

- (54) Title  
**MLK4 gene, a new diagnostic and prognostic marker in cancers**
- (51) International Patent Classification(s)  
**C12Q 1/68** (2006.01)
- (21) Application No: **2010355523** (22) Date of Filing: **2010.06.15**
- (87) WIPO No: **WO11/158061**
- (43) Publication Date: **2011.12.22**  
(44) Accepted Journal Date: **2014.09.11**
- (71) Applicant(s)  
**Miriam Martini;Alberto Bardelli**
- (72) Inventor(s)  
**Bardelli, Alberto;Martini, Miriam**
- (74) Agent / Attorney  
**Shelston IP, L 21 60 Margaret St, Sydney, NSW, 2000**
- (56) Related Art  
**Martini M et al. MLK4, a novel oncogenic kinase in the RAS pathway. SIBBM Seminar "Frontiers in Molecular Biology", Padua, 3-5 June 2010, P26.**  
**Bardelli A. Alterazioni genetiche e analisi funzionale della proteina chinasi MLK4 nei tumori umani. OSIRIS - Centro di Documentazione e Condivisione delle Conoscenze. 2009**

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
22 December 2011 (22.12.2011)

PCT

(10) International Publication Number  
**WO 2011/158061 A1**

(51) International Patent Classification:  
*C12Q 1/68* (2006.01)

(21) International Application Number:  
PCT/IB2010/052684

(22) International Filing Date:  
15 June 2010 (15.06.2010)

(25) Filing Language: English

(26) Publication Language: English

(72) Inventors; and

(71) Applicants : BARDELLI, Alberto [IT/IT]; Via Guinice-  
celli, 11/4, I-10132 Torino (IT). MARTINI, Miriam [IT/  
IT]; Via Sciesa, 22, I-10078 Venaria (Torino) (IT).

(74) Agent: FREYRIA FAVA, Cristina; c/o Buzzi, Notaro &  
Antonielli d'Oulx S.r.l., Via Maria Vittoria, 18, I-10123  
Torino (IT).

(81) Designated States (*unless otherwise indicated, for every  
kind of national protection available*): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,  
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,  
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,  
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,  
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,  
NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD,  
SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR,  
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every  
kind of regional protection available*): ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,  
ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,  
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,  
LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK,  
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))



WO 2011/158061 A1

(54) Title: MLK4 GENE, A NEW DIAGNOSTIC AND PROGNOSTIC MARKER IN CANCERS

(57) Abstract: An in vitro diagnostic method for determining invasive potential of cancer comprising measuring MLK4 gene expression in a cancer cell sample, wherein MLK4 gene overexpression is indicative of an invasive cancer, preferably a colorectal, bladder, breast, gastric, melanoma, lung, ovary or GMB cancer.

## MLK4 gene, a new diagnostic and prognostic marker in cancers

### Technical field of the invention

5 The present disclosure concerns the identification of a new diagnostic and prognostic marker in cancers. More specifically, such a new marker belongs to the family of Mixed Lineage Kinases (MLKs). This marker plays a central role in the diagnosis, prognosis and the response to therapy of tumors bearing mutations in the KRAS and BRAF genes. Such tumors are characterized by an extremely poor prognosis and lack of response to therapy.

### 10 Background art

Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of common general knowledge in the field.

15 Targeting of deregulated protein kinases has proven effective for multiple cancers types. Using high-throughput mutational profiling of kinase genes, we previously identified somatic mutations in Mixed Lineage Kinase 4 (MLK4) in colorectal tumors (CRC) (1). MLK4 belongs to the family of MLK serine-threonine kinases thought to activate multiple intracellular signaling pathways. They are characterized by an amino-terminal SRC-homology domain (SH3), followed  
20 sequentially by a kinase domain, a leucine-zipper region and a Cdc42/Rac Interactive Binding (CRIB) motif. The carboxyl terminus of all MLKs is proline-rich but diverge significantly among different members of the family, indicating that these regions might serve different regulatory functions. The discovery of MLK4 somatic mutations in CRCs  
25 is presently known about the biochemical and cellular properties of MLK4 in normal and neoplastic cells.

It is an object of the present invention to overcome or ameliorate at least one of the disadvantages of the prior art, or to provide a useful alternative.

**Summary of the invention**

According to a first aspect, the present invention provides an *in vitro* diagnostic method comprising measuring MLK4 gene expression in a tumor cell sample material of a patient, wherein measuring MLK4 gene overexpression in said tumor cell sample material is indicative of an invasive potential of said tumor.

