
Asn	Asp	Lys	Arg	Glu	Val	Lys	Val	Ala	Gln	Leu	Ala	Gly	Ser	Val	Ala			
	755					760						765						
Glu	Met	Ser	Ala	Tyr	His	His	Gly	Glu	Gln	Ala	Leu	Met	Met	Thr	Ile			
	770				775					780								
Val	Asn	Leu	Ala	Gln	Asn	Phe	Val	Gly	Ser	Asn	Asn	Ile	Asn	Leu	Leu			
785				790					795						800			
Gln	Pro	Ile	Gly	Gln	Phe	Gly	Thr	Arg	Leu	His	Gly	Gly	Lys	Asp	Ala			
			805					810						815				
Ala	Ser	Pro	Arg	Tyr	Ile	Phe	Thr	Met	Leu	Ser	Thr	Leu	Ala	Arg	Leu			
			820					825					830					

Leu	Phe	Pro	Ala	Val	Asp	Asp	Asn	Leu	Leu	Lys	Phe	Leu	Tyr	Asp	Asp		
		835					840					845					
Asn	Gln	Arg	Val	Glu	Pro	Glu	Trp	Tyr	Ile	Pro	Ile	Ile	Pro	Met	Val		
	850					855					860						
Leu	Ile	Asn	Gly	Ala	Glu	Gly	Ile	Gly	Thr	Gly	Trp	Ala	Cys	Lys	Leu		
865					870					875					880		
Pro	Asn	Tyr	Asp	Ala	Arg	Glu	Ile	Val	Asn	Asn	Val	Arg	Arg	Met	Leu		
			885						890					895			
Asp	Gly	Leu	Asp	Pro	His	Pro	Met	Leu	Pro	Asn	Tyr	Lys	Asn	Phe	Lys		
		900						905					910				
Gly	Thr	Ile	Gln	Glu	Leu	Gly	Gln	Asn	Gln	Tyr	Ala	Val	Ser	Gly	Glu		
	915						920					925					
Ile	Phe	Val	Val	Asp	Arg	Asn	Thr	Val	Glu	Ile	Thr	Glu	Leu	Pro	Val		
	930					935					940						
Arg	Thr	Trp	Thr	Gln	Val	Tyr	Lys	Glu	Gln	Val	Leu	Glu	Pro	Met	Leu		
945				950						955					960		
Asn	Gly	Thr	Asp	Lys	Thr	Pro	Ala	Leu	Ile	Ser	Asp	Tyr	Lys	Glu	Tyr		
			965						970					975			
His	Thr	Asp	Thr	Thr	Val	Lys	Phe	Val	Val	Lys	Met	Thr	Glu	Glu	Lys		
		980						985					990				
Leu	Ala	Gln	Ala	Glu	Ala	Ala	Gly	Leu	His	Lys	Val	Phe	Lys	Leu	Gln		
	995						1000					1005					
Thr	Thr	Leu	Thr	Cys	Asn	Ser	Met	Val	Leu	Phe	Asp	His	Met	Gly	Cys		
	1010					1015					1020						
Leu	Lys	Lys	Tyr	Glu	Thr	Val	Gln	Asp	Ile	Leu	Lys	Glu	Phe	Phe	Asp		
1025				1030					1035						1040		
Leu	Arg	Leu	Ser	Tyr	Tyr	Gly	Leu	Arg	Lys	Glu	Trp	Leu	Val	Gly	Met		
			1045					1050						1055			
Leu	Gly	Ala	Glu	Ser	Thr	Lys	Leu	Asn	Asn	Gln	Ala	Arg	Phe	Ile	Leu		
		1060						1065					1070				
Glu	Lys	Ile	Gln	Gly	Lys	Ile	Thr	Ile	Glu	Asn	Arg	Ser	Lys	Lys	Asp		
	1075						1080					1085					
Leu	Ile	Gln	Met	Leu	Val	Gln	Arg	Gly	Tyr	Glu	Ser	Asp	Pro	Val	Lys		
	1090					1095						1100					
Ala	Trp	Lys	Glu	Ala	Gln	Glu	Lys	Ala	Ala	Glu	Glu	Asp	Glu	Thr	Gln		
1105					1110					1115					1120		
Asn	Gln	His	Asp	Asp	Ser	Ser	Ser	Asp	Ser	Gly	Thr	Pro	Ser	Gly	Pro		
			1125						1130					1135			
Asp	Phe	Asn	Tyr	Ile	Leu	Asn	Met	Ser	Leu	Trp	Ser	Leu	Thr	Lys	Glu		
		1140					1145						1150				
Lys	Val	Glu	Glu	Leu	Ile	Lys	Gln	Arg	Asp	Ala	Lys	Gly	Arg	Glu	Val		
	1155						1160					1165					
Asn	Asp	Leu	Lys	Arg	Lys	Ser	Pro	Ser	Asp	Leu	Trp	Lys	Glu	Asp	Leu		
	1170					1175					1180						
Ala	Ala	Phe	Val	Glu	Glu	Leu	Asp	Lys	Val	Glu	Ser	Gln	Glu	Arg	Glu		
1185					1190					1195					1200		
Asp	Val	Leu	Ala	Gly	Met	Ser	Gly	Lys	Ala	Ile	Lys	Gly	Lys	Val	Gly		
			1205					1210						1215			
Lys	Pro	Lys	Val	Lys	Lys	Leu	Gln	Leu	Glu	Glu	Thr	Met	Pro	Ser	Pro		
		1220						1225					1230				
Tyr	Gly	Arg	Arg	Ile	Ile	Pro	Glu	Ile	Thr	Ala	Met	Lys	Ala	Asp	Ala		
	1235						1240					1245					
Ser	Lys	Lys	Leu	Leu	Lys	Lys	Lys	Lys	Gly	Asp	Leu	Asp	Thr	Ala	Ala		
	1250					1255					1260						

Val Lys Val Glu Phe Asp Glu Glu Phe Ser Gly Ala Pro Val Glu Gly
1265 1270 1275 1280
Ala Gly Glu Glu Ala Leu Thr Pro Ser Val Pro Ile Asn Lys Gly Pro
1285 1290 1295
Lys Pro Lys Arg Glu Lys Lys Glu Pro Gly Thr Arg Val Arg Lys Thr
1300 1305 1310
Pro Thr Ser Ser Gly Lys Pro Ser Ala Lys Lys Val Lys Lys Arg Asn
1315 1320 1325
Pro Trp Ser Asp Asp Glu Ser Lys Ser Glu Ser Asp Leu Glu Glu Thr
1330 1335 1340
Glu Pro Val Val Ile Pro Arg Asp Ser Leu Leu Arg Arg Ala Ala Ala
1345 1350 1355 1360
Glu Arg Pro Lys Tyr Thr Phe Asp Phe Ser Glu Glu Glu Asp Asp Asp
1365 1370 1375
Ala Asp Asp Asp Asp Asp Asp Asn Asn Asp Leu Glu Glu Leu Lys Val
1380 1385 1390
Lys Ala Ser Pro Ile Thr Asn Asp Gly Glu Asp Glu Phe Val Pro Ser
1395 1400 1405
Asp Gly Leu Asp Lys Asp Glu Tyr Thr Phe Ser Pro Gly Lys Ser Lys
1410 1415 1420
Ala Thr Pro Glu Lys Ser Leu His Asp Lys Lys Ser Gln Asp Phe Gly
1425 1430 1435 1440
Asn Leu Phe Ser Phe Pro Ser Tyr Ser Gln Lys Ser Glu Asp Asp Ser
1445 1450 1455
Ala Lys Phe Asp Ser Asn Glu Glu Asp Ser Ala Ser Val Phe Ser Pro
1460 1465 1470
Ser Phe Gly Leu Lys Gln Thr Asp Lys Val Pro Ser Lys Thr Val Ala
1475 1480 1485
Ala Lys Lys Gly Lys Pro Ser Ser Asp Thr Val Pro Lys Pro Lys Arg
1490 1495 1500
Ala Pro Lys Gln Lys Lys Val Val Glu Ala Val Asn Ser Asp Ser Asp
1505 1510 1515 1520
Ser Glu Phe Gly Ile Pro Lys Lys Thr Thr Thr Pro Lys Gly Lys Gly
1525 1530 1535
Arg Gly Ala Lys Lys Arg Lys Ala Ser Gly Ser Glu Asn Glu Gly Asp
1540 1545 1550
Tyr Asn Pro Gly Arg Lys Thr Ser Lys Thr Thr Ser Lys Lys Pro Lys
1555 1560 1565
Lys Thr Ser Phe Asp Gln Asp Ser Asp Val Asp Ile Phe Pro Ser Asp
1570 1575 1580
Phe Pro Thr Glu Pro Pro Ser Leu Pro Arg Thr Gly Arg Ala Arg Lys
1585 1590 1595 1600
Glu Val Lys Tyr Phe Ala Glu Ser Asp Glu Glu Glu Asp Asp Val Asp
1605 1610 1615
Phe Ala Met Phe Asn
1620

<210> 208

<211> 1522

<212> PRT

<213> Homo sapiens

<220>

<223> >TOPBP1|ENSG00000163781|ENST00000260810|4569

<400> 208

Met	Ser	Arg	Asn	Asp	Lys	Glu	Pro	Phe	Phe	Val	Lys	Phe	Leu	Lys	Ser	1	5	10	15
Ser	Asp	Asn	Ser	Lys	Cys	Phe	Phe	Lys	Ala	Leu	Glu	Ser	Ile	Lys	Glu	20	25	30	
Phe	Gln	Ser	Glu	Glu	Tyr	Leu	Gln	Ile	Ile	Thr	Glu	Glu	Glu	Ala	Leu	35	40	45	
Lys	Ile	Lys	Glu	Asn	Asp	Arg	Ser	Leu	Tyr	Ile	Cys	Asp	Pro	Phe	Ser	50	55	60	
Gly	Val	Val	Phe	Asp	His	Leu	Lys	Lys	Leu	Gly	Cys	Arg	Ile	Val	Gly	65	70	75	80
Pro	Gln	Val	Val	Ile	Phe	Cys	Met	His	His	Gln	Arg	Cys	Val	Pro	Arg	85	90	95	
Ala	Glu	His	Pro	Val	Tyr	Asn	Met	Val	Met	Ser	Asp	Val	Thr	Ile	Ser	100	105	110	
Cys	Thr	Ser	Leu	Glu	Lys	Glu	Lys	Arg	Glu	Glu	Val	His	Lys	Tyr	Val	115	120	125	
Gln	Met	Met	Gly	Gly	Arg	Val	Tyr	Arg	Asp	Leu	Asn	Val	Ser	Val	Thr	130	135	140	
His	Leu	Ile	Ala	Gly	Glu	Val	Gly	Ser	Lys	Lys	Tyr	Leu	Val	Ala	Ala	145	150	155	160
Asn	Leu	Lys	Lys	Pro	Ile	Leu	Leu	Pro	Ser	Trp	Ile	Lys	Thr	Leu	Trp	165	170	175	
Glu	Lys	Ser	Gln	Glu	Lys	Lys	Ile	Thr	Arg	Tyr	Thr	Asp	Ile	Asn	Met	180	185	190	
Glu	Asp	Phe	Lys	Cys	Pro	Ile	Phe	Leu	Gly	Cys	Ile	Ile	Cys	Val	Thr	195	200	205	
Gly	Leu	Cys	Gly	Leu	Asp	Arg	Lys	Glu	Val	Gln	Gln	Leu	Thr	Val	Lys	210	215	220	
His	Gly	Gly	Gln	Tyr	Met	Gly	Gln	Leu	Lys	Met	Asn	Glu	Cys	Thr	His	225	230	235	240
Leu	Ile	Val	Gln	Glu	Pro	Lys	Gly	Gln	Lys	Tyr	Glu	Cys	Ala	Lys	Arg	245	250	255	
Trp	Asn	Val	His	Cys	Val	Thr	Thr	Gln	Trp	Phe	Phe	Asp	Ser	Ile	Glu	260	265	270	
Lys	Gly	Phe	Cys	Gln	Asp	Glu	Ser	Ile	Tyr	Lys	Thr	Glu	Pro	Arg	Pro	275	280	285	
Glu	Ala	Lys	Thr	Met	Pro	Asn	Ser	Ser	Thr	Pro	Thr	Ser	Gln	Ile	Asn	290	295	300	
Thr	Ile	Asp	Ser	Arg	Thr	Leu	Ser	Asp	Val	Ser	Asn	Ile	Ser	Asn	Ile	305	310	315	320
Asn	Ala	Ser	Cys	Val	Ser	Glu	Ser	Ile	Cys	Asn	Ser	Leu	Asn	Ser	Lys	325	330	335	
Leu	Glu	Pro	Thr	Leu	Glu	Asn	Leu	Glu	Asn	Leu	Asp	Val	Ser	Ala	Phe	340	345	350	
Gln	Ala	Pro	Glu	Asp	Leu	Leu	Asp	Gly	Cys	Arg	Ile	Tyr	Leu	Cys	Gly	355	360	365	
Phe	Ser	Gly	Arg	Lys	Leu	Asp	Lys	Leu	Arg	Arg	Leu	Ile	Asn	Ser	Gly	370	375	380	
Gly	Gly	Val	Arg	Phe	Asn	Gln	Leu	Asn	Glu	Asp	Val	Thr	His	Val	Ile	385	390	395	400
Val	Gly	Asp	Tyr	Asp	Asp	Glu	Leu	Lys	Gln	Phe	Trp	Asn	Lys	Ser	Ala				