According to a second aspect, the present invention provides a method of treating cancer in a patient whose tumor bears a mutation in the nucleotide sequence of the gene KRAS, said method comprising administering an inhibitor of MLK4 enzymatic activity to said patient to reduce or abrogate MLK4 RNA or protein expression.

According to a third aspect, the present invention provides a method of treating cancer in a patient having a tumor, wherein said tumor bears at least one mutation in the KRAS gene encoded protein, and wherein said at least one mutation of the KRAS gene encoded protein is selected among G12S, G12V, G12D, G12A, G12C, G13A and G13D, said method comprising administering an antagonist and/or an inhibitor of MLK4 expression and/or enzymatic activity to said patient.

According to a fourth aspect, the present invention provides use of an inhibitor of MLK4 enzymatic activity in the manufacture of a medicament for reducing or abrogating MLK4 RNA or protein expression in a patient whose tumor bears a mutation in the nucleotide sequence of the gene KRAS.

According to a fifth aspect, the present invention provides use of an antagonist and/or inhibitor of MLK4 expression and/or MLK4 enzymatic activity in the manufacture of a medicament for the treatment of cancer in a patient whose tumor bears at least one mutation in the KRAS gene encoded protein, and wherein said at least one mutation of the KRAS gene encoded protein is selected among G12S, G12V, G12D, G12A, G12C, G13A and G13D.

Unless the context clearly requires otherwise, throughout the description and the claims, the words “comprise”, “comprising”, and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of “including, but not limited to”.

In one or more preferred embodiments, the present invention relates to a new diagnostic and prognostic marker in cancers.

The claims form an integral part of the technical teaching herein provided in relation to the present invention.

In some preferred embodiments the invention relates to an *in vitro* diagnostic method which, by measuring of MLK4 gene expression, more specifically MLK4 gene overexpression in a cancer cell sample material, is able to determine the potential of invasiveness of the cancer.

In an embodiment, the present disclosure concerns measuring MLK4 protein expression and/or MLK4 coding nucleic acid, preferably RNA, expression.

In a further embodiment, the present disclosure concerns measuring MLK4 gene overexpression in a bioptic sample, wherein the bioptic sample corresponds to the invasive portion of the cancer cell sample material.

In a further embodiment, the present disclosure concerns measuring wild-type MLK4 and/or mutated MLK4 gene expression, wherein the mutated MLK4 gene contains i) at least one somatic mutation at the amino acid level coding for at least one amino acid mutation selected from H261Y, H261Q, G291E, A293E, W296Stp, R470C, R553Stp, R555Stp, N596I, K629E, P843S, H890P, and M894T mutations or ii) at least one somatic mutation at the nucleic acid level selected from C781T, C783G, G872A, C878A, G888A, C1408T, C1657T, C1663T, A1787T, A1885G, C2788T, A2669C, T2681C mutation.

In a still further embodiment, the present disclosure concerns an *in vitro* diagnostic method, wherein the MLK4 gene expression measurement is performed using a reagent able to selectively bind to MLK4 protein and/or MLK4 coding nucleic acid, which is directly or indirectly labeled with a detectable substance.

In a further embodiment, the present disclosure concerns an *in vitro* diagnostic method, wherein the MLK4 gene expression measurement is performed in combination with the KRAS gene mutations and expression measurement.

In a further embodiment, the present disclosure concerns inactivation of the MLK4 protein expression or its biochemical function as a mean to therapeutically treat tumors bearings KRAS gene mutations.

#### **Brief description of the drawings**

The present invention will now be described in detail in relation to some \_\_\_\_\_

preferred embodiments by way of non limiting examples, referring to the annexed figures, in which:

- **Figure 1. Analysis of MLK4 in human tumors.** A) Somatic mutations identified in the MLK4 gene in CRCs and glioblastomas. B) Moderately-differentiated infiltrating colorectal carcinoma. Several neoplastic glandular structures that vary in size and shape can be seen. At the lateral border (identified by arrows) the tumor cells show intense membrane and sub-membrane staining showing infiltration. The arrowheads indicate areas of strong MLK4 expression in neoplastic cells that are infiltrating the stroma. Note that noninfiltrating cells are negative or show weak apical staining. C) High magnification of the section presented in B. D) Metastasis of CRC in a regional lymph node. The fat tissue surrounding the lymph node is infiltrated by the metastatic cells. E) Higher magnification of the section presented in C. F) Liver metastasis of colorectal carcinoma. Only the cells at the periphery of the metastatic nodule (right side) are immunostained.