				405				410				415			
His	Arg	Pro	His	Val	Val	Gly	Ala	Lys	Trp	Leu	Leu	Glu	Cys	Phe	Ser
			420					425					430		
Lys	Gly	Tyr	Met	Leu	Ser	Glu	Glu	Pro	Tyr	Ile	His	Ala	Asn	Tyr	Gln
		435					440					445			
Pro	Val	Glu	Ile	Pro	Val	Ser	His	Lys	Pro	Glu	Ser	Lys	Ala	Ala	Leu
	450					455					460				
Leu	Lys	Lys	Lys	Asn	Ser	Ser	Phe	Ser	Lys	Lys	Asp	Phe	Ala	Pro	Ser
465				470						475					480
Glu	Lys	His	Glu	Gln	Ala	Asp	Glu	Asp	Leu	Leu	Ser	Gln	Tyr	Glu	Asn
			485					490						495	
Gly	Ser	Ser	Thr	Val	Val	Glu	Ala	Lys	Thr	Ser	Glu	Ala	Arg	Pro	Phe
			500					505					510		
Asn	Asp	Ser	Thr	His	Ala	Glu	Pro	Leu	Asn	Asp	Ser	Thr	His	Ile	Ser
		515					520					525			
Leu	Gln	Glu	Glu	Asn	Gln	Ser	Ser	Val	Ser	His	Cys	Val	Pro	Asp	Val
	530				535						540				
Ser	Thr	Ile	Thr	Glu	Glu	Gly	Leu	Phe	Ser	Gln	Lys	Ser	Phe	Leu	Val
545					550					555					560
Leu	Gly	Phe	Ser	Asn	Glu	Asn	Glu	Ser	Asn	Ile	Ala	Asn	Ile	Ile	Lys
				565					570					575	
Glu	Asn	Ala	Gly	Lys	Ile	Met	Ser	Leu	Ser	Arg	Thr	Val	Ala	Asp	
		580						585				590			
Tyr	Ala	Val	Val	Pro	Leu	Leu	Gly	Cys	Glu	Val	Glu	Ala	Thr	Val	Gly
	595						600					605			
Glu	Val	Val	Thr	Asn	Thr	Trp	Leu	Val	Thr	Cys	Ile	Asp	Tyr	Gln	Thr
	610					615					620				
Leu	Phe	Asp	Pro	Lys	Ser	Asn	Pro	Leu	Phe	Thr	Pro	Val	Pro	Val	Met
625					630					635					640
Thr	Gly	Met	Thr	Pro	Leu	Glu	Asp	Cys	Val	Ile	Ser	Phe	Ser	Gln	Cys
				645					650					655	
Ala	Gly	Ala	Glu	Lys	Glu	Ser	Leu	Thr	Phe	Leu	Ala	Asn	Leu	Leu	Gly
		660						665					670		
Ala	Ser	Val	Gln	Glu	Tyr	Phe	Val	Arg	Lys	Ser	Asn	Ala	Lys	Lys	Gly
		675					680					685			
Met	Phe	Ala	Ser	Thr	His	Leu	Ile	Leu	Lys	Glu	Arg	Gly	Gly	Ser	Lys
	690					695					700				
Tyr	Glu	Ala	Ala	Lys	Lys	Trp	Asn	Leu	Pro	Ala	Val	Thr	Ile	Ala	Trp
705					710					715					720
Leu	Leu	Glu	Thr	Ala	Arg	Thr	Gly	Lys	Arg	Ala	Asp	Glu	Ser	His	Phe
				725					730					735	
Leu	Ile	Glu	Asn	Ser	Thr	Lys	Glu	Glu	Arg	Ser	Leu	Glu	Thr	Glu	Ile
			740					745					750		
Thr	Asn	Gly	Ile	Asn	Leu	Asn	Ser	Asp	Thr	Ala	Glu	His	Pro	Gly	Thr
		755					760					765			
Arg	Leu	Gln	Thr	His	Arg	Lys	Thr	Val	Val	Thr	Pro	Leu	Asp	Met	Asn
	770					775					780				
Arg	Phe	Gln	Ser	Lys	Ala	Phe	Arg	Ala	Val	Val	Ser	Gln	His	Ala	Arg
785					790					795					800
Gln	Val	Ala	Ala	Ser	Pro	Ala	Val	Gly	Gln	Pro	Leu	Gln	Lys	Glu	Pro
			805					810						815	
Ser	Leu	His	Leu	Asp	Thr	Pro	Ser	Lys	Phe	Leu	Ser	Lys	Asp	Lys	Leu
		820						825					830		
Phe	Lys	Pro	Ser	Phe	Asp	Val	Lys	Asp	Ala	Leu	Ala	Ala	Leu	Glu	Thr

1265		1270		1275		1280
Ile Asp Tyr Cys His	Leu Ile Glu Lys	Leu Gly Gly Leu Val	Ile Glu			
	1285	1290	1295			
Lys Gln Cys Phe Asp	Pro Thr Cys Thr	His Ile Val Val	Gly His Pro			
	1300	1305	1310			
Leu Arg Asn Glu Lys	Tyr Leu Ala Ser	Val Ala Ala Gly	Lys Trp Val			
	1315	1320	1325			
Leu His Arg Ser Tyr	Leu Glu Ala Cys	Arg Thr Ala Gly	His Phe Val			
	1330	1335	1340			
Gln Glu Glu Asp Tyr	Glu Trp Gly Ser	Ser Ser Ile Leu	Asp Val Leu			
1345	1350	1355	1360			
Thr Gly Ile Asn Val	Gln Gln Arg Arg	Leu Ala Leu Ala	Ala Met Arg			
	1365	1370	1375			
Trp Arg Lys Lys Ile	Gln Gln Arg Gln	Glu Ser Gly Ile	Val Glu Gly			
	1380	1385	1390			
Ala Phe Ser Gly Trp	Lys Val Ile Leu	His Val Asp Gln	Ser Arg Glu			
	1395	1400	1405			
Ala Gly Phe Lys Arg	Leu Leu Gln Ser	Gly Gly Ala Lys	Val Leu Pro			
	1410	1415	1420			
Gly His Ser Val Pro	Leu Phe Lys Glu	Ala Thr His Leu	Phe Ser Asp			
1425	1430	1435	1440			
Leu Asn Lys Leu Lys	Pro Asp Asp Ser	Gly Val Asn Ile	Ala Glu Ala			
	1445	1450	1455			
Ala Ala Gln Asn Val	Tyr Cys Leu Arg	Thr Glu Tyr Ile	Ala Asp Tyr			
	1460	1465	1470			
Leu Met Gln Glu Ser	Pro Pro His Val	Glu Asn Tyr Cys	Leu Pro Glu			
	1475	1480	1485			
Ala Ile Ser Phe Ile	Gln Asn Asn Lys	Glu Leu Gly Thr	Gly Leu Ser			
	1490	1495	1500			
Gln Lys Arg Lys Ala	Pro Thr Glu Lys	Asn Lys Ile Lys	Arg Pro Arg			
1505	1510	1515	1520			
Val His						

<210> 209
 <211> 1977
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >TP53BP1|ENSG00000067369|ENST00000382044|5934

<400> 209															
Met Pro Gly Glu Gln Met Asp Pro Thr Gly Ser Gln Leu Asp Ser Asp															
1				5				10					15		
Phe Ser Gln Gln Asp Thr Pro Cys Leu Ile Ile Glu Asp Ser Gln Pro															
			20				25					30			
Glu Ser Gln Val Leu Glu Asp Asp Ser Gly Ser His Phe Ser Met Leu															
			35			40					45				
Ser Arg His Leu Pro Asn Leu Gln Thr His Lys Glu Asn Pro Val Leu															
			50		55			60							
Asp Val Val Ser Asn Pro Glu Gln Thr Ala Gly Glu Glu Arg Gly Asp															
65				70			75						80		

Gly	Asn	Ser	Gly	Phe	Asn	Glu	His	Leu	Lys	Glu	Asn	Lys	Val	Ala	Asp	85	90	95
Pro	Val	Asp	Ser	Ser	Asn	Leu	Asp	Thr	Cys	Gly	Ser	Ile	Ser	Gln	Val	100	105	110
Ile	Glu	Gln	Leu	Pro	Gln	Pro	Asn	Arg	Thr	Ser	Ser	Val	Leu	Gly	Met	115	120	125
Ser	Val	Glu	Ser	Ala	Pro	Ala	Val	Glu	Glu	Glu	Lys	Gly	Glu	Glu	Leu	130	135	140
Glu	Gln	Lys	Glu	Lys	Glu	Lys	Glu	Glu	Asp	Thr	Ser	Gly	Asn	Thr	Thr	145	150	155
His	Ser	Leu	Gly	Ala	Glu	Asp	Thr	Ala	Ser	Ser	Gln	Leu	Gly	Phe	Gly	165	170	175
Val	Leu	Glu	Leu	Ser	Gln	Ser	Gln	Asp	Val	Glu	Glu	Asn	Thr	Val	Pro	180	185	190
Tyr	Glu	Val	Asp	Lys	Glu	Gln	Leu	Gln	Ser	Val	Thr	Thr	Asn	Ser	Gly	195	200	205
Tyr	Thr	Arg	Leu	Ser	Asp	Val	Asp	Ala	Asn	Thr	Ala	Ile	Lys	His	Glu	210	215	220
Glu	Gln	Ser	Asn	Glu	Asp	Ile	Pro	Ile	Ala	Glu	Gln	Ser	Ser	Lys	Asp	225	230	235
Ile	Pro	Val	Thr	Ala	Gln	Pro	Ser	Lys	Asp	Val	His	Val	Val	Lys	Glu	245	250	255
Gln	Asn	Pro	Pro	Pro	Ala	Arg	Ser	Glu	Asp	Met	Pro	Phe	Ser	Pro	Lys	260	265	270
Ala	Ser	Val	Ala	Ala	Met	Glu	Ala	Lys	Glu	Gln	Leu	Ser	Ala	Gln	Glu	275	280	285
Leu	Met	Glu	Ser	Gly	Leu	Gln	Ile	Gln	Lys	Ser	Pro	Glu	Pro	Glu	Val	290	295	300
Leu	Ser	Thr	Gln	Glu	Asp	Leu	Phe	Asp	Gln	Ser	Asn	Lys	Thr	Val	Ser	305	310	315
Ser	Asp	Gly	Cys	Ser	Thr	Pro	Ser	Arg	Glu	Glu	Gly	Gly	Cys	Ser	Leu	325	330	335
Ala	Ser	Thr	Pro	Ala	Thr	Thr	Leu	His	Leu	Leu	Gln	Leu	Ser	Gly	Gln	340	345	350
Arg	Ser	Leu	Val	Gln	Asp	Ser	Leu	Ser	Thr	Asn	Ser	Ser	Asp	Leu	Val	355	360	365
Ala	Pro	Ser	Pro	Asp	Ala	Phe	Arg	Ser	Thr	Pro	Phe	Ile	Val	Pro	Ser	370	375	380
Ser	Pro	Thr	Glu	Gln	Glu	Gly	Arg	Gln	Asp	Lys	Pro	Met	Asp	Thr	Ser	385	390	395
Val	Leu	Ser	Glu	Glu	Gly	Gly	Glu	Pro	Phe	Gln	Lys	Lys	Leu	Gln	Ser	405	410	415
Gly	Glu	Pro	Val	Glu	Leu	Glu	Asn	Pro	Pro	Leu	Leu	Pro	Glu	Ser	Thr	420	425	430
Val	Ser	Pro	Gln	Ala	Ser	Thr	Pro	Ile	Ser	Gln	Ser	Thr	Pro	Val	Phe	435	440	445
Pro	Pro	Gly	Ser	Leu	Pro	Ile	Pro	Ser	Gln	Pro	Gln	Phe	Ser	His	Asp	450	455	460
Ile	Phe	Ile	Pro	Ser	Pro	Ser	Leu	Glu	Glu	Gln	Ser	Asn	Asp	Gly	Lys	465	470	475
Lys	Asp	Gly	Asp	Met	His	Ser	Ser	Ser	Leu	Thr	Val	Glu	Cys	Ser	Lys	485	490	495
Thr	Ser	Glu	Ile	Glu	Pro	Lys	Asn	Ser	Pro	Glu	Asp	Leu	Gly	Leu	Ser	500	505	510

Leu	Thr	Gly	Asp	Ser	Cys	Lys	Leu	Met	Leu	Ser	Thr	Ser	Glu	Tyr	Ser	515	520	525
Gln	Ser	Pro	Lys	Met	Glu	Ser	Leu	Ser	Ser	His	Arg	Ile	Asp	Glu	Asp	530	535	540
Gly	Glu	Asn	Thr	Gln	Ile	Glu	Asp	Thr	Glu	Pro	Met	Ser	Pro	Val	Leu	545	550	555
Asn	Ser	Lys	Phe	Val	Pro	Ala	Glu	Asn	Asp	Ser	Ile	Leu	Met	Asn	Pro	565	570	575
Ala	Gln	Asp	Gly	Glu	Val	Gln	Leu	Ser	Gln	Asn	Asp	Asp	Lys	Thr	Lys	580	585	590
Gly	Asp	Asp	Thr	Asp	Thr	Arg	Asp	Asp	Ile	Ser	Ile	Leu	Ala	Thr	Gly	595	600	605
Cys	Lys	Gly	Arg	Glu	Glu	Thr	Val	Ala	Glu	Asp	Val	Cys	Ile	Asp	Leu	610	615	620
Thr	Cys	Asp	Ser	Gly	Ser	Gln	Ala	Val	Pro	Ser	Pro	Ala	Thr	Arg	Ser	625	630	635
Glu	Ala	Leu	Ser	Ser	Val	Leu	Asp	Gln	Glu	Glu	Ala	Met	Glu	Ile	Lys	645	650	655
Glu	His	His	Pro	Glu	Glu	Gly	Ser	Ser	Gly	Ser	Glu	Val	Glu	Glu	Ile	660	665	670
Pro	Glu	Thr	Pro	Cys	Glu	Ser	Gln	Gly	Glu	Glu	Leu	Lys	Glu	Glu	Asn	675	680	685
Met	Glu	Ser	Val	Pro	Leu	His	Leu	Ser	Leu	Thr	Glu	Thr	Gln	Ser	Gln	690	695	700
Gly	Leu	Cys	Leu	Gln	Lys	Glu	Met	Pro	Lys	Lys	Glu	Cys	Ser	Glu	Ala	705	710	715
Met	Glu	Val	Glu	Thr	Ser	Val	Ile	Ser	Ile	Asp	Ser	Pro	Gln	Lys	Leu	725	730	735
Ala	Ile	Leu	Asp	Gln	Glu	Leu	Glu	His	Lys	Glu	Gln	Glu	Ala	Trp	Glu	740	745	750
Glu	Ala	Thr	Ser	Glu	Asp	Ser	Ser	Val	Val	Ile	Val	Asp	Val	Lys	Glu	755	760	765
Pro	Ser	Pro	Arg	Val	Asp	Val	Ser	Cys	Glu	Pro	Leu	Glu	Gly	Val	Glu	770	775	780
Lys	Cys	Ser	Asp	Ser	Gln	Ser	Trp	Glu	Asp	Ile	Ala	Pro	Glu	Ile	Glu	785	790	795
Pro	Cys	Ala	Glu	Asn	Arg	Leu	Asp	Thr	Lys	Glu	Glu	Lys	Ser	Val	Glu	805	810	815
Tyr	Glu	Gly	Asp	Leu	Lys	Ser	Gly	Thr	Ala	Glu	Thr	Glu	Pro	Val	Glu	820	825	830
Gln	Asp	Ser	Ser	Gln	Pro	Ser	Leu	Pro	Leu	Val	Arg	Ala	Asp	Asp	Pro	835	840	845
Leu	Arg	Leu	Asp	Gln	Glu	Leu	Gln	Gln	Pro	Gln	Thr	Gln	Glu	Lys	Thr	850	855	860
Ser	Asn	Ser	Leu	Thr	Glu	Asp	Ser	Lys	Met	Ala	Asn	Ala	Lys	Gln	Leu	865	870	875
Ser	Ser	Asp	Ala	Glu	Ala	Gln	Lys	Leu	Gly	Lys	Pro	Ser	Ala	His	Ala	885	890	895
Ser	Gln	Ser	Phe	Cys	Glu	Ser	Ser	Ser	Glu	Thr	Pro	Phe	His	Phe	Thr	900	905	910
Leu	Pro	Lys	Glu	Gly	Asp	Ile	Ile	Pro	Pro	Leu	Thr	Gly	Ala	Thr	Pro	915	920	925
Pro	Leu	Ile	Gly	His	Leu	Lys	Leu	Glu	Pro	Lys	Arg	His	Ser	Thr	Pro	930	935	940

Ile	Gly	Ile	Ser	Asn	Tyr	Pro	Glu	Ser	Thr	Ile	Ala	Thr	Ser	Asp	Val	945	950	955	960
Met	Ser	Glu	Ser	Met	Val	Glu	Thr	His	Asp	Pro	Ile	Leu	Gly	Ser	Gly	965	970	975	
Lys	Gly	Asp	Ser	Gly	Ala	Ala	Pro	Asp	Val	Asp	Asp	Lys	Leu	Cys	Leu	980	985	990	
Arg	Met	Lys	Leu	Val	Ser	Pro	Glu	Thr	Glu	Ala	Ser	Glu	Glu	Ser	Leu	995	1000	1005	
Gln	Phe	Asn	Leu	Glu	Lys	Pro	Ala	Thr	Gly	Glu	Arg	Lys	Asn	Gly	Ser	1010	1015	1020	
Thr	Ala	Val	Ala	Glu	Ser	Val	Ala	Ser	Pro	Gln	Lys	Thr	Met	Ser	Val	1025	1030	1035	1040
Leu	Ser	Cys	Ile	Cys	Glu	Ala	Arg	Gln	Glu	Asn	Glu	Ala	Arg	Ser	Glu	1045	1050	1055	
Asp	Pro	Pro	Thr	Thr	Pro	Ile	Arg	Gly	Asn	Leu	Leu	His	Phe	Pro	Ser	1060	1065	1070	
Ser	Gln	Gly	Glu	Glu	Glu	Lys	Glu	Lys	Leu	Glu	Gly	Asp	His	Thr	Ile	1075	1080	1085	
Arg	Gln	Ser	Gln	Gln	Pro	Met	Lys	Pro	Ile	Ser	Pro	Val	Lys	Asp	Pro	1090	1095	1100	
Val	Ser	Pro	Ala	Ser	Gln	Lys	Met	Val	Ile	Gln	Gly	Pro	Ser	Ser	Pro	1105	1110	1115	1120
Gln	Gly	Glu	Ala	Met	Val	Thr	Asp	Val	Leu	Glu	Asp	Gln	Lys	Glu	Gly	1125	1130	1135	
Arg	Ser	Thr	Asn	Lys	Glu	Asn	Pro	Ser	Lys	Ala	Leu	Ile	Glu	Arg	Pro	1140	1145	1150	
Ser	Gln	Asn	Asn	Ile	Gly	Ile	Gln	Thr	Met	Glu	Cys	Ser	Leu	Arg	Val	1155	1160	1165	
Pro	Glu	Thr	Val	Ser	Ala	Ala	Thr	Gln	Thr	Ile	Lys	Asn	Val	Cys	Glu	1170	1175	1180	
Gln	Gly	Thr	Ser	Thr	Val	Asp	Gln	Asn	Phe	Gly	Lys	Gln	Asp	Ala	Thr	1185	1190	1195	1200
Val	Gln	Thr	Glu	Arg	Gly	Ser	Gly	Glu	Lys	Pro	Val	Ser	Ala	Pro	Gly	1205	1210	1215	
Asp	Asp	Thr	Glu	Ser	Leu	His	Ser	Gln	Gly	Glu	Glu	Glu	Phe	Asp	Met	1220	1225	1230	
Pro	Gln	Pro	Pro	His	Gly	His	Val	Leu	His	Arg	His	Met	Arg	Thr	Ile	1235	1240	1245	
Arg	Glu	Val	Arg	Thr	Leu	Val	Thr	Arg	Val	Ile	Thr	Asp	Val	Tyr	Tyr	1250	1255	1260	
Val	Asp	Gly	Thr	Glu	Val	Glu	Arg	Lys	Val	Thr	Glu	Glu	Thr	Glu	Glu	1265	1270	1275	1280
Pro	Ile	Val	Glu	Cys	Gln	Glu	Cys	Glu	Thr	Glu	Val	Ser	Pro	Ser	Gln	1285	1290	1295	
Thr	Gly	Gly	Ser	Ser	Gly	Asp	Leu	Gly	Asp	Ile	Ser	Ser	Phe	Ser	Ser	1300	1305	1310	
Lys	Ala	Ser	Ser	Leu	His	Arg	Thr	Ser	Ser	Gly	Thr	Ser	Leu	Ser	Ala	1315	1320	1325	
Met	His	Ser	Ser	Gly	Ser	Ser	Gly	Lys	Gly	Ala	Gly	Pro	Leu	Arg	Gly	1330	1335	1340	
Lys	Thr	Ser	Gly	Thr	Glu	Pro	Ala	Asp	Phe	Ala	Leu	Pro	Ser	Ser	Arg	1345	1350	1355	1360
Gly	Gly	Pro	Gly	Lys	Leu	Ser	Pro	Arg	Lys	Gly	Val	Ser	Gln	Thr	Gly	1365	1370	1375	

Thr	Pro	Val	Cys	Glu	Glu	Asp	Gly	Asp	Ala	Gly	Leu	Gly	Ile	Arg	Gln		
			1380					1385					1390				
Gly	Gly	Lys	Ala	Pro	Val	Thr	Pro	Arg	Gly	Arg	Gly	Arg	Arg	Gly	Arg		
		1395					1400				1405						
Pro	Pro	Ser	Arg	Thr	Thr	Gly	Thr	Arg	Glu	Thr	Ala	Val	Pro	Gly	Pro		
		1410				1415					1420						
Leu	Gly	Ile	Glu	Asp	Ile	Ser	Pro	Asn	Leu	Ser	Pro	Asp	Asp	Lys	Ser		
1425				1430					1435					1440			
Phe	Ser	Arg	Val	Val	Pro	Arg	Val	Pro	Asp	Ser	Thr	Arg	Arg	Thr	Asp		
		1445						1450			1455						
Val	Gly	Ala	Gly	Ala	Leu	Arg	Arg	Ser	Asp	Ser	Pro	Glu	Ile	Pro	Phe		
		1460						1465			1470						
Gln	Ala	Ala	Ala	Gly	Pro	Ser	Asp	Gly	Leu	Asp	Ala	Ser	Ser	Pro	Gly		
		1475					1480				1485						
Asn	Ser	Phe	Val	Gly	Leu	Arg	Val	Val	Ala	Lys	Trp	Ser	Ser	Asn	Gly		
1490					1495				1500								
Tyr	Phe	Tyr	Ser	Gly	Lys	Ile	Thr	Arg	Asp	Val	Gly	Ala	Gly	Lys	Tyr		
1505				1510					1515					1520			
Lys	Leu	Leu	Phe	Asp	Asp	Gly	Tyr	Glu	Cys	Asp	Val	Leu	Gly	Lys	Asp		
			1525					1530					1535				
Ile	Leu	Leu	Cys	Asp	Pro	Ile	Pro	Leu	Asp	Thr	Glu	Val	Thr	Ala	Leu		
		1540					1545						1550				
Ser	Glu	Asp	Glu	Tyr	Phe	Ser	Ala	Gly	Val	Val	Lys	Gly	His	Arg	Lys		
		1555					1560				1565						
Glu	Ser	Gly	Glu	Leu	Tyr	Tyr	Ser	Ile	Glu	Lys	Glu	Gly	Gln	Arg	Lys		
1570				1575					1580								
Trp	Tyr	Lys	Arg	Met	Ala	Val	Ile	Leu	Ser	Leu	Glu	Gln	Gly	Asn	Arg		
1585				1590					1595					1600			
Leu	Arg	Glu	Gln	Tyr	Gly	Leu	Gly	Pro	Tyr	Glu	Ala	Val	Thr	Pro	Leu		
		1605						1610					1615				
Thr	Lys	Ala	Ala	Asp	Ile	Ser	Leu	Asp	Asn	Leu	Val	Glu	Gly	Lys	Arg		
		1620						1625					1630				
Lys	Arg	Arg	Ser	Asn	Val	Ser	Ser	Pro	Ala	Thr	Pro	Thr	Ala	Ser	Ser		
		1635					1640				1645						
Ser	Ser	Ser	Thr	Thr	Pro	Thr	Arg	Lys	Ile	Thr	Glu	Ser	Pro	Arg	Ala		
		1650				1655					1660						
Ser	Met	Gly	Val	Leu	Ser	Gly	Lys	Arg	Lys	Leu	Ile	Thr	Ser	Glu	Glu		
1665				1670					1675					1680			
Glu	Arg	Ser	Pro	Ala	Lys	Arg	Gly	Arg	Lys	Ser	Ala	Thr	Val	Lys	Pro		
		1685						1690					1695				
Gly	Ala	Val	Gly	Ala	Gly	Glu	Phe	Val	Ser	Pro	Cys	Glu	Ser	Gly	Asp		
		1700						1705					1710				
Asn	Thr	Gly	Glu	Pro	Ser	Ala	Leu	Glu	Glu	Gln	Arg	Gly	Pro	Leu	Pro		
		1715					1720				1725						
Leu	Asn	Lys	Thr	Leu	Phe	Leu	Gly	Tyr	Ala	Phe	Leu	Leu	Thr	Met	Ala		
1730						1735					1740						
Thr	Thr	Ser	Asp	Lys	Leu	Ala	Ser	Arg	Ser	Lys	Leu	Pro	Asp	Gly	Pro		
1745				1750					1755					1760			
Thr	Gly	Ser	Ser	Glu	Glu	Glu	Glu	Glu	Phe	Leu	Glu	Ile	Pro	Pro	Phe		
			1765					1770					1775				
Asn	Lys	Gln	Tyr	Thr	Glu	Ser	Gln	Leu	Arg	Ala	Gly	Ala	Gly	Tyr	Ile		
		1780					1785						1790				
Leu	Glu	Asp	Phe	Asn	Glu	Ala	Gln	Cys	Asn	Thr	Ala	Tyr	Gln	Cys	Leu		
		1795					1800						1805				

Leu Ile Ala Asp Gln His Cys Arg Thr Arg Lys Tyr Phe Leu Cys Leu
 1810 1815 1820
 Ala Ser Gly Ile Pro Cys Val Ser His Val Trp Val His Asp Ser Cys
 1825 1830 1835 1840
 His Ala Asn Gln Leu Gln Asn Tyr Arg Asn Tyr Leu Leu Pro Ala Gly
 1845 1850 1855
 Tyr Ser Leu Glu Glu Gln Arg Ile Leu Asp Trp Gln Pro Arg Glu Asn
 1860 1865 1870
 Pro Phe Gln Asn Leu Lys Val Leu Leu Val Ser Asp Gln Gln Gln Asn
 1875 1880 1885
 Phe Leu Glu Leu Trp Ser Glu Ile Leu Met Thr Gly Gly Ala Ala Ser
 1890 1895 1900
 Val Lys Gln His His Ser Ser Ala His Asn Lys Asp Ile Ala Leu Gly
 1905 1910 1915 1920
 Val Phe Asp Val Val Val Thr Asp Pro Ser Cys Pro Ala Ser Val Leu
 1925 1930 1935
 Lys Cys Ala Glu Ala Leu Gln Leu Pro Val Val Ser Gln Glu Trp Val
 1940 1945 1950
 Ile Gln Cys Leu Ile Val Gly Glu Arg Ile Gly Phe Lys Gln His Pro
 1955 1960 1965
 Lys Tyr Lys His Asp Tyr Val Ser His
 1970 1975

<210> 210
 <211> 393
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >TP53|ENSG00000141510|ENST00000269305|1182

<400> 210
 Met Glu Glu Pro Gln Ser Asp Pro Ser Val Glu Pro Pro Leu Ser Gln
 1 5 10 15
 Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu Asn Asn Val Leu
 20 25 30
 Ser Pro Leu Pro Ser Gln Ala Met Asp Asp Leu Met Leu Ser Pro Asp
 35 40 45
 Asp Ile Glu Gln Trp Phe Thr Glu Asp Pro Gly Pro Asp Glu Ala Pro
 50 55 60
 Arg Met Pro Glu Ala Ala Pro Pro Val Ala Pro Ala Pro Ala Ala Pro
 65 70 75 80
 Thr Pro Ala Ala Pro Ala Pro Ala Pro Ser Trp Pro Leu Ser Ser Ser
 85 90 95
 Val Pro Ser Gln Lys Thr Tyr Gln Gly Ser Tyr Gly Phe Arg Leu Gly
 100 105 110
 Phe Leu His Ser Gly Thr Ala Lys Ser Val Thr Cys Thr Tyr Ser Pro
 115 120 125
 Ala Leu Asn Lys Met Phe Cys Gln Leu Ala Lys Thr Cys Pro Val Gln
 130 135 140
 Leu Trp Val Asp Ser Thr Pro Pro Pro Gly Thr Arg Val Arg Ala Met
 145 150 155 160
 Ala Ile Tyr Lys Gln Ser Gln His Met Thr Glu Val Val Arg Arg Cys

				165					170				175				
Pro	His	His	Glu	Arg	Cys	Ser	Asp	Ser	Asp	Gly	Leu	Ala	Pro	Pro	Gln		
			180					185					190				
His	Leu	Ile	Arg	Val	Glu	Gly	Asn	Leu	Arg	Val	Glu	Tyr	Leu	Asp	Asp		
		195					200					205					
Arg	Asn	Thr	Phe	Arg	His	Ser	Val	Val	Val	Pro	Tyr	Glu	Pro	Pro	Glu		
	210					215					220						
Val	Gly	Ser	Asp	Cys	Thr	Thr	Ile	His	Tyr	Asn	Tyr	Met	Cys	Asn	Ser		
225					230					235					240		
Ser	Cys	Met	Gly	Gly	Met	Asn	Arg	Arg	Pro	Ile	Leu	Thr	Ile	Ile	Thr		
			245						250					255			
Leu	Glu	Asp	Ser	Ser	Gly	Asn	Leu	Leu	Gly	Arg	Asn	Ser	Phe	Glu	Val		
			260					265					270				
Arg	Val	Cys	Ala	Cys	Pro	Gly	Arg	Asp	Arg	Arg	Thr	Glu	Glu	Glu	Asn		
		275				280						285					
Leu	Arg	Lys	Lys	Gly	Glu	Pro	His	His	Glu	Leu	Pro	Pro	Gly	Ser	Thr		
	290					295					300						
Lys	Arg	Ala	Leu	Pro	Asn	Asn	Thr	Ser	Ser	Ser	Pro	Gln	Pro	Lys	Lys		
305					310					315					320		
Lys	Pro	Leu	Asp	Gly	Glu	Tyr	Phe	Thr	Leu	Gln	Ile	Arg	Gly	Arg	Glu		
			325						330					335			
Arg	Phe	Glu	Met	Phe	Arg	Glu	Leu	Asn	Glu	Ala	Leu	Glu	Leu	Lys	Asp		
			340					345					350				
Ala	Gln	Ala	Gly	Lys	Glu	Pro	Gly	Gly	Ser	Arg	Ala	His	Ser	Ser	His		
		355					360				365						
Leu	Lys	Ser	Lys	Lys	Gly	Gln	Ser	Thr	Ser	Arg	His	Lys	Lys	Leu	Met		
	370					375					380						
Phe	Lys	Thr	Glu	Gly	Pro	Asp	Ser	Asp									
385					390												

<210> 211
 <211> 3859
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >TRRAP|ENSG00000196367|ENST00000359863|11580

<400> 211																	
Met	Ala	Phe	Val	Ala	Thr	Gln	Gly	Ala	Thr	Val	Val	Asp	Gln	Thr	Thr		
1				5				10					15				
Leu	Met	Lys	Lys	Tyr	Leu	Gln	Phe	Val	Ala	Ala	Leu	Thr	Asp	Val	Asn		
		20						25					30				
Thr	Pro	Asp	Glu	Thr	Lys	Leu	Lys	Met	Met	Gln	Glu	Val	Ser	Glu	Asn		
		35					40					45					
Phe	Glu	Asn	Val	Thr	Ser	Ser	Pro	Gln	Tyr	Ser	Thr	Phe	Leu	Glu	His		
	50				55						60						
Ile	Ile	Pro	Arg	Phe	Leu	Thr	Phe	Leu	Gln	Asp	Gly	Glu	Val	Gln	Phe		
65					70					75					80		
Leu	Gln	Glu	Lys	Pro	Ala	Gln	Gln	Leu	Arg	Lys	Leu	Val	Leu	Glu	Ile		
			85					90					95				
Ile	His	Arg	Ile	Pro	Thr	Asn	Glu	His	Leu	Arg	Pro	His	Thr	Lys	Asn		
			100					105					110				