- **Figure 2. Functional interaction between oncogenic RAS and MLK4 on transformation, tumor growth and response to therapy.** A) Relative kinase activity of MLK4 wt and mutant alleles. The relative kinase activity was calculated as the ratio between total MLK4 protein and the amount of ATP consumption. Results were normalized using the relative activity of wt MLK4. B) Focus forming assay in NIH3T3 cells transfected with wt and mutant MLK4 alleles. Ras (V12) is used as a positive control. \*\*p value  $\leq 0.01$ ; error bars represent s.d. C) Focus forming assay in NIH3T3 cells co-transfected with mutant Ras (V12) and wt/mutant MLK4 alleles. \*\*p value  $\leq 0.01$ ; error bars represent s.d. D) Biochemical validation of HKE3 cancer cells expressing the indicated MLK4 variants. E) Tumor formation by HKE3 cancer cells in xenograft mouse models expressing the indicated MLK4 variants. Cells were injected in the side of nude mice and tumor growth was measured at the indicated time points. Error bars represent s.e.m. F) Biochemical validation of DKO3 cancer cells expressing the indicated MLK4 variants. G) Tumor formation by DKO3 cancer cells expressing the indicated MLK4 variants in xenograft mouse models. Cells were injected in the side of nude mice and tumor growth was measured at the indicated time points. Error bars represent s.e.m.

- **Figure 3. Invasive, tumorigenic and metastatic potential of MLK4 mutated alleles.** A) A549 cancer cells lines were transduced with lentiviral

vectors expressing wild type or two MLK4 mutants. MLK4 expression levels were assessed by western blotting on lysates of cell transduced with empty vector or the indicated constructs. The invasive potential was assessed by migration through a matrigel coated membrane in absence or upon stimulation with Hepatocyte Growth Factor (HGF). \*p value  $\leq 0.05$ ; \*\*p value  $\leq 0.01$ ; error bars represent s.d. B) Tumor formation by A549 cancer cells expressing wild type or two MLK4 mutants in xenograft mouse models. Cells were injected in the side of nude mice and tumor growth was measured at the indicated time points. Error bars represent s.e.m. C) Upon sacrifice the lungs were labeled by airway perfusion with the India ink, and superficial metastases were counted under a stereoscopic microscope. \*\*p value  $\leq 0.01$ ; error bars represent s.e.m. D) Invasive potential of transduced DLD1 tumor cell line. Experiments were performed as in A. \*p value  $\leq 0.05$ ; \*\*p value  $\leq 0.01$ ; error bars represent s.d. E) Tumor formation by DLD1 cancer cells expressing wild type or two MLK4 mutants in xenograft mouse models. Experiments were performed as in B. Error bars represent s.e.m. F) Analysis of DLD1-derived superficial lung metastases. Experiments were performed as in C. \*\*p value  $\leq 0.01$ ; error bars represent s.e.m.

- **Figure 4. Downregulation or abrogation of MLK4 expression impairs the transforming and invasive potential of human cancer cells.** A) Two independent ShRNAs targeting the MLK4 gene were used to downregulate MLK4 expression in multiple cell lines (A549, Colo205, DLD1, HCT116) from different tumor types. Expression levels of the MLK4 protein were assessed by western blotting. Anchorage-independent growth (soft agar) assay performed on parental and MLK4 knock down/knock out cells. \*p value  $\leq 0.05$ ; \*\*p value  $\leq 0.01$ ; error bars represent s.d. B) Schematic representation of the vector used to KO the MLK4 genes from the genome of the HCT116 cell line. A sequential targeting strategy was used to knock out (KO) both of the MLK4 alleles. L-ITR: Left-Inverted Terminal Repeats, P: neomycin promoter, Nco: neomycin resistance cassette, pA: polyadenylation sequence, R-ITR: Right-Inverted Terminal Repeats. C) Expression levels of the MLK4 protein assessed by western blotting in parental and knock out HCT116 cells. D) The invasive potential of parental and MLK4 knock-down or KO cells was assessed by migration through a matrigel coated membrane in absence or upon stimulation with Hepatocyte Growth Factor (HGF). \*p value  $\leq 0.05$ ; \*\*p value  $\leq 0.01$ ; error bars represent s.d. E) Control and MLK4 knock-down DLD1 cancer cells were injected in the side of nude mice and tumor

growth was measured at the indicated time points. Error bars represent s.e.m. F) Control and MLK4 knock-out HCT116 cancer cells were injected in the side of nude mice and tumor growth was measured at the indicated time points. Error bars represent s.e.m.