Val	Leu	Ser	Val	Met	Phe	Arg	Phe	Leu	Glu	Thr	Glu	Asn	Glu	Glu	Asn	
		115					120					125				
Val	Leu	Ile	Cys	Leu	Arg	Ile	Ile	Ile	Glu	Leu	His	Lys	Gln	Phe	Arg	
		130				135					140					
Pro	Pro	Ile	Thr	Gln	Glu	Ile	His	His	Phe	Leu	Asp	Phe	Val	Lys	Gln	
145					150					155					160	
Ile	Tyr	Lys	Glu	Leu	Pro	Lys	Val	Val	Asn	Arg	Tyr	Phe	Glu	Asn	Pro	
				165					170					175		
Gln	Val	Ile	Pro	Glu	Asn	Thr	Val	Pro	Pro	Glu	Met	Val	Gly	Met		
			180					185				190				
Ile	Thr	Thr	Ile	Ala	Val	Lys	Val	Asn	Pro	Glu	Arg	Glu	Asp	Ser	Glu	
		195					200					205				
Thr	Arg	Thr	His	Ser	Ile	Ile	Pro	Arg	Gly	Ser	Leu	Ser	Leu	Lys	Val	
		210				215					220					
Leu	Ala	Glu	Leu	Pro	Ile	Ile	Val	Val	Leu	Met	Tyr	Gln	Leu	Tyr	Lys	
225					230				235						240	
Leu	Asn	Ile	His	Asn	Val	Val	Ala	Glu	Phe	Val	Pro	Leu	Ile	Met	Asn	
				245					250					255		
Thr	Ile	Ala	Ile	Gln	Val	Ser	Ala	Gln	Ala	Arg	Gln	His	Lys	Leu	Tyr	
			260					265					270			
Asn	Lys	Glu	Leu	Tyr	Ala	Asp	Phe	Ile	Ala	Ala	Gln	Ile	Lys	Thr	Leu	
		275					280					285				
Ser	Phe	Leu	Ala	Tyr	Ile	Ile	Arg	Ile	Tyr	Gln	Glu	Leu	Val	Thr	Lys	
	290					295					300					
Tyr	Ser	Gln	Gln	Met	Val	Lys	Gly	Met	Leu	Gln	Leu	Leu	Ser	Asn	Cys	
305				310					315						320	
Pro	Ala	Glu	Thr	Ala	His	Leu	Arg	Lys	Glu	Leu	Leu	Ile	Ala	Ala	Lys	
				325					330					335		
His	Ile	Leu	Thr	Thr	Glu	Leu	Arg	Asn	Gln	Phe	Ile	Pro	Cys	Met	Asp	
			340					345					350			
Lys	Leu	Phe	Asp	Glu	Ser	Ile	Leu	Ile	Gly	Ser	Gly	Tyr	Thr	Ala	Arg	
		355					360					365				
Glu	Thr	Leu	Arg	Pro	Leu	Ala	Tyr	Ser	Thr	Leu	Ala	Asp	Leu	Val	His	
		370				375					380					
His	Val	Arg	Gln	His	Leu	Pro	Leu	Ser	Asp	Leu	Ser	Leu	Ala	Val	Gln	
385					390					395					400	
Leu	Phe	Ala	Lys	Asn	Ile	Asp	Asp	Glu	Ser	Leu	Pro	Ser	Ser	Ile	Gln	
				405					410					415		
Thr	Met	Ser	Cys	Lys	Leu	Leu	Leu	Asn	Leu	Val	Asp	Cys	Ile	Arg	Ser	
			420					425					430			
Lys	Ser	Glu	Gln	Glu	Ser	Gly	Asn	Gly	Arg	Asp	Val	Leu	Met	Arg	Met	
		435					440					445				
Leu	Glu	Val	Phe	Val	Leu	Lys	Phe	His	Thr	Ile	Ala	Arg	Tyr	Gln	Leu	
						455					460					
Ser	Ala	Ile	Phe	Lys	Lys	Cys	Lys	Pro	Gln	Ser	Glu	Leu	Gly	Ala	Val	
465					470					475					480	
Glu	Ala	Ala	Leu	Pro	Gly	Val	Pro	Thr	Ala	Pro	Ala	Ala	Pro	Gly	Pro	
				485					490					495		
Ala	Pro	Ser	Pro	Ala	Pro	Val	Pro	Ala	Pro	Pro	Pro	Pro	Pro	Pro	Pro	
			500					505					510			
Pro	Pro	Pro	Ala	Thr	Pro	Val	Thr	Pro	Ala	Pro	Val	Pro	Pro	Phe	Glu	
		515					520					525				
Lys	Gln	Gly	Glu	Lys	Asp	Lys	Glu	Asp	Lys	Gln	Thr	Phe	Gln	Val	Thr	
		530				535					540					

Asp	Cys	Arg	Ser	Leu	Val	Lys	Thr	Leu	Val	Cys	Gly	Val	Lys	Thr	Ile
545					550					555					560
Thr	Trp	Gly	Ile	Thr	Ser	Cys	Lys	Ala	Pro	Gly	Glu	Ala	Gln	Phe	Ile
				565					570						575
Pro	Asn	Lys	Gln	Leu	Gln	Pro	Lys	Glu	Thr	Gln	Ile	Tyr	Ile	Lys	Leu
			580					585							590
Val	Lys	Tyr	Ala	Met	Gln	Ala	Leu	Asp	Ile	Tyr	Gln	Val	Gln	Ile	Ala
		595					600					605			
Gly	Asn	Gly	Gln	Thr	Tyr	Ile	Arg	Val	Ala	Asn	Cys	Gln	Thr	Val	Arg
610						615					620				
Met	Lys	Glu	Glu	Lys	Glu	Val	Leu	Glu	His	Phe	Ala	Gly	Val	Phe	Thr
625					630					635					640
Met	Met	Asn	Pro	Leu	Thr	Phe	Lys	Glu	Ile	Phe	Gln	Thr	Thr	Val	Pro
				645					650						655
Tyr	Met	Val	Glu	Arg	Ile	Ser	Lys	Asn	Tyr	Ala	Leu	Gln	Ile	Val	Ala
			660					665							670
Asn	Ser	Phe	Leu	Ala	Asn	Pro	Thr	Thr	Ser	Ala	Leu	Phe	Ala	Thr	Ile
		675					680					685			
Leu	Val	Glu	Tyr	Leu	Leu	Asp	Arg	Leu	Pro	Glu	Met	Gly	Ser	Asn	Val
	690					695					700				
Glu	Leu	Ser	Asn	Leu	Tyr	Leu	Lys	Leu	Phe	Lys	Leu	Val	Phe	Gly	Ser
705					710					715					720
Val	Ser	Leu	Phe	Ala	Ala	Glu	Asn	Glu	Gln	Met	Leu	Lys	Pro	His	Leu
				725					730						735
His	Lys	Ile	Val	Asn	Ser	Ser	Met	Glu	Leu	Ala	Gln	Thr	Ala	Lys	Glu
			740					745							750
Pro	Tyr	Asn	Tyr	Phe	Leu	Leu	Leu	Arg	Ala	Leu	Phe	Arg	Ser	Ile	Gly
		755					760					765			
Gly	Gly	Ser	His	Asp	Leu	Leu	Tyr	Gln	Glu	Phe	Leu	Pro	Leu	Leu	Pro
	770					775					780				
Asn	Leu	Leu	Gln	Gly	Leu	Asn	Met	Leu	Gln	Ser	Gly	Leu	His	Lys	Gln
785					790					795					800
His	Met	Lys	Asp	Leu	Phe	Val	Glu	Leu	Cys	Leu	Thr	Val	Pro	Val	Arg
			805						810						815
Leu	Ser	Ser	Leu	Leu	Pro	Tyr	Leu	Pro	Met	Leu	Met	Asp	Pro	Leu	Val
			820					825					830		
Ser	Ala	Leu	Asn	Gly	Ser	Gln	Thr	Leu	Val	Ser	Gln	Gly	Leu	Arg	Thr
		835					840					845			
Leu	Glu	Leu	Cys	Val	Asp	Asn	Leu	Gln	Pro	Asp	Phe	Leu	Tyr	Asp	His
	850					855					860				
Ile	Gln	Pro	Val	Arg	Ala	Glu	Leu	Met	Gln	Ala	Leu	Trp	Arg	Thr	Leu
865					870					875					880
Arg	Asn	Pro	Ala	Asp	Ser	Ile	Ser	His	Val	Ala	Tyr	Arg	Val	Leu	Gly
				885					890						895
Lys	Phe	Gly	Gly	Ser	Asn	Arg	Lys	Met	Leu	Lys	Glu	Ser	Gln	Lys	Leu
			900					905					910		
His	Tyr	Val	Val	Thr	Glu	Val	Gln	Gly	Pro	Ser	Ile	Thr	Val	Glu	Phe
		915					920					925			
Ser	Asp	Cys	Lys	Ala	Ser	Leu	Gln	Leu	Pro	Met	Glu	Lys	Ala	Ile	Glu
	930					935					940				
Thr	Ala	Leu	Asp	Cys	Leu	Lys	Ser	Ala	Asn	Thr	Glu	Pro	Tyr	Tyr	Arg
945					950					955					960
Arg	Gln	Ala	Trp	Glu	Val	Ile	Lys	Cys	Phe	Leu	Val	Ala	Met	Met	Ser
				965					970						975

Leu	Glu	Asp	Asn	Lys	His	Ala	Leu	Tyr	Gln	Leu	Leu	Ala	His	Pro	Asn	
			980					985					990			
Phe	Thr	Glu	Lys	Thr	Ile	Pro	Asn	Val	Ile	Ile	Ser	His	Arg	Tyr	Lys	
		995					1000					1005				
Ala	Gln	Asp	Thr	Pro	Ala	Arg	Lys	Thr	Phe	Glu	Gln	Ala	Leu	Thr	Gly	
	1010					1015					1020					
Ala	Phe	Met	Ser	Ala	Val	Ile	Lys	Asp	Leu	Arg	Pro	Ser	Ala	Leu	Pro	
1025					1030					1035					1040	
Phe	Val	Ala	Ser	Leu	Ile	Arg	His	Tyr	Thr	Met	Val	Ala	Val	Ala	Gln	
			1045						1050						1055	
Gln	Cys	Gly	Pro	Phe	Leu	Leu	Pro	Cys	Tyr	Gln	Val	Gly	Ser	Gln	Pro	
		1060						1065					1070			
Ser	Thr	Ala	Met	Phe	His	Ser	Glu	Glu	Asn	Gly	Ser	Lys	Gly	Met	Asp	
		1075					1080					1085				
Pro	Leu	Val	Leu	Ile	Asp	Ala	Ile	Ala	Ile	Cys	Met	Ala	Tyr	Glu	Glu	
	1090					1095					1100					
Lys	Glu	Leu	Cys	Lys	Ile	Gly	Glu	Val	Ala	Leu	Ala	Val	Ile	Phe	Asp	
1105					1110					1115					1120	
Val	Ala	Ser	Ile	Ile	Leu	Gly	Ser	Lys	Glu	Arg	Ala	Cys	Gln	Leu	Pro	
			1125						1130						1135	
Leu	Phe	Ser	Tyr	Ile	Val	Glu	Arg	Leu	Cys	Ala	Cys	Cys	Tyr	Glu	Gln	
		1140					1145						1150			
Ala	Trp	Tyr	Ala	Lys	Leu	Gly	Gly	Val	Val	Ser	Ile	Lys	Phe	Leu	Met	
	1155					1160						1165				
Glu	Arg	Leu	Pro	Leu	Thr	Trp	Val	Leu	Gln	Asn	Gln	Gln	Thr	Phe	Leu	
	1170					1175				1180						
Lys	Ala	Leu	Leu	Phe	Val	Met	Met	Asp	Leu	Thr	Gly	Glu	Val	Ser	Asn	
1185					1190					1195					1200	
Gly	Ala	Val	Ala	Met	Ala	Lys	Thr	Thr	Leu	Glu	Gln	Leu	Leu	Met	Arg	
			1205						1210					1215		
Cys	Ala	Thr	Pro	Leu	Lys	Asp	Glu	Glu	Arg	Ala	Glu	Glu	Ile	Val	Ala	
		1220					1225						1230			
Ala	Gln	Glu	Lys	Ser	Phe	His	His	Val	Thr	His	Asp	Leu	Val	Arg	Glu	
	1235					1240					1245					
Val	Thr	Ser	Pro	Asn	Ser	Thr	Val	Arg	Lys	Gln	Ala	Met	His	Ser	Leu	
	1250					1255					1260					
Gln	Val	Leu	Ala	Gln	Val	Thr	Gly	Lys	Ser	Val	Thr	Val	Ile	Met	Glu	
1265				1270						1275					1280	
Pro	His	Lys	Glu	Val	Leu	Gln	Asp	Met	Val	Pro	Pro	Lys	Lys	His	Leu	
			1285					1290						1295		
Leu	Arg	His	Gln	Pro	Ala	Asn	Ala	Gln	Ile	Gly	Leu	Met	Glu	Gly	Asn	
		1300					1305					1310				
Thr	Phe	Cys	Thr	Thr	Leu	Gln	Pro	Arg	Leu	Phe	Thr	Met	Asp	Leu	Asn	
	1315					1320						1325				
Val	Val	Glu	His	Lys	Val	Phe	Tyr	Thr	Glu	Leu	Leu	Asn	Leu	Cys	Glu	
	1330				1335						1340					
Ala	Glu	Asp	Ser	Ala	Leu	Thr	Lys	Leu	Pro	Cys	Tyr	Lys	Ser	Leu	Pro	
1345					1350				1355						1360	
Ser	Leu	Val	Pro	Leu	Arg	Ile	Ala	Ala	Leu	Asn	Ala	Leu	Ala	Ala	Cys	
			1365					1370					1375			
Asn	Tyr	Leu	Pro	Gln	Ser	Arg	Glu	Lys	Ile	Ile	Ala	Ala	Leu	Phe	Lys	
		1380					1385					1390				
Ala	Leu	Asn	Ser	Thr	Asn	Ser	Glu	Leu	Gln	Glu	Ala	Gly	Glu	Ala	Cys	
	1395						1400					1405				

Met	Arg	Lys	Phe	Leu	Glu	Gly	Ala	Thr	Ile	Glu	Val	Asp	Gln	Ile	His	
1410						1415						1420				
Thr	His	Met	Arg	Pro	Leu	Leu	Met	Met	Leu	Gly	Asp	Tyr	Arg	Ser	Leu	
1425						1430						1435				
Thr	Leu	Asn	Val	Val	Asn	Arg	Leu	Thr	Ser	Val	Thr	Arg	Leu	Phe	Pro	
					1445						1450					
Asn	Ser	Phe	Asn	Asp	Lys	Phe	Cys	Asp	Gln	Met	Met	Gln	His	Leu	Arg	
					1460						1465					
Lys	Trp	Met	Glu	Val	Val	Val	Ile	Thr	His	Lys	Gly	Gly	Gln	Arg	Ser	
					1475						1480					
Asp	Gly	Asn	Glu	Ser	Ile	Ser	Glu	Cys	Gly	Arg	Cys	Pro	Leu	Ser	Pro	
					1490						1495					
Phe	Cys	Gln	Phe	Glu	Glu	Met	Lys	Ile	Cys	Ser	Ala	Ile	Ile	Asn	Leu	
1505						1510						1515				
Phe	His	Leu	Ile	Pro	Ala	Ala	Pro	Gln	Thr	Leu	Val	Lys	Pro	Leu	Leu	
					1525						1530					
Glu	Val	Val	Met	Lys	Thr	Glu	Arg	Ala	Met	Leu	Ile	Glu	Ala	Gly	Ser	
					1540						1545					
Pro	Phe	Arg	Glu	Pro	Leu	Ile	Lys	Phe	Leu	Thr	Arg	His	Pro	Ser	Gln	
					1555						1560					
Thr	Val	Glu	Leu	Phe	Met	Met	Glu	Ala	Thr	Leu	Asn	Asp	Pro	Gln	Trp	
1570						1575						1580				
Ser	Arg	Met	Phe	Met	Ser	Phe	Leu	Lys	His	Lys	Asp	Ala	Arg	Pro	Leu	
1585						1590						1595				
Arg	Asp	Val	Leu	Ala	Ala	Asn	Pro	Asn	Arg	Phe	Ile	Thr	Leu	Leu	Leu	
					1605						1610					
Pro	Gly	Gly	Ala	Gln	Thr	Ala	Val	Arg	Pro	Gly	Ser	Pro	Ser	Thr	Ser	
					1620						1625					
Thr	Met	Arg	Leu	Asp	Leu	Gln	Phe	Gln	Ala	Ile	Lys	Ile	Ile	Ser	Ile	
					1635						1640					
Ile	Val	Lys	Asn	Asp	Asp	Ser	Trp	Leu	Ala	Ser	Gln	His	Ser	Leu	Val	
1650						1655						1660				
Ser	Gln	Leu	Arg	Arg	Val	Trp	Val	Ser	Glu	Asn	Phe	Gln	Glu	Arg	His	
1665						1670						1675				
Arg	Lys	Glu	Asn	Met	Ala	Ala	Thr	Asn	Trp	Lys	Glu	Pro	Lys	Leu	Leu	
					1685						1690					
Ala	Tyr	Cys	Leu	Leu	Asn	Tyr	Cys	Lys	Arg	Asn	Tyr	Gly	Asp	Ile	Glu	
					1700						1705					
Leu	Leu	Phe	Gln	Leu	Leu	Arg	Ala	Phe	Thr	Gly	Arg	Phe	Leu	Cys	Asn	
					1715						1720					
Met	Thr	Phe	Leu	Lys	Glu	Tyr	Met	Glu	Glu	Glu	Ile	Pro	Lys	Asn	Tyr	
1730						1735						1740				
Ser	Ile	Ala	Gln	Lys	Arg	Ala	Leu	Phe	Phe	Arg	Phe	Val	Asp	Phe	Asn	
1745						1750						1755				
Asp	Pro	Asn	Phe	Gly	Asp	Glu	Leu	Lys	Ala	Lys	Val	Leu	Gln	His	Ile	
					1765						1770					
Leu	Asn	Pro	Ala	Phe	Leu	Tyr	Ser	Phe	Glu	Lys	Gly	Glu	Gly	Glu	Gln	
					1780						1785					
Leu	Leu	Gly	Pro	Pro	Asn	Pro	Glu	Gly	Asp	Asn	Pro	Glu	Ser	Ile	Thr	
					1795						1800					
Ser	Val	Phe	Ile	Thr	Lys	Val	Leu	Asp	Pro	Glu	Lys	Gln	Ala	Asp	Met	
					1810						1815					
Leu	Asp	Ser	Leu	Arg	Ile	Tyr	Leu	Leu	Gln	Tyr	Ala	Thr	Leu	Leu	Val	
1825						1830						1835				

Page 611

Met Ile Leu Lys Ser Ala Cys Ser Asn Asn Pro Ser Tyr Ile Asp Arg
 2275 2280 2285
 Leu Ile Ser Val Phe Met Arg Ser Leu Gln Lys Met Val Arg Glu His
 2290 2295 2300
 Leu Asn Pro Gln Ala Ala Ser Gly Ser Thr Glu Ala Thr Ser Gly Thr
 2305 2310 2315 2320
 Ser Glu Leu Val Met Leu Ser Leu Glu Leu Val Lys Thr Arg Leu Ala
 2325 2330 2335
 Val Met Ser Met Glu Met Arg Lys Asn Phe Ile Gln Ala Ile Leu Thr
 2340 2345 2350
 Ser Leu Ile Glu Lys Ser Pro Asp Ala Lys Ile Leu Arg Ala Val Val
 2355 2360 2365
 Lys Ile Val Glu Glu Trp Val Lys Asn Asn Ser Pro Met Ala Ala Asn
 2370 2375 2380
 Gln Thr Pro Thr Leu Arg Glu Lys Ser Ile Leu Leu Val Lys Met Met
 2385 2390 2395 2400
 Thr Tyr Ile Glu Lys Arg Phe Pro Glu Asp Leu Glu Leu Asn Ala Gln
 2405 2410 2415
 Phe Leu Asp Leu Val Asn Tyr Val Tyr Arg Asp Glu Thr Leu Ser Gly
 2420 2425 2430
 Ser Glu Leu Thr Ala Lys Leu Glu Pro Ala Phe Leu Ser Gly Leu Arg
 2435 2440 2445
 Cys Ala Gln Pro Leu Ile Arg Ala Lys Phe Phe Glu Val Phe Asp Asn
 2450 2455 2460
 Ser Met Lys Arg Arg Val Tyr Glu Arg Leu Leu Tyr Val Thr Cys Ser
 2465 2470 2475 2480
 Gln Asn Trp Glu Ala Met Gly Asn His Phe Trp Ile Lys Gln Cys Ile
 2485 2490 2495
 Glu Leu Leu Leu Ala Val Cys Glu Lys Ser Thr Pro Ile Gly Thr Ser
 2500 2505 2510
 Cys Gln Gly Ala Met Leu Pro Ser Ile Thr Asn Val Ile Asn Leu Ala
 2515 2520 2525
 Asp Ser His Asp Arg Ala Ala Phe Ala Met Val Thr His Val Lys Gln
 2530 2535 2540
 Glu Pro Arg Glu Arg Glu Asn Ser Glu Ser Lys Glu Glu Asp Val Glu
 2545 2550 2555 2560
 Ile Asp Ile Glu Leu Ala Pro Gly Asp Gln Thr Ser Thr Pro Lys Thr
 2565 2570 2575
 Lys Glu Leu Ser Glu Lys Asp Ile Gly Asn Gln Leu His Met Leu Thr
 2580 2585 2590
 Asn Arg His Asp Lys Phe Leu Asp Thr Leu Arg Glu Val Lys Thr Gly
 2595 2600 2605
 Ala Leu Leu Ser Ala Phe Val Gln Leu Cys His Ile Ser Thr Thr Leu
 2610 2615 2620
 Ala Glu Lys Thr Trp Val Gln Leu Phe Pro Arg Leu Trp Lys Ile Leu
 2625 2630 2635 2640
 Ser Asp Arg Gln Gln His Ala Leu Ala Gly Glu Ile Ser Pro Phe Leu
 2645 2650 2655
 Cys Ser Gly Ser His Gln Val Gln Arg Asp Cys Gln Pro Ser Ala Leu
 2660 2665 2670
 Asn Cys Phe Val Glu Ala Met Ser Gln Cys Val Pro Pro Ile Pro Ile
 2675 2680 2685
 Arg Pro Cys Val Leu Lys Tyr Leu Gly Lys Thr His Asn Leu Trp Phe
 2690 2695 2700

Arg Ser Thr Leu Met Leu Glu His Gln Ala Phe Glu Lys Gly Leu Ser
 2705 2710 2715 2720
 Leu Gln Ile Lys Pro Lys Gln Thr Thr Glu Phe Tyr Glu Gln Glu Ser
 2725 2730 2735
 Ile Thr Pro Pro Gln Gln Glu Ile Leu Asp Ser Leu Ala Glu Leu Tyr
 2740 2745 2750
 Ser Leu Leu Gln Glu Glu Asp Met Trp Ala Gly Leu Trp Gln Lys Arg
 2755 2760 2765
 Cys Lys Tyr Ser Glu Thr Ala Thr Ala Ile Ala Tyr Glu Gln His Gly
 2770 2775 2780
 Phe Phe Glu Gln Ala Gln Glu Ser Tyr Glu Lys Ala Met Asp Lys Ala
 2785 2790 2795 2800
 Lys Lys Glu His Glu Arg Ser Asn Ala Ser Pro Ala Ile Phe Pro Glu
 2805 2810 2815
 Tyr Gln Leu Trp Glu Asp His Trp Ile Arg Cys Ser Lys Glu Leu Asn
 2820 2825 2830
 Gln Trp Glu Ala Leu Thr Glu Tyr Gly Gln Ser Lys Gly His Ile Asn
 2835 2840 2845
 Pro Tyr Leu Val Leu Glu Cys Ala Trp Arg Val Ser Asn Trp Thr Ala
 2850 2855 2860
 Met Lys Glu Ala Leu Val Gln Val Glu Val Ser Cys Pro Lys Glu Met
 2865 2870 2875 2880
 Ala Trp Lys Val Asn Met Tyr Arg Gly Tyr Leu Ala Ile Cys His Pro
 2885 2890 2895
 Glu Glu Gln Gln Leu Ser Phe Ile Glu Arg Leu Val Glu Met Ala Ser
 2900 2905 2910
 Ser Leu Ala Ile Arg Glu Trp Arg Arg Leu Pro His Val Val Ser His
 2915 2920 2925
 Val His Thr Pro Leu Leu Gln Ala Ala Gln Gln Ile Ile Glu Leu Gln
 2930 2935 2940
 Glu Ala Ala Gln Ile Asn Ala Gly Leu Gln Pro Thr Asn Leu Gly Arg
 2945 2950 2955 2960
 Asn Asn Ser Leu His Asp Met Lys Thr Val Val Lys Thr Trp Arg Asn
 2965 2970 2975
 Arg Leu Pro Ile Val Ser Asp Asp Leu Ser His Trp Ser Ser Ile Phe
 2980 2985 2990
 Met Trp Arg Gln His His Tyr Gln Gly Lys Pro Thr Trp Ser Gly Met
 2995 3000 3005
 His Ser Ser Ser Ile Val Thr Ala Tyr Glu Asn Ser Ser Gln His Asp
 3010 3015 3020
 Pro Ser Ser Asn Asn Ala Met Leu Gly Val His Ala Ser Ala Ser Ala
 3025 3030 3035 3040
 Ile Ile Gln Tyr Gly Lys Ile Ala Arg Lys Gln Gly Leu Val Asn Val
 3045 3050 3055
 Ala Leu Asp Ile Leu Ser Arg Ile His Thr Ile Pro Thr Val Pro Ile
 3060 3065 3070
 Val Asp Cys Phe Gln Lys Ile Arg Gln Gln Val Lys Cys Tyr Leu Gln
 3075 3080 3085
 Leu Ala Gly Val Met Gly Lys Asn Glu Cys Met Gln Gly Leu Glu Val
 3090 3095 3100
 Ile Glu Ser Thr Asn Leu Lys Tyr Phe Thr Lys Glu Met Thr Ala Glu
 3105 3110 3115 3120
 Phe Tyr Ala Leu Lys Gly Met Phe Leu Ala Gln Ile Asn Lys Ser Glu
 3125 3130 3135

Glu Ala Asn Lys Ala Phe Ser Ala Ala Val Gln Met His Asp Val Leu
 3140 3145 3150
 Val Lys Ala Trp Ala Met Trp Gly Asp Tyr Leu Glu Asn Ile Phe Val
 3155 3160 3165
 Lys Glu Arg Gln Leu His Leu Gly Val Ser Ala Ile Thr Cys Tyr Leu
 3170 3175 3180
 His Ala Cys Arg His Gln Asn Glu Ser Lys Ser Arg Lys Tyr Leu Ala
 3185 3190 3195 3200
 Lys Val Leu Trp Leu Leu Ser Phe Asp Asp Asp Lys Asn Thr Leu Ala
 3205 3210 3215
 Asp Ala Val Asp Lys Tyr Cys Ile Gly Val Pro Pro Ile Gln Trp Leu
 3220 3225 3230
 Ala Trp Ile Pro Gln Leu Leu Thr Cys Leu Val Gly Ser Glu Gly Lys
 3235 3240 3245
 Leu Leu Leu Asn Leu Ile Ser Gln Val Gly Arg Val Tyr Pro Gln Ala
 3250 3255 3260
 Val Tyr Phe Pro Ile Arg Thr Leu Tyr Leu Thr Leu Lys Ile Glu Gln
 3265 3270 3275 3280
 Arg Glu Arg Tyr Lys Ser Asp Pro Gly Pro Ile Arg Ala Thr Ala Pro
 3285 3290 3295
 Met Trp Arg Cys Ser Arg Ile Met His Met Gln Arg Glu Leu His Pro
 3300 3305 3310
 Thr Leu Leu Ser Ser Leu Glu Gly Ile Val Asp Gln Met Val Trp Phe
 3315 3320 3325
 Arg Glu Asn Trp His Glu Glu Val Leu Arg Gln Leu Gln Gln Gly Leu
 3330 3335 3340
 Ala Lys Cys Tyr Ser Val Ala Phe Glu Lys Ser Gly Ala Val Ser Asp
 3345 3350 3355 3360
 Ala Lys Ile Thr Pro His Thr Leu Asn Phe Val Lys Lys Leu Val Ser
 3365 3370 3375
 Thr Phe Gly Val Gly Leu Glu Asn Val Ser Asn Val Ser Thr Met Phe
 3380 3385 3390
 Ser Ser Ala Ala Ser Glu Ser Leu Ala Arg Arg Ala Gln Ala Thr Ala
 3395 3400 3405
 Gln Asp Pro Val Phe Gln Lys Leu Lys Gly Gln Phe Thr Thr Asp Phe
 3410 3415 3420
 Asp Phe Ser Val Pro Gly Ser Met Lys Leu His Asn Leu Ile Ser Lys
 3425 3430 3435 3440
 Leu Lys Lys Trp Ile Lys Ile Leu Glu Ala Lys Thr Lys Gln Leu Pro
 3445 3450 3455
 Lys Phe Phe Leu Ile Glu Glu Lys Cys Arg Phe Leu Ser Asn Phe Ser
 3460 3465 3470
 Ala Gln Thr Ala Glu Val Glu Ile Pro Gly Glu Phe Leu Met Pro Lys
 3475 3480 3485
 Pro Thr His Tyr Tyr Ile Lys Ile Ala Arg Phe Met Pro Arg Val Glu
 3490 3495 3500
 Ile Val Gln Lys His Asn Thr Ala Ala Arg Arg Leu Tyr Ile Arg Gly
 3505 3510 3515 3520
 His Asn Gly Lys Ile Tyr Pro Tyr Leu Val Met Asn Asp Ala Cys Leu
 3525 3530 3535
 Thr Glu Ser Arg Arg Glu Glu Arg Val Leu Gln Leu Leu Arg Leu Leu
 3540 3545 3550
 Asn Pro Cys Leu Glu Lys Arg Lys Glu Thr Thr Lys Arg His Leu Phe
 3555 3560 3565