5           - **Figure 5. Effects of MLK4 abrogation on the KRAS-ERK signaling pathway.** A) Measurement of 'active' (GTP bound) RAS assessed by a pull-down assay with the RAS binding domain of C-RAF in parental and MLK4 knock out cells in the absence or presence of serum. Total RAS is used as a reference for sample loading. B) Biochemical analysis of MEK1/2 and ERK1/2 kinases  
10 phosphorylation levels. Cell lysates were subjected to western blotting with the indicated antibodies. C) Biochemical analysis of p38, SAPK/JNK and AKT  
15 kinases phosphorylation levels. Cell lysates were subjected to western blotting with the indicated antibodies. D) Comparison of MEK1/2 and ERK1/2 kinases  
phosphorylation levels in parental, MLK4 knock out and mutant KRAS KO  
HCT116 cells. E) Analysis of the level of expression of SHC and Grb2 adaptor  
proteins detected by western-blot analysis in parental and MLK4 knock out  
HCT116 cells.

          - **Figure 6.** A) Rates of cell proliferation of A549 and DLD1 cells transduced with wt and mutant MLK4. Error bars represent s.d. B) Cell  
20 proliferation rates of parental and MLK4 KO HCT116 cancer cells. \*p value  $\leq$  0.05; \*\*p value  $\leq$  0.01; error bars represent s.d. C) Schematic representation of the  
MLK4 signaling pathway according to this work.

          - **Figure 7.** The colorectal cancer cell line DiFi was transduced with lentiviral vectors expressing wild type or two MLK4 mutants and RAS V12. The  
25 indicated cell lines were treated with 0.1  $\mu$ g/ml of Cetuximab (Erbix) for 72 hours, after which cell viability was assessed by the ATP assay. Results shown are  
normalized to untreated cells; \*\*\*p value  $\leq$  0.001; error bars represent s.d.

### **Detailed description of embodiments of the invention**

30

The present invention will now be described in detail in relation to some preferred embodiments by way of non limiting examples.

In the following description, numerous specific details are given to provide a thorough understanding of embodiments. The embodiments can be practiced  
35 without one or more of the specific details, or with other methods, components,

materials, etc. In other instances, well-known structures, materials, or operations are not shown or described in detail to avoid obscuring aspects of the embodiments.

The headings provided herein are for convenience only and do not interpret the scope or meaning of the embodiments.

Using molecular analysis of tumor samples combined with biochemical and functional approaches we have assessed the role of MLK4 in cancer cells. We employed two complementary strategies: (i) ectopic expression of wt and mutated MLK4 alleles (forward genetics) and (ii) targeting MLK4 expression using ShRNA-mediated down regulation or somatic gene knock out (KO) of MLK4 gene (reverse genetics) in normal and cancer cell lines.

Mutational profiling of the coding region of the MLK4 gene was performed in multiple tumor types including bladder, breast, gastric, melanoma, lung, ovary and glioblastoma (GBM) samples (Table 1).

**Table 1.**

Tumor type	Histotype	Number of samples analysed
<b>Bladder</b>	<b>Total</b>	<b>35</b>
	transitional cell carcinoma	33
	cell line	2
<b>Breast</b>	<b>Total</b>	<b>88</b>
	ductal carcinoma	26
	lobular carcinoma	29
	medullary carcinoma	17
	mucinous carcinoma	16
<b>Gastric</b>	<b>Total</b>	<b>12</b>
	adenocarcinoma of gastroesophageal junction	12
<b>Glioma</b>	<b>Total</b>	<b>113</b>
	Glioblastoma	105
	anaplastic astrocytoma	2
	anaplastic oligodendroglioma	2
	high grade glioma cell lines	4
<b>Lung</b>	<b>Total</b>	<b>86</b>
	Adenocarcinoma	86
<b>Melanoma</b>	<b>Total</b>	<b>22</b>
	Primary	1
	nodal metastasis	12
	cutaneous metastasis	7
	visceral metastasis	2
<b>Ovary</b>	<b>Total</b>	<b>48</b>
	serous adenocarcinoma	48

In addition to those previously reported (1) we found novel somatic

mutations affecting the *MLK4* gene in GBMs (Fig. 1A). Overall, the mutation frequency in CRC and GBMs were 9/192 (5%) and 4/140 (3%), respectively. Most of the non-synonymous mutations involved evolutionarily conserved amino acids some of which were affected multiple times. In addition, a subset of the mutations occurred at amino acids previously shown to be pathogenic in other protein kinases (1). Furthermore, two nonsense mutations were identified that affected arginines R553 and R555 in the C-terminus. It is possible that latter truncating mutations influence the function of *MLK4* by removing the putative C-terminal inhibitory domain. Finally, the nonsense mutation (W296 stop), located in the kinase domain, may represent a “passenger”, rather than a “driver” mutation.