Phe Thr Val Pro Arg Val Val Ala Val Ser Pro Gln Met Arg Leu Val
 3570 3575 3580
 Glu Asp Asn Pro Ser Ser Leu Ser Leu Val Glu Ile Tyr Lys Gln Arg
 3585 3590 3595 3600
 Cys Ala Lys Lys Gly Ile Glu His Asp Asn Pro Ile Ser Arg Tyr Tyr
 3605 3610 3615
 Asp Arg Leu Ala Thr Val Gln Ala Arg Gly Thr Gln Ala Ser His Gln
 3620 3625 3630
 Val Leu Arg Asp Ile Leu Lys Glu Val Gln Ser Asn Met Val Pro Arg
 3635 3640 3645
 Ser Met Leu Lys Glu Trp Ala Leu His Thr Phe Pro Asn Ala Thr Asp
 3650 3655 3660
 Tyr Trp Thr Phe Arg Lys Met Phe Thr Ile Gln Leu Ala Leu Ile Gly
 3665 3670 3675 3680
 Phe Ala Glu Phe Val Leu His Leu Asn Arg Leu Asn Pro Glu Met Leu
 3685 3690 3695
 Gln Ile Ala Gln Asp Thr Gly Lys Leu Asn Val Ala Tyr Phe Arg Phe
 3700 3705 3710
 Asp Ile Asn Asp Ala Thr Gly Asp Leu Asp Ala Asn Arg Pro Val Pro
 3715 3720 3725
 Phe Arg Leu Thr Pro Asn Ile Ser Glu Phe Leu Thr Thr Ile Gly Val
 3730 3735 3740
 Ser Gly Pro Leu Thr Ala Ser Met Ile Ala Val Ala Arg Cys Phe Ala
 3745 3750 3755 3760
 Gln Pro Asn Phe Lys Val Asp Gly Ile Leu Lys Thr Val Leu Arg Asp
 3765 3770 3775
 Glu Ile Ile Ala Trp His Lys Lys Thr Gln Glu Asp Thr Ser Ser Pro
 3780 3785 3790
 Leu Ser Ala Ala Gly Gln Pro Glu Asn Met Asp Ser Gln Gln Leu Val
 3795 3800 3805
 Ser Leu Val Gln Lys Ala Val Thr Ala Ile Met Thr Arg Leu His Asn
 3810 3815 3820
 Leu Ala Gln Phe Glu Gly Gly Glu Ser Lys Val Asn Thr Leu Val Ala
 3825 3830 3835 3840
 Ala Ala Asn Ser Leu Asp Asn Leu Cys Arg Met Asp Pro Ala Trp His
 3845 3850 3855
 Pro Trp Leu

<210> 212
 <211> 152
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >UBE2N|ENSG00000177889|ENST00000318066|459

<400> 212
 Met Ala Gly Leu Pro Arg Arg Ile Ile Lys Glu Thr Gln Arg Leu Leu
 1 5 10 15
 Ala Glu Pro Val Pro Gly Ile Lys Ala Glu Pro Asp Glu Ser Asn Ala
 20 25 30
 Arg Tyr Phe His Val Val Ile Ala Gly Pro Gln Asp Ser Pro Phe Glu

		35				40				45					
Gly	Gly	Thr	Phe	Lys	Leu	Glu	Leu	Phe	Leu	Pro	Glu	Glu	Tyr	Pro	Met
	50					55					60				
Ala	Ala	Pro	Lys	Val	Arg	Phe	Met	Thr	Lys	Ile	Tyr	His	Pro	Asn	Val
65					70					75					80
Asp	Lys	Leu	Gly	Arg	Ile	Cys	Leu	Asp	Ile	Leu	Lys	Asp	Lys	Trp	Ser
				85					90					95	
Pro	Ala	Leu	Gln	Ile	Arg	Thr	Val	Leu	Leu	Ser	Ile	Gln	Ala	Leu	Leu
			100					105					110		
Ser	Ala	Pro	Asn	Pro	Asp	Asp	Pro	Leu	Ala	Asn	Asp	Val	Ala	Glu	Gln
		115					120					125			
Trp	Lys	Thr	Asn	Glu	Ala	Gln	Ala	Ile	Glu	Thr	Ala	Arg	Ala	Trp	Thr
	130					135					140				
Arg	Leu	Tyr	Ala	Met	Asn	Asn	Ile								
145					150										

<210> 213

<211> 719

<212> PRT

<213> Homo sapiens

<220>

<223> >UIMC1|ENSG00000087206|ENST00000377227|2160

<400> 213

Met	Pro	Arg	Arg	Lys	Lys	Lys	Val	Lys	Glu	Val	Ser	Glu	Ser	Arg	Asn
1				5					10					15	
Leu	Glu	Lys	Lys	Asp	Val	Glu	Thr	Thr	Ser	Ser	Val	Ser	Val	Lys	Arg
			20					25					30		
Lys	Arg	Arg	Leu	Glu	Asp	Ala	Phe	Ile	Val	Ile	Ser	Asp	Ser	Asp	Gly
		35					40					45			
Glu	Glu	Pro	Lys	Glu	Glu	Asn	Gly	Leu	Gln	Lys	Thr	Lys	Thr	Lys	Gln
	50					55					60				
Ser	Asn	Arg	Ala	Lys	Cys	Leu	Ala	Lys	Arg	Lys	Ile	Ala	Gln	Met	Thr
65					70					75					80
Glu	Glu	Glu	Gln	Phe	Ala	Leu	Ala	Leu	Lys	Met	Ser	Glu	Gln	Glu	Ala
			85						90					95	
Arg	Glu	Val	Asn	Ser	Gln	Glu	Glu	Glu	Glu	Glu	Glu	Leu	Leu	Arg	Lys
			100					105					110		
Ala	Ile	Ala	Glu	Ser	Leu	Asn	Ser	Cys	Arg	Pro	Ser	Asp	Ala	Ser	Ala
		115					120					125			
Thr	Arg	Ser	Arg	Pro	Leu	Ala	Thr	Gly	Pro	Ser	Ser	Gln	Ser	His	Gln
	130					135						140			
Glu	Lys	Thr	Thr	Asp	Ser	Gly	Leu	Thr	Glu	Gly	Ile	Trp	Gln	Leu	Val
145					150					155					160
Pro	Pro	Ser	Leu	Phe	Lys	Gly	Ser	His	Ile	Ser	Gln	Gly	Asn	Glu	Ala
				165					170					175	
Glu	Glu	Arg	Glu	Glu	Pro	Trp	Asp	His	Thr	Glu	Lys	Thr	Glu	Glu	Glu
			180					185					190		
Pro	Val	Ser	Gly	Ser	Ser	Gly	Ser	Trp	Asp	Gln	Ser	Ser	Gln	Pro	Val
		195					200					205			
Phe	Glu	Asn	Val	Asn	Val	Lys	Ser	Phe	Asp	Arg	Cys	Thr	Gly	His	Ser
	210					215					220				

Ala	Glu	His	Thr	Gln	Cys	Gly	Lys	Pro	Gln	Glu	Ser	Thr	Gly	Arg	Gly	225	230	235	240
Ser	Ala	Phe	Leu	Lys	Ala	Val	Gln	Gly	Ser	Gly	Asp	Thr	Ser	Arg	His	245	250	255	
Cys	Leu	Pro	Thr	Leu	Ala	Asp	Ala	Lys	Gly	Leu	Gln	Asp	Thr	Gly	Gly	260	265	270	
Thr	Val	Asn	Tyr	Phe	Trp	Gly	Ile	Pro	Phe	Cys	Pro	Asp	Gly	Val	Asp	275	280	285	
Pro	Asn	Gln	Tyr	Thr	Lys	Val	Ile	Leu	Cys	Gln	Leu	Glu	Val	Tyr	Gln	290	295	300	
Lys	Ser	Leu	Lys	Met	Ala	Gln	Arg	Gln	Leu	Leu	Asn	Lys	Lys	Gly	Phe	305	310	315	320
Gly	Glu	Pro	Val	Leu	Pro	Arg	Pro	Pro	Ser	Leu	Ile	Gln	Asn	Glu	Cys	325	330	335	
Gly	Gln	Gly	Glu	Gln	Ala	Ser	Glu	Lys	Asn	Glu	Cys	Ile	Ser	Glu	Asp	340	345	350	
Met	Gly	Asp	Glu	Asp	Lys	Glu	Glu	Arg	Gln	Glu	Ser	Arg	Ala	Ser	Asp	355	360	365	
Trp	His	Ser	Lys	Thr	Lys	Asp	Phe	Gln	Glu	Ser	Ser	Ile	Lys	Ser	Leu	370	375	380	
Lys	Glu	Lys	Leu	Leu	Leu	Glu	Glu	Glu	Pro	Thr	Thr	Ser	His	Gly	Gln	385	390	395	400
Ser	Ser	Gln	Gly	Ile	Val	Glu	Glu	Thr	Ser	Glu	Glu	Gly	Asn	Ser	Val	405	410	415	
Pro	Ala	Ser	Gln	Ser	Val	Ala	Ala	Leu	Thr	Ser	Lys	Arg	Ser	Leu	Val	420	425	430	
Leu	Met	Pro	Glu	Ser	Ser	Ala	Glu	Glu	Ile	Thr	Val	Cys	Pro	Glu	Thr	435	440	445	
Gln	Leu	Ser	Ser	Ser	Glu	Thr	Phe	Asp	Leu	Glu	Arg	Glu	Val	Ser	Pro	450	455	460	
Gly	Ser	Arg	Asp	Ile	Leu	Asp	Gly	Val	Arg	Ile	Ile	Met	Ala	Asp	Lys	465	470	475	480
Glu	Val	Gly	Asn	Lys	Glu	Asp	Ala	Glu	Lys	Glu	Val	Ala	Ile	Ser	Thr	485	490	495	
Phe	Ser	Ser	Ser	Asn	Gln	Val	Ser	Cys	Pro	Leu	Cys	Asp	Gln	Cys	Phe	500	505	510	
Pro	Pro	Thr	Lys	Ile	Glu	Arg	His	Ala	Met	Tyr	Cys	Asn	Gly	Leu	Met	515	520	525	
Glu	Glu	Asp	Thr	Val	Leu	Thr	Arg	Arg	Gln	Lys	Glu	Ala	Lys	Thr	Lys	530	535	540	
Ser	Asp	Ser	Gly	Thr	Ala	Ala	Gln	Thr	Ser	Leu	Asp	Ile	Asp	Lys	Asn	545	550	555	560
Glu	Lys	Cys	Tyr	Leu	Cys	Lys	Ser	Leu	Val	Pro	Phe	Arg	Glu	Tyr	Gln	565	570	575	
Cys	His	Val	Asp	Ser	Cys	Leu	Gln	Leu	Ala	Lys	Ala	Asp	Gln	Gly	Asp	580	585	590	
Gly	Pro	Glu	Gly	Ser	Gly	Arg	Ala	Cys	Ser	Thr	Val	Glu	Gly	Lys	Trp	595	600	605	
Gln	Gln	Arg	Leu	Lys	Asn	Pro	Lys	Glu	Lys	Gly	His	Ser	Glu	Gly	Arg	610	615	620	
Leu	Leu	Ser	Phe	Leu	Glu	Gln	Ser	Glu	His	Lys	Thr	Ser	Asp	Ala	Asp	625	630	635	640
Ile	Lys	Ser	Ser	Glu	Thr	Gly	Ala	Phe	Arg	Val	Pro	Ser	Pro	Gly	Met	645	650	655	

Glu	Glu	Ala	Gly	Cys	Ser	Arg	Glu	Met	Gln	Ser	Ser	Phe	Thr	Arg	Arg
			660					665					670		
Asp	Leu	Asn	Glu	Ser	Pro	Val	Lys	Ser	Phe	Val	Ser	Ile	Ser	Glu	Ala
		675					680					685			
Thr	Asp	Cys	Leu	Val	Asp	Phe	Lys	Lys	Gln	Val	Thr	Val	Gln	Pro	Gly
	690					695					700				
Ser	Arg	Thr	Arg	Thr	Lys	Ala	Gly	Arg	Gly	Arg	Arg	Lys	Phe		
705					710					715					

<210> 214
 <211> 785
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >USP1|ENSG00000162607|ENST00000339950|2358

<400> 214

Met	Pro	Gly	Val	Ile	Pro	Ser	Glu	Ser	Asn	Gly	Leu	Ser	Arg	Gly	Ser
1			5					10					15		
Pro	Ser	Lys	Lys	Asn	Arg	Leu	Ser	Leu	Lys	Phe	Phe	Gln	Lys	Lys	Glu
		20					25					30			
Thr	Lys	Arg	Ala	Leu	Asp	Phe	Thr	Asp	Ser	Gln	Glu	Asn	Glu	Glu	Lys
	35					40					45				
Ala	Ser	Glu	Tyr	Arg	Ala	Ser	Glu	Ile	Asp	Gln	Val	Val	Pro	Ala	Ala
	50					55				60					
Gln	Ser	Ser	Pro	Ile	Asn	Cys	Glu	Lys	Arg	Glu	Asn	Leu	Leu	Pro	Phe
65				70					75					80	
Val	Gly	Leu	Asn	Asn	Leu	Gly	Asn	Thr	Cys	Tyr	Leu	Asn	Ser	Ile	Leu
			85					90						95	
Gln	Val	Leu	Tyr	Phe	Cys	Pro	Gly	Phe	Lys	Ser	Gly	Val	Lys	His	Leu
		100						105					110		
Phe	Asn	Ile	Ile	Ser	Arg	Lys	Lys	Glu	Ala	Leu	Lys	Asp	Glu	Ala	Asn
	115						120					125			
Gln	Lys	Asp	Lys	Gly	Asn	Cys	Lys	Glu	Asp	Ser	Leu	Ala	Ser	Tyr	Glu
	130					135					140				
Leu	Ile	Cys	Ser	Leu	Gln	Ser	Leu	Ile	Ile	Ser	Val	Glu	Gln	Leu	Gln
145					150					155					160
Ala	Ser	Phe	Leu	Leu	Asn	Pro	Glu	Lys	Tyr	Thr	Asp	Glu	Leu	Ala	Thr
			165						170					175	
Gln	Pro	Arg	Arg	Leu	Leu	Asn	Thr	Leu	Arg	Glu	Leu	Asn	Pro	Met	Tyr
		180						185					190		
Glu	Gly	Tyr	Leu	Gln	His	Asp	Ala	Gln	Glu	Val	Leu	Gln	Cys	Ile	Leu
	195						200					205			
Gly	Asn	Ile	Gln	Glu	Thr	Cys	Gln	Leu	Leu	Lys	Lys	Glu	Glu	Val	Lys
	210					215					220				
Asn	Val	Ala	Glu	Leu	Pro	Thr	Lys	Val	Glu	Glu	Ile	Pro	His	Pro	Lys
225					230					235					240
Glu	Glu	Met	Asn	Gly	Ile	Asn	Ser	Ile	Glu	Met	Asp	Ser	Met	Arg	His
			245						250					255	
Ser	Glu	Asp	Phe	Lys	Glu	Lys	Leu	Pro	Lys	Gly	Asn	Gly	Lys	Arg	Lys
		260						265					270		
Ser	Asp	Thr	Glu	Phe	Gly	Asn	Met	Lys	Lys	Lys	Val	Lys	Leu	Ser	Lys

705					710					715					720
Gln	Thr	Gly	Ile	Asn	Ile	Ser	Gly	Phe	Glu	Asn	Lys	Ile	Ser	Tyr	Val
				725					730					735	
Val	Gln	Ser	Leu	Lys	Glu	Tyr	Glu	Gly	Lys	Trp	Leu	Leu	Phe	Asp	Asp
			740					745					750		
Ser	Glu	Val	Lys	Val	Thr	Glu	Glu	Lys	Asp	Phe	Leu	Asn	Ser	Leu	Ser
		755					760				765				
Pro	Ser	Thr	Ser	Pro	Thr	Ser	Thr	Pro	Tyr	Leu	Leu	Phe	Tyr	Lys	Lys
	770					775					780				
Leu															
785															

<210> 215
 <211> 677
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >WDR48|ENSG00000114742|ENST00000302313|2034