To assess whether *MLK4* deregulation occurred by mechanisms other than somatic mutations, we developed two polyclonal antibodies and used them to measure the expression pattern of *MLK4* in CRC samples. Immunohistochemical (IHC) analysis was performed on a panel of 38 CRCs. The results demonstrated that most tumors expressed low or undetectable levels of *MLK4* expression. However, 9 CRC samples (24%) showed clear *MLK4* positivity, which was very distinctive being limited at the invasive portion of the tumor (Fig. 1B). Areas of intensely stained neoplastic cells were typically observed infiltrating the stroma (Fig. 1C). In contrast, non-infiltrating cells were negative or showed a very weak expression signal. For one sample lymph node and liver metastasis were available in addition to the primary tumour (Fig. 1 D-F); in both, *MLK4* expression was again confined to the invasive portion of the tumor. DNA sequence analysis excluded that the *MLK4* gene was somatically mutated in the CRC samples that demonstrated increased expression. The *MLK4* expression pattern suggests that *MLK4* plays a central role in tumor invasion.

Lentiviral vectors were engineered to stably transduce *MLK4* wt cDNA and two of the mutant alleles, specifically H261Y, G291E and R470C, identified in the CRC sequencing screens (Table 1). The H261Y was selected for further studies since histidine 261 was found mutated twice (H261Y and H261Q). The G291E allele was selected as it is located in a region corresponding to the BRAF V600E oncogenic mutation found in multiple tumor (2). The R470C allele was chosen as this variant has been identified in two independent CRC samples (Table 1).

Biochemical *in vitro* assays were used to assess the catalytic activity of wt and mutated MLK4. As a recipient for these experiments we used colorectal cancer cells in which the MLK4 gene was genetically inactivated and as a result the endogenous MLK4 protein was absent (see below Figure 4B). These cells  
5 were transduced with wild type or the H261Y, G291E and R470C MLK4 variants. Mutated MLK4 proteins displayed increased kinase activity as compared to the wt counterpart (Fig. 2A).

Next we sought to establish the role of MLK4 on cellular phenotypes associated with tumorigenesis using both forward and reverse genetics strategies.  
10 We initially used the standard NIH3T3 focus forming assay to assess the transforming potential of wt and mutated *MLK4* alleles. Although a positive control (*RAS* G12V) was capable of promoting focus formation at high efficiency, no effects were seen when wt or mutated MLK4 were used (Fig. 2B). These results suggest that wt and mutant *MLK4* alleles are unable, *per se*, to support  
15 cellular transformation. On the contrary, when mutant MLK4 was co-transfected in NIH3T3 cells with the oncogenic version of *RAS* (G12V) a striking synergistic effect in focus formation was observed (Fig. 2C). Importantly, two of the mutants (G291E and R470C) but not wild type MLK4 increased the transforming potential of *RAS* G12V.

20 This behavior is reminiscent of other cancer genes (such as *Pokemon*) whose transforming potential becomes detectable only when assessed in cooperation with other oncogenes (3).

To further assess the role of MLK4 in *KRAS* mediated oncogenesis, we employed an isogenic variant (HKE3) of a tumorigenic human colorectal cancer  
25 cell line, HCT116, in which the mutant *KRAS* allele is disrupted by homologous recombination (4). Consequently, HKE3 cells are a suitable model to establish whether mutated *MLK4* alleles can rescue the tumorigenic phenotype of oncogenic *KRAS*. HKE3 expressing wt or the MLK4 mutants (G291E and R470C) that cooperated with *RAS* in the transformation assays (Fig. 2D) were  
30 subcutaneously injected into nude mice. Cells expressing the empty vector were used as controls. Two weeks after inoculation, HKE3 cells transduced with the empty vector or wt *MLK4* did not generate tumors (Fig. 2E). On the contrary mice inoculated with the G291E and R470C *MLK4* mutants developed subcutaneous tumors within 2 weeks. The observation that mutated *MLK4* alleles

rescue the tumorigenic phenotype of cancer cells, in which mutant *KRAS* was deleted, supports the concept that MLK4 plays a critical role in the same pathway.

To validate this observation in a different cellular system, we used an isogenic variant (DKO3) of another highly tumorigenic human colorectal cancer cell line (DLD-1), in which the mutant *KRAS* allele has been disrupted by homologous recombination. Similar *in vivo* experiments were performed in DKO3 cells transduced with MLK4 mutant alleles with comparable results (Fig. 2F-G).

The pivotal role of the KRAS pathway in human cancers has been further underlined by the discovery that KRAS mutated colorectal tumors do not respond to the anti EGFR monoclonal antibodies cetuximab (Erbix) and panitumumab (Vectibix) (5, WO-A-2008/112274). These findings were rapidly translated into the clinic and detection of KRAS mutations is now routinely used to select patients eligible for cetuximab therapy. We previously reported that the ectopic expression of mutated KRAS and BRAF impairs the response to cetuximab and panitumumab of colorectal cancer cells (6, 7). We reasoned that if MLK4 acts downstream to KRAS its oncogenic activation could also impairs the effects of cetuximab.

To test this hypothesis we used the DiFi colorectal cancer cell line which is inhibited by nanomolar concentration of cetuximab. In DiFi cells the KRAS pathway is activated by amplification/overexpression of the *EGFR* gene. DiFi cells are wild type for *MLK4*, *KRAS*, *BRAF*, *PIK3CA* and *PTEN* mutations, thus representing a suitable model for our experiments. Wt or mutated MLK4 proteins were expressed in DiFi cells and mutated KRAS (G12V) was used as positive control. Cells expressing mutated *MLK4* or *KRAS* alleles were less sensitive to cetuximab as indicated by their markedly increased survival upon treatment with this monoclonal antibody (Fig. 7).

These data indicate that the presence of oncogenic MLK4 can substitute for *KRAS* mutations in conferring resistance to EGFR targeted therapies, supporting the concept that KRAS and MLK4 are involved in the same signal transduction pathway (Figure 6D).

We next sought to assess the role of MLK4 alleles in *invasive-growth*, a genetic program leading to metastasis formation. To formally test this hypothesis, we evaluated whether ectopic expression of wt or mutated MLK4 might affect the invasive properties of cancer cells (A549 and DLD1). Both wt and mutant MLK4 mildly affected cell invasion measured by the matrigel Boyden chamber assays in

standard growth medium. We then assessed the invasiveness upon stimulation with the Hepatocyte Growth Factor (HGF) which has previously been linked to the *invasive growth* properties of cancer cells. In the presence of HGF, expression of both mutated *MLK4* alleles (G291E and R470C) markedly increased invasion as compared to the non-transduced cells or those expressing wt *MLK4* (Fig. 3A and 3D). To determine whether *MLK4* promoted cell proliferation might have contributed to the augmented invasive potential, cell proliferation rates were analyzed. Ectopic expression of *MLK4* did not affect proliferation (Fig. 6A).

To further evaluate the tumorigenic potential of *MLK4* mutations, A549 and DLD1 cells expressing wt *MLK4*, G291E and R470C mutants were subcutaneously injected into nude mice. Cells expressing the empty vector were used as controls. Mice inoculated with wt or mutated *MLK4* cells developed subcutaneous tumors within 3 weeks. Both A549 and DLD1 cells transduced with mutant *MLK4* gave rise to larger tumors compared to control cells (Fig. 3B and 3E). Metastasis formation in the lungs of mice bearing *MLK4* over-expressing tumors was then quantified. We found that the number of metastasis was increased in *MLK4* mutant expressing tumors versus controls (Fig. 3C and 3F) in both A549 and DLD1 models.

When considered together, the forward genetic approaches indicate that mutated *MLK4* alleles cooperate or substitute oncogenic *KRAS* in transformation and invasion and promote metastasis formation.

As a complementary approach, reverse genetics was used to evaluate how reduced expression or deletion of the *MLK4* gene affected the tumorigenic properties of cancer cells. We first identified ShRNA that targeted the *MLK4* sequence leading to efficient downregulation of its expression. Two independent *MLK4* ShRNAs were selected based on their efficiency in reducing *MLK4* expression in multiple cell lines derived from colon and lung cancers, including those utilized in the forward genetic approach (Fig. 4A, upper panel). We first investigated the effect of *MLK4* down-regulation on anchorage independent growth, a key feature of the neoplastic phenotype. Anchorage-independent growth of all cancer cell lines was significantly reduced or almost completely impaired (Fig. 4A).