<400> 215															
Met	Ala	Ala	His	His	Arg	Gln	Asn	Thr	Ala	Gly	Arg	Arg	Lys	Val	Gln
1			5						10					15	
Val	Ser	Tyr	Val	Ile	Arg	Asp	Glu	Val	Glu	Lys	Tyr	Asn	Arg	Asn	Gly
			20					25				30			
Val	Asn	Ala	Leu	Gln	Leu	Asp	Pro	Ala	Leu	Asn	Arg	Leu	Phe	Thr	Ala
		35				40					45				
Gly	Arg	Asp	Ser	Ile	Ile	Arg	Ile	Trp	Ser	Val	Asn	Gln	His	Lys	Gln
	50				55					60					
Asp	Pro	Tyr	Ile	Ala	Ser	Met	Glu	His	His	Thr	Asp	Trp	Val	Asn	Asp
65					70				75					80	
Ile	Val	Leu	Cys	Cys	Asn	Gly	Lys	Thr	Leu	Ile	Ser	Ala	Ser	Ser	Asp
			85						90					95	
Thr	Thr	Val	Lys	Val	Trp	Asn	Ala	His	Lys	Gly	Phe	Cys	Met	Ser	Thr
		100					105						110		
Leu	Arg	Thr	His	Lys	Asp	Tyr	Val	Lys	Ala	Leu	Ala	Tyr	Ala	Lys	Asp
	115					120					125				
Lys	Glu	Leu	Val	Ala	Ser	Ala	Gly	Leu	Asp	Arg	Gln	Ile	Phe	Leu	Trp
	130			135					140						
Asp	Val	Asn	Thr	Leu	Thr	Ala	Leu	Thr	Ala	Ser	Asn	Asn	Thr	Val	Thr
145				150					155					160	
Thr	Ser	Ser	Leu	Ser	Gly	Asn	Lys	Asp	Ser	Ile	Tyr	Ser	Leu	Ala	Met
			165				170						175		
Asn	Gln	Leu	Gly	Thr	Ile	Ile	Val	Ser	Gly	Ser	Thr	Glu	Lys	Val	Leu
	180						185					190			
Arg	Val	Trp	Asp	Pro	Arg	Thr	Cys	Ala	Lys	Leu	Met	Lys	Leu	Lys	Gly
	195					200					205				
His	Thr	Asp	Asn	Val	Lys	Ala	Leu	Leu	Leu	Asn	Arg	Asp	Gly	Thr	Gln
	210				215					220					
Cys	Leu	Ser	Gly	Ser	Ser	Asp	Gly	Thr	Ile	Arg	Leu	Trp	Ser	Leu	Gly
225				230					235					240	
Gln	Gln	Arg	Cys	Ile	Ala	Thr	Tyr	Arg	Val	His	Asp	Glu	Gly	Val	Trp
			245					250					255		

Ala	Leu	Gln	Val	Asn	Asp	Ala	Phe	Thr	His	Val	Tyr	Ser	Gly	Gly	Arg		
			260					265					270				
Asp	Arg	Lys	Ile	Tyr	Cys	Thr	Asp	Leu	Arg	Asn	Pro	Asp	Ile	Arg	Val		
		275					280					285					
Leu	Ile	Cys	Glu	Glu	Lys	Ala	Pro	Val	Leu	Lys	Met	Glu	Leu	Asp	Arg		
	290					295					300						
Ser	Ala	Asp	Pro	Pro	Pro	Ala	Ile	Trp	Val	Ala	Thr	Thr	Lys	Ser	Thr		
305					310					315					320		
Val	Asn	Lys	Trp	Thr	Leu	Lys	Gly	Ile	His	Asn	Phe	Arg	Ala	Ser	Gly		
				325				330						335			
Asp	Tyr	Asp	Asn	Asp	Cys	Thr	Asn	Pro	Ile	Thr	Pro	Leu	Cys	Thr	Gln		
		340					345						350				
Pro	Asp	Gln	Val	Ile	Lys	Gly	Gly	Ala	Ser	Ile	Ile	Gln	Cys	His	Ile		
		355					360					365					
Leu	Asn	Asp	Lys	Arg	His	Ile	Leu	Thr	Lys	Asp	Thr	Asn	Asn	Asn	Val		
	370					375					380						
Ala	Tyr	Trp	Asp	Val	Leu	Lys	Ala	Cys	Lys	Val	Glu	Asp	Leu	Gly	Lys		
385					390					395					400		
Val	Asp	Phe	Glu	Asp	Glu	Ile	Lys	Lys	Arg	Phe	Lys	Met	Val	Tyr	Val		
			405						410					415			
Pro	Asn	Trp	Phe	Ser	Val	Asp	Leu	Lys	Thr	Gly	Met	Leu	Thr	Ile	Thr		
		420					425						430				
Leu	Asp	Glu	Ser	Asp	Cys	Phe	Ala	Ala	Trp	Val	Ser	Ala	Lys	Asp	Ala		
	435					440						445					
Gly	Phe	Ser	Ser	Pro	Asp	Gly	Ser	Asp	Pro	Lys	Leu	Asn	Leu	Gly	Gly		
	450				455					460							
Leu	Leu	Leu	Gln	Ala	Leu	Leu	Glu	Tyr	Trp	Pro	Arg	Thr	His	Val	Asn		
465					470					475					480		
Pro	Met	Asp	Glu	Glu	Glu	Asn	Glu	Val	Asn	His	Val	Asn	Gly	Glu	Gln		
			485					490						495			
Glu	Asn	Arg	Val	Gln	Lys	Gly	Asn	Gly	Tyr	Phe	Gln	Val	Pro	Pro	His		
		500					505						510				
Thr	Pro	Val	Ile	Phe	Gly	Glu	Ala	Gly	Gly	Arg	Thr	Leu	Phe	Arg	Leu		
		515					520					525					
Leu	Cys	Arg	Asp	Ser	Gly	Gly	Glu	Thr	Glu	Ser	Met	Leu	Leu	Asn	Glu		
	530				535						540						
Thr	Val	Pro	Gln	Trp	Val	Ile	Asp	Ile	Thr	Val	Asp	Lys	Asn	Met	Pro		
545					550					555					560		
Lys	Phe	Asn	Lys	Ile	Pro	Phe	Tyr	Leu	Gln	Pro	His	Ala	Ser	Ser	Gly		
			565				570							575			
Ala	Lys	Thr	Leu	Lys	Lys	Asp	Arg	Leu	Ser	Ala	Ser	Asp	Met	Leu	Gln		
		580					585						590				
Val	Arg	Lys	Val	Met	Glu	His	Val	Tyr	Glu	Lys	Ile	Ile	Asn	Leu	Asp		
		595					600					605					
Asn	Glu	Ser	Gln	Thr	Thr	Ser	Ser	Ser	Asn	Asn	Glu	Lys	Pro	Gly	Glu		
	610				615						620						
Gln	Glu	Lys	Glu	Glu	Asp	Ile	Ala	Val	Leu	Ala	Glu	Glu	Lys	Ile	Glu		
625					630					635					640		
Leu	Leu	Cys	Gln	Asp	Gln	Val	Leu	Asp	Pro	Asn	Met	Asp	Leu	Arg	Thr		
			645						650					655			
Val	Lys	His	Phe	Ile	Trp	Lys	Ser	Gly	Gly	Asp	Leu	Thr	Leu	His	Tyr		
			660					665					670				
Arg	Gln	Lys	Ser	Thr													
		675															

<210> 216
 <211> 1432
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >WRN|ENSG00000165392|ENST00000298139|4299

<400> 216
 Met Ser Glu Lys Lys Leu Glu Thr Thr Ala Gln Gln Arg Lys Cys Pro
 1 5 10 15
 Glu Trp Met Asn Val Gln Asn Lys Arg Cys Ala Val Glu Glu Arg Lys
 20 25 30
 Ala Cys Val Arg Lys Ser Val Phe Glu Asp Asp Leu Pro Phe Leu Glu
 35 40 45
 Phe Thr Gly Ser Ile Val Tyr Ser Tyr Asp Ala Ser Asp Cys Ser Phe
 50 55 60
 Leu Ser Glu Asp Ile Ser Met Ser Leu Ser Asp Gly Asp Val Val Gly
 65 70 75 80
 Phe Asp Met Glu Trp Pro Pro Leu Tyr Asn Arg Gly Lys Leu Gly Lys
 85 90 95
 Val Ala Leu Ile Gln Leu Cys Val Ser Glu Ser Lys Cys Tyr Leu Phe
 100 105 110
 His Val Ser Ser Met Ser Val Phe Pro Gln Gly Leu Lys Met Leu Leu
 115 120 125
 Glu Asn Lys Ala Val Lys Lys Ala Gly Val Gly Ile Glu Gly Asp Gln
 130 135 140
 Trp Lys Leu Leu Arg Asp Phe Asp Ile Lys Leu Lys Asn Phe Val Glu
 145 150 155 160
 Leu Thr Asp Val Ala Asn Lys Lys Leu Lys Cys Thr Glu Thr Trp Ser
 165 170 175
 Leu Asn Ser Leu Val Lys His Leu Leu Gly Lys Gln Leu Leu Lys Asp
 180 185 190
 Lys Ser Ile Arg Cys Ser Asn Trp Ser Lys Phe Pro Leu Thr Glu Asp
 195 200 205
 Gln Lys Leu Tyr Ala Ala Thr Asp Ala Tyr Ala Gly Phe Ile Ile Tyr
 210 215 220
 Arg Asn Leu Glu Ile Leu Asp Asp Thr Val Gln Arg Phe Ala Ile Asn
 225 230 235 240
 Lys Glu Glu Glu Ile Leu Leu Ser Asp Met Asn Lys Gln Leu Thr Ser
 245 250 255
 Ile Ser Glu Glu Val Met Asp Leu Ala Lys His Leu Pro His Ala Phe
 260 265 270
 Ser Lys Leu Glu Asn Pro Arg Arg Val Ser Ile Leu Leu Lys Asp Ile
 275 280 285
 Ser Glu Asn Leu Tyr Ser Leu Arg Arg Met Ile Ile Gly Ser Thr Asn
 290 295 300
 Ile Glu Thr Glu Leu Arg Pro Ser Asn Asn Leu Asn Leu Leu Ser Phe
 305 310 315 320
 Glu Asp Ser Thr Thr Gly Gly Val Gln Gln Lys Gln Ile Arg Glu His
 325 330 335
 Glu Val Leu Ile His Val Glu Asp Glu Thr Trp Asp Pro Thr Leu Asp

Met	Asp	Ser	Tyr	Leu	Met	Asn	His	Phe	Asp	Leu	Pro	Thr	Cys	Asp	Asn
	115					120					125				
Cys	Arg	Asp	Ala	Asp	Asp	Lys	His	Lys	Leu	Ile	Thr	Lys	Thr	Glu	Ala
	130					135					140				
Lys	Gln	Glu	Tyr	Leu	Leu	Lys	Asp	Cys	Asp	Leu	Glu	Lys	Arg	Glu	Pro
145				150						155					160
Pro	Leu	Lys	Phe	Ile	Val	Lys	Lys	Asn	Pro	His	His	Ser	Gln	Trp	Gly
			165					170						175	
Asp	Met	Lys	Leu	Tyr	Leu	Lys	Leu	Gln	Ile	Val	Lys	Arg	Ser	Leu	Glu
	180						185					190			
Val	Trp	Gly	Ser	Gln	Glu	Ala	Leu	Glu	Glu	Ala	Lys	Glu	Val	Arg	Gln
	195					200					205				
Glu	Asn	Arg	Glu	Lys	Met	Lys	Gln	Lys	Lys	Phe	Asp	Lys	Lys	Val	Lys
	210					215				220					
Glu	Leu	Arg	Arg	Ala	Val	Arg	Ser	Ser	Val	Trp	Lys	Arg	Glu	Thr	Ile
225					230					235					240
Val	His	Gln	His	Glu	Tyr	Gly	Pro	Glu	Glu	Asn	Leu	Glu	Asp	Asp	Met
			245					250						255	
Tyr	Arg	Lys	Thr	Cys	Thr	Met	Cys	Gly	His	Glu	Leu	Thr	Tyr	Glu	Lys
			260					265						270	

Met

<210> 218
 <211> 633
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >XRCC1|ENSG00000073050|ENST00000262887|1902

<400> 218

Met	Pro	Glu	Ile	Arg	Leu	Arg	His	Val	Val	Ser	Cys	Ser	Ser	Gln	Asp
1				5				10						15	
Ser	Thr	His	Cys	Ala	Glu	Asn	Leu	Leu	Lys	Ala	Asp	Thr	Tyr	Arg	Lys
			20					25					30		
Trp	Arg	Ala	Ala	Lys	Ala	Gly	Glu	Lys	Thr	Ile	Ser	Val	Val	Leu	Gln
	35					40					45				
Leu	Glu	Lys	Glu	Glu	Gln	Ile	His	Ser	Val	Asp	Ile	Gly	Asn	Asp	Gly
	50				55					60					
Ser	Ala	Phe	Val	Glu	Val	Leu	Val	Gly	Ser	Ser	Ala	Gly	Gly	Ala	Gly
65				70				75						80	
Glu	Gln	Asp	Tyr	Glu	Val	Leu	Leu	Val	Thr	Ser	Ser	Phe	Met	Ser	Pro
			85					90					95		
Ser	Glu	Ser	Arg	Ser	Gly	Ser	Asn	Pro	Asn	Arg	Val	Arg	Met	Phe	Gly
			100				105						110		
Pro	Asp	Lys	Leu	Val	Arg	Ala	Ala	Ala	Glu	Lys	Arg	Trp	Asp	Arg	Val
	115					120						125			
Lys	Ile	Val	Cys	Ser	Gln	Pro	Tyr	Ser	Lys	Asp	Ser	Pro	Phe	Gly	Leu
	130					135					140				
Ser	Phe	Val	Arg	Phe	His	Ser	Pro	Pro	Asp	Lys	Asp	Glu	Ala	Glu	Ala
145				150					155						160
Pro	Ser	Gln	Lys	Val	Thr	Val	Thr	Lys	Leu	Gly	Gln	Phe	Arg	Val	Lys

		595					600					605					
Pro	Arg	Trp	Ile	Tyr	Ser	Cys	Asn	Glu	Lys	Gln	Lys	Leu	Leu	Pro	His		
	610					615					620						
Gln	Leu	Tyr	Gly	Val	Val	Pro	Gln	Ala									
625					630												

<210> 219
 <211> 280
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >XRCC2|ENSG00000196584|ENST00000359321|843

<400> 219																	
Met	Cys	Ser	Ala	Phe	His	Arg	Ala	Glu	Ser	Gly	Thr	Glu	Leu	Leu	Ala		
1				5					10					15			
Arg	Leu	Glu	Gly	Arg	Ser	Ser	Leu	Lys	Glu	Ile	Glu	Pro	Asn	Leu	Phe		
			20					25					30				
Ala	Asp	Glu	Asp	Ser	Pro	Val	His	Gly	Asp	Ile	Leu	Glu	Phe	His	Gly		
		35					40					45					
Pro	Glu	Gly	Thr	Gly	Lys	Thr	Glu	Met	Leu	Tyr	His	Leu	Thr	Ala	Arg		
	50				55						60						
Cys	Ile	Leu	Pro	Lys	Ser	Glu	Gly	Gly	Leu	Glu	Val	Glu	Val	Leu	Phe		
65				70					75					80			
Ile	Asp	Thr	Asp	Tyr	His	Phe	Asp	Met	Leu	Arg	Leu	Val	Thr	Ile	Leu		
			85					90					95				
Glu	His	Arg	Leu	Ser	Gln	Ser	Ser	Glu	Glu	Ile	Ile	Lys	Tyr	Cys	Leu		
			100					105					110				
Gly	Arg	Phe	Phe	Leu	Val	Tyr	Cys	Ser	Ser	Ser	Thr	His	Leu	Leu	Leu		
		115					120					125					
Thr	Leu	Tyr	Ser	Leu	Glu	Ser	Met	Phe	Cys	Ser	His	Pro	Ser	Leu	Cys		
		130				135					140						
Leu	Leu	Ile	Leu	Asp	Ser	Leu	Ser	Ala	Phe	Tyr	Trp	Ile	Asp	Arg	Val		
				150						155				160			
Asn	Gly	Gly	Glu	Ser	Val	Asn	Leu	Gln	Glu	Ser	Thr	Leu	Arg	Lys	Cys		
			165					170						175			
Ser	Gln	Cys	Leu	Glu	Lys	Leu	Val	Asn	Asp	Tyr	Arg	Leu	Val	Leu	Phe		
		180					185					190					
Ala	Thr	Thr	Gln	Thr	Ile	Met	Gln	Lys	Ala	Ser	Ser	Ser	Ser	Glu	Glu		
		195				200						205					
Pro	Ser	His	Ala	Ser	Arg	Arg	Leu	Cys	Asp	Val	Asp	Ile	Asp	Tyr	Arg		
	210				215						220						
Pro	Tyr	Leu	Cys	Lys	Ala	Trp	Gln	Gln	Leu	Val	Lys	His	Arg	Met	Phe		
225				230					235					240			
Phe	Ser	Lys	Gln	Asp	Asp	Ser	Gln	Ser	Ser	Asn	Gln	Phe	Ser	Leu	Val		
			245					250					255				
Ser	Arg	Cys	Leu	Lys	Ser	Asn	Ser	Leu	Lys	Lys	His	Phe	Phe	Ile	Ile		
		260					265						270				
Gly	Glu	Ser	Gly	Val	Glu	Phe	Cys										
	275					280											