Experiments based on ShRNA mediated down regulation may be misleading due to non-specific gene targeting. We therefore used homologous recombination to delete exon 1 (where the *MLK4* kinase domain is located) from

the HCT116 cells to generate cells with permanently abrogated MLK4 expression (Fig. 4B). A two-step genetic strategy was used to obtain first heterozygous and then homozygous cells in which both alleles of the *MLK4* locus were correctly targeted. Knocked out (KO) cells were viable yet lacked MLK4 expression confirmed by PCR and immunoblotting with the anti-MLK4 antibody (Fig. 4C). Soft agar growth of MLK4 KO cells was severely impaired (Fig. 4A) confirming and substantiating the results obtained with the Sh-mediated MLK4-silencing. Similarly, the invasive potential of knocked down and KO MLK4 cells was dramatically impaired (Fig. 4D).

Next, we determined the effect of reduced or abrogated MLK4 expression on the tumorigenic potential in colorectal cancer cells *in vivo*. To this end we took advantage of DLD1 and HCT116, two colorectal cancer cell lines which harbor a single copy of the activated *KRAS* oncogene. MLK4 knock down cells and those transduced with scrambled ShRNA were evaluated for the ability to growth in xenograft experiments. When inoculated into nude mice, DLD1 control cells rapidly formed solid tumors while MLK4 knock down cells did not (Fig. 4E). To independently confirm these results we used the HCT116 cell in which *MLK4* was genetically deleted. Xenograft experiments showed that cancer cells lacking MLK4 expression were virtually unable to form subcutaneous tumors while the corresponding isogenic wt cells rapidly grew forming large tumor masses (Fig. 4F).

To gain mechanistic insights into the role of MLK4 in RAS signaling, we took advantage of wt and *MLK4* KO HCT116 cells which carry constitutively active KRAS (G13D). It has been reported that MLKs are serine/threonine protein kinases that regulate signaling by mitogen-activated-protein kinase (MAPK) pathways. We reasoned that lack of MLK4 could affect different steps of the KRAS-MAPK signaling cascade. We initially analyzed the levels of 'active' (GTP bound) KRAS and the activation (phosphorylation) of its downstream effectors MEK, ERK, JNK, p38 and AKT.

We found that in the absence of MLK4 the amount of active (GTP-bound) RAS and the levels MEK and ERK phosphorylation were reduced (Fig. 5A and 5B). The absence of MLK4 did not affect the phosphorylation levels of JNK, p38 and AKT (Fig. 5C). Of note, in HCT116 cells the knock out of MLK4 or the deletion of mutant KRAS resulted in comparable reduction of MEK and ERK phosphorylation (Fig. 5D). Interestingly the reduction of KRAS-MAPK activation in

MLK4 knock out cells was accompanied by lower expression of SHC to a less extent Grb2, two known amplifier of the RAS signaling cascade (Fig. 5E).

Together with the functional experiments showing KRAS-MLK4 cooperation in promoting tumorigenesis, these results indicate that the main signaling axis (the ERK/MAPK cascade) activated by oncogenic KRAS is impaired in the absence of the MLK4 kinase (Figure 5F).

A key biological output resulting from MAPK activation is the induction of cell proliferation. We therefore evaluated the growing rate of parental and KO *MLK4* cells and found that the absence of MLK4 affected the proliferation of HCT116 cells (Fig. 6B). These findings further support the concept that MLK4 is involved in the proliferation of cells that are dependent on mutant *KRAS*.

This work was prompted by the finding that *MLK4* is somatically mutated in CRCs and GBMs. Mutations of the related MLK3 kinase were recently detected in 21% of microsatellite unstable (MSI) colorectal tumors (8). We sequenced the MLK4 exons where we previously detected somatic mutations in 30 MSI colorectal tumors and did not find any mutations.

Mutant alleles of MLK4 and KRAS work in concert to trigger cellular transformation and mutant MLK4 rescues the tumorigenic potential of cells in which mutant KRAS is deleted. When considered together these findings implicate MLK4 in KRAS mediated oncogenesis. Remarkably, knock down or knock out of MLK4 abrogates the growth of cancer cells carrying mutated KRAS.

We and others have previously targeted the *MET*, *KRAS* and *PIK3CA* oncogenes by homologous recombination in HCT116 cancer cells. While targeting of each of these oncogenes drives distinct biological phenotypes, only the deletion of mutated *KRAS* and MLK4 results in abrogation of tumorigenesis.

In conclusion, these data identify the MLK4 kinase as a novel effector of KRAS, a presently undruggable human oncogene (Fig. 6C). Pharmacological targeting of the MLK4 kinase is amenable and provides a novel therapeutic prospect for KRAS mutated tumors.

## **Materials and Methods**

### **Mutational Analysis**

PCR primers were designed using Primer 3 ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)), and synthesized by Invitrogen/Life Technologies, Inc. (Paisley, England). The PCR and sequencing primers are listed in Table 2.

**Table 2.**

Exon		Forward primer	Reverse primer	Sequence primer
1	1_1	CACCCGATCCCGATTCCTAC SEQ ID No.:1	TCCGAAGCCACCTGTACC SEQ ID No.:2	TCCCTACCCCCCTCCCTTC SEQ ID No.:3
	1_2	CTCAGCGTCCTCGTCGTC SEQ ID No.:4	CGCAGCTCGATGATGTTC SEQ ID No.:5	GCAGCTCGATGATGTTCG SEQ ID No.:6
	1_3	CACCTTCCCGCCAACTAC SEQ ID No.:7	GGCCAGTTGACCAGCAC SEQ ID No.:8	TCTTCCCGCCAACTACG SEQ ID No.:9
	1_4	CAGSAGGTGGCCGTGAAG SEQ ID No.:10	CSCCCTAAGCCTGACTCC SEQ ID No.:11	CTCCCACCTCTTCCATCTG SEQ ID No.:12
	1_5	GCACGTGCTGGTCAACTG SEQ ID No.:13	CSCCCTAAGCCTGACTCC SEQ ID No.:14	ACTGGCCCGTGCAGATAG SEQ ID No.:15
2		CAGGCACAGATAACCCACTAAAG SEQ ID No.:16	AAACCATAAAAAATTCATTCAAAAAGG SEQ ID No.:17	CATGCCACAGAAAAATGTG SEQ ID No.:18
	5	GGGAATTCAGCTTCCTCTG SEQ ID No.:19	CCCTATCTCTTCCCTGGTCTG SEQ ID No.:20	GGGGLCGCCAGTGTAGTGG SEQ ID No.:21
6	6_1	CGATCCTGCAATATTCTTTGC SEQ ID No.:22	GTGGCACCTACGAATGTCC SEQ ID No.:23	TTGCTGTGTCTTAATACATGG SEQ ID No.:24
	7	TTTCTCATTGTTTTCATGTTTGGAG SEQ ID No.:25	GCATGATAAATTCATCCTAGAAAATTG SEQ ID No.:26	TTTTGAGTAAGTAGTTTCACAACTG SEQ ID No.:27
8		TTGTGAGCCCTTCAAAACTG SEQ ID No.:28	CTTGCTAATCTCCCTCTGTC SEQ ID No.:29	TGCTCTGTATCCACCAAAACC SEQ ID No.:30
	9	GTA AAAACGACGGCCAGT GAAAAGGGCTGCATGTGTTT SEQ ID No.:31	CTCTCTGAGCATCTTCCCAA SEQ ID No.:32	GTA AAAACGACGGCCAGT SEQ ID No.:33
9	9_2	GCCTACATTGATCTACCTCTTGG SEQ ID No.:34	GTA AAAACGACGGCCAGT TCTTCTCTTCTTGGGCAAC SEQ ID No.:35	GTA AAAACGACGGCCAGT SEQ ID No.:36
	9_3	CATAAACACACGGCTCCTCA SEQ ID No.:37	GTA AAAACGACGGCCAGT AGCAGGCACATTGTGGAGAA SEQ ID No.:38	CTAAAACACCCCACT SEQ ID No.:39
9	9_4	GTA AAAACGACGGCCAGT CCACACCTTCTTCTCCACAA SEQ ID No.:40	CATTCCACATGTCTGCTGT SEQ ID No.:41	GTA AAAACGACGGCCAGT SEQ ID No.:42
	9_5	GTA AAAACGACGGCCAGT CAAGAACTTGCCGTCTTCC SEQ ID No.:43	CAAATPATTAAATGTATCACCAGGA SEQ ID No.:44	GTA AAAACGACGGCCAGT SEQ ID No.:45

5 PCR primers that amplify the selected exons and the flanking intronic sequences, including splicing donor and acceptor regions, were used and PCR products were on average 381 bps in length. PCR was performed in both 384- and 96-well formats in 5- or 10-uL reaction volumes, respectively, containing 0.25 mmol/L deoxynucleotide triphosphates, 1 umol/L each of the forward