<210> 220

<211> 346
 <212> PRT
 <213> Homo sapiens

<220>

<223> >XRCC3|ENSG00000126215|ENST00000553264|1041

<400> 220

```

Met Asp Leu Asp Leu Leu Asp Leu Asn Pro Arg Ile Ile Ala Ala Ile
1      5      10      15
Lys Lys Ala Lys Leu Lys Ser Val Lys Glu Val Leu His Phe Ser Gly
20      25      30
Pro Asp Leu Lys Arg Leu Thr Asn Leu Ser Ser Pro Glu Val Trp His
35      40      45
Leu Leu Arg Thr Ala Ser Leu His Leu Arg Gly Ser Ser Ile Leu Thr
50      55      60
Ala Leu Gln Leu His Gln Gln Lys Glu Arg Phe Pro Thr Gln His Gln
65      70      75      80
Arg Leu Ser Leu Gly Cys Pro Val Leu Asp Ala Leu Leu Arg Gly Gly
85      90      95
Leu Pro Leu Asp Gly Ile Thr Glu Leu Ala Gly Arg Ser Ser Ala Gly
100     105     110
Lys Thr Gln Leu Ala Leu Gln Leu Cys Leu Ala Val Gln Phe Pro Arg
115     120     125
Gln His Gly Gly Leu Glu Ala Gly Ala Val Tyr Ile Cys Thr Glu Asp
130     135     140
Ala Phe Pro His Lys Arg Leu Gln Gln Leu Met Ala Gln Gln Pro Arg
145     150     155     160
Leu Arg Thr Asp Val Pro Gly Glu Leu Leu Gln Lys Leu Arg Phe Gly
165     170     175
Ser Gln Ile Phe Ile Glu His Val Ala Asp Val Asp Thr Leu Leu Glu
180     185     190
Cys Val Asn Lys Lys Val Pro Val Leu Leu Ser Arg Gly Met Ala Arg
195     200     205
Leu Val Val Ile Asp Ser Val Ala Ala Pro Phe Arg Cys Glu Phe Asp
210     215     220
Ser Gln Ala Ser Ala Pro Arg Ala Arg His Leu Gln Ser Leu Gly Ala
225     230     235     240
Thr Leu Arg Glu Leu Ser Ser Ala Phe Gln Ser Pro Val Leu Cys Ile
245     250     255
Asn Gln Val Thr Glu Ala Met Glu Glu Gln Gly Ala Ala His Gly Pro
260     265     270
Leu Gly Phe Trp Asp Glu Arg Val Ser Pro Ala Leu Gly Ile Thr Trp
275     280     285
Ala Asn Gln Leu Leu Val Arg Leu Leu Ala Asp Arg Leu Arg Glu Glu
290     295     300
Glu Ala Ala Leu Gly Cys Pro Ala Arg Thr Leu Arg Val Leu Ser Ala
305     310     315     320
Pro His Leu Pro Pro Ser Ser Cys Ser Tyr Thr Ile Ser Ala Glu Gly
325     330     335
Val Arg Gly Thr Pro Gly Thr Gln Ser His
340     345

```

<210> 221
 <211> 336
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >XRCC4|ENSG00000152422|ENST00000511817|1011

<400> 221
 Met Glu Arg Lys Ile Ser Arg Ile His Leu Val Ser Glu Pro Ser Ile
 1 5 10 15
 Thr His Phe Leu Gln Val Ser Trp Glu Lys Thr Leu Glu Ser Gly Phe
 20 25 30
 Val Ile Thr Leu Thr Asp Gly His Ser Ala Trp Thr Gly Thr Val Ser
 35 40 45
 Glu Ser Glu Ile Ser Gln Glu Ala Asp Asp Met Ala Met Glu Lys Gly
 50 55 60
 Lys Tyr Val Gly Glu Leu Arg Lys Ala Leu Leu Ser Gly Ala Gly Pro
 65 70 75 80
 Ala Asp Val Tyr Thr Phe Asn Phe Ser Lys Glu Ser Cys Tyr Phe Phe
 85 90 95
 Phe Glu Lys Asn Leu Lys Asp Val Ser Phe Arg Leu Gly Ser Phe Asn
 100 105 110
 Leu Glu Lys Val Glu Asn Pro Ala Glu Val Ile Arg Glu Leu Ile Cys
 115 120 125
 Tyr Cys Leu Asp Thr Ile Ala Glu Asn Gln Ala Lys Asn Glu His Leu
 130 135 140
 Gln Lys Glu Asn Glu Arg Leu Leu Arg Asp Trp Asn Asp Val Gln Gly
 145 150 155 160
 Arg Phe Glu Lys Cys Val Ser Ala Lys Glu Ala Leu Glu Thr Asp Leu
 165 170 175
 Tyr Lys Arg Phe Ile Leu Val Leu Asn Glu Lys Lys Thr Lys Ile Arg
 180 185 190
 Ser Leu His Asn Lys Leu Leu Asn Ala Ala Gln Glu Arg Glu Lys Asp
 195 200 205
 Ile Lys Gln Glu Gly Glu Thr Ala Ile Cys Ser Glu Met Thr Ala Asp
 210 215 220
 Arg Asp Pro Val Tyr Asp Glu Ser Thr Asp Glu Glu Ser Glu Asn Gln
 225 230 235 240
 Thr Asp Leu Ser Gly Leu Ala Ser Ala Ala Val Ser Lys Asp Asp Ser
 245 250 255
 Ile Ile Ser Ser Leu Asp Val Thr Asp Ile Ala Pro Ser Arg Lys Arg
 260 265 270
 Arg Gln Arg Met Gln Arg Asn Leu Gly Thr Glu Pro Lys Met Ala Pro
 275 280 285
 Gln Glu Asn Gln Leu Gln Glu Lys Glu Asn Ser Arg Pro Asp Ser Ser
 290 295 300
 Leu Pro Glu Thr Ser Lys Lys Glu His Ile Ser Ala Glu Asn Met Ser
 305 310 315 320
 Leu Glu Thr Leu Arg Asn Ser Ser Pro Glu Asp Leu Phe Asp Glu Ile
 325 330 335

<210> 222
 <211> 609
 <212> PRT
 <213> Homo sapiens

<220>
 <223> >XRCC6|ENSG00000196419|ENST00000359308|1830

<400> 222
 Met Ser Gly Trp Glu Ser Tyr Tyr Lys Thr Glu Gly Asp Glu Glu Ala
 1 5 10 15
 Glu Glu Glu Gln Glu Glu Asn Leu Glu Ala Ser Gly Asp Tyr Lys Tyr
 20 25 30
 Ser Gly Arg Asp Ser Leu Ile Phe Leu Val Asp Ala Ser Lys Ala Met
 35 40 45
 Phe Glu Ser Gln Ser Glu Asp Glu Leu Thr Pro Phe Asp Met Ser Ile
 50 55 60
 Gln Cys Ile Gln Ser Val Tyr Ile Ser Lys Ile Ile Ser Ser Asp Arg
 65 70 75 80
 Asp Leu Leu Ala Val Val Phe Tyr Gly Thr Glu Lys Asp Lys Asn Ser
 85 90 95
 Val Asn Phe Lys Asn Ile Tyr Val Leu Gln Glu Leu Asp Asn Pro Gly
 100 105 110
 Ala Lys Arg Ile Leu Glu Leu Asp Gln Phe Lys Gly Gln Gln Gly Gln
 115 120 125
 Lys Arg Phe Gln Asp Met Met Gly His Gly Ser Asp Tyr Ser Leu Ser
 130 135 140
 Glu Val Leu Trp Val Cys Ala Asn Leu Phe Ser Asp Val Gln Phe Lys
 145 150 155 160
 Met Ser His Lys Arg Ile Met Leu Phe Thr Asn Glu Asp Asn Pro His
 165 170 175
 Gly Asn Asp Ser Ala Lys Ala Ser Arg Ala Arg Thr Lys Ala Gly Asp
 180 185 190
 Leu Arg Asp Thr Gly Ile Phe Leu Asp Leu Met His Leu Lys Lys Pro
 195 200 205
 Gly Gly Phe Asp Ile Ser Leu Phe Tyr Arg Asp Ile Ile Ser Ile Ala
 210 215 220
 Glu Asp Glu Asp Leu Arg Val His Phe Glu Glu Ser Ser Lys Leu Glu
 225 230 235 240
 Asp Leu Leu Arg Lys Val Arg Ala Lys Glu Thr Arg Lys Arg Ala Leu
 245 250 255
 Ser Arg Leu Lys Leu Lys Leu Asn Lys Asp Ile Val Ile Ser Val Gly
 260 265 270
 Ile Tyr Asn Leu Val Gln Lys Ala Leu Lys Pro Pro Pro Ile Lys Leu
 275 280 285
 Tyr Arg Glu Thr Asn Glu Pro Val Lys Thr Lys Thr Arg Thr Phe Asn
 290 295 300
 Thr Ser Thr Gly Gly Leu Leu Leu Pro Ser Asp Thr Lys Arg Ser Gln
 305 310 315 320
 Ile Tyr Gly Ser Arg Gln Ile Ile Leu Glu Lys Glu Glu Thr Glu Glu
 325 330 335
 Leu Lys Arg Phe Asp Asp Pro Gly Leu Met Leu Met Gly Phe Lys Pro

Claims

1. An inhibitor of ATR kinase, for use in a method of treating a hyper-proliferative disease in a subject, wherein said subject or the hyper-proliferative disease is characterized by one or more biomarker(s) selected from
 - a) one or more functional mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, APC, ATG5, ARID1A, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRAF, BRCA2, BRIP1, CCND1, CCNE1, CCNE2, CDC7, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, DYRK1A, EGFR, ERBB2, ERBB3, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, HDAC2, H2AFX, HRAS, KRAS, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, MYC, NBN, NRAS, PALB2, PARP1, PARP2, PARP3, PARP4, PCNA, PIK3CA, PMS2, POLA1, POLB, POLH, POLL, POLN, POLQ, PRKDC, PTEN, RAD9A, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TMRSS2, TMRSS2-ERG, TOPBP1, TOP2A, TOP2B, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4 and/or XRCC6 gene/protein; and/or
 - b) the activation of the ALT pathway; and/or
 - c) microsatellite instability.
2. The inhibitor of ATR kinase for use according to claim 1, which is 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof.
3. The inhibitor of ATR kinase for use according to claims 1 or 2, wherein the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), which comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, APC, ARID1A, ATG5, ATM, ATR, ATRIP, ATRX, BAP1, BARD1, BLM, BRCA2, BRIP1, CDK12, CHEK1, CHEK2, DCLRE1A, DCLRE1B, DCLRE1C, ERCC2, ERCC3, ERCC4, ERCC5, FAM175A, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FBXO18, FBXW7, FEN1, GEN1, H2AFX, HDAC2, LIG4, MDC1, MLH1, MLH3, MRE11A, MSH2, MSH3, MSH6, NBN, PALB2, PARP1, PARP2, PARP3, PARP4, PMS2, POLA1, POLB, POLH, POLN, POLN, POLQ, PRKDC, PTEN, RAD17, RAD18, RAD50, RAD51, RAD52, RAD54B, RAD54L, RAD9A, RB1, REV3L, RPA1, RPA2, SLX4, TDP1, TDP2, TP53, TP53BP1, TRRAP, UBE2N, UIMC1, USP1, WDR48, WRN, XPA, XRCC1, XRCC2, XRCC3, XRCC4

and/or XRCC6.

4. The inhibitor of ATR kinase for use according to claims 1 or 2, wherein the subject or the hyper-proliferative disease is characterized by one or more biomarker(s), which comprise(s) one or more deleterious mutation(s) in one or more gene(s)/protein(s) selected from BRCA1, ATM, BLM, BRCA2, ERCC5, FEN1, FANCD2, FANCG, H2AFX, PARP1, PCNA, POLL, POLB/POLL, RAD9A, RAD17, RAD52, REV3L, TDP2, TP53BP1, UBE2N, XPA.
5. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is (are) selected from BRCA1, ATM, BLM, ERCC5, FEN1, FANCD2, H2AFX, PARP1, PCNA, RAD9A, RAD17, REV3L, TP53BP1, UBE2N.
6. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is (are) selected from BRCA1, ATM, FANCD2, H2AFX, RAD17, UBE2N.
7. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is (are) selected from BRCA1, ATM, FEN1, H2AFX, PCNA.
8. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is BRCA1.
9. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is ATM.
10. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is FANCD2.
11. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is H2AFX.
12. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is RAD17.
13. The inhibitor of ATR kinase for use according to claim 4, wherein the one or more gene(s)/protein(s) is UBE2N.

14. The inhibitor of ATR kinase for use according to any one of claims 1 to 13, wherein the method of treating a hyper-proliferative disease comprises the steps:

- a) determining if one or more of the biomarker(s) defined in any one of claims 1 to 13 are present in a sample of the subject;
- 5 b) administering a therapeutically effective amount of the inhibitor of ATR kinase, preferably of 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt, to the subject, if one or more of the biomarker(s) determined by step (b) is (are) determined positively.

10

15. The inhibitor of ATR kinase for use according to claim 14, wherein the sample of the subject is an in vitro sample.

16. Use of one or more of the biomarker(s) defined in any one of claims 1 to 13 for identifying a
15 subject with a hyper-proliferative disease who is disposed to respond favorably to Compound A.

17. A kit comprising 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt together with means to detect one or more of the biomarker(s) defined in any one
20 of claims 1 to 13.

18. A method for identifying a subject having a hyper-proliferative disease disposed to respond favorably to an inhibitor of ATR kinase, preferably to 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine or a tautomer, an N-oxide, a
25 hydrate, a solvate, or a pharmaceutically acceptable salt thereof, wherein the method comprises the detection of one or more of the biomarker(s) defined in any one of claims 1 to 13 in a sample of the subject.

19. A method of determining whether a subject having a hyper-proliferative disease will respond to the
30 treatment with an inhibitor of ATR kinase, preferably with 2-[(3R)-3-methylmorpholin-4-yl]-4-(1-methyl-1H-pyrazol-5-yl)-8-(1H-pyrazol-5-yl)-1,7-naphthyridine or a tautomer, an N-oxide, a hydrate, a solvate, or a pharmaceutically acceptable salt thereof, wherein the method comprises the detection of one or more of the biomarker(s) defined in any one of claims 1 to 13 in a sample of the subject.

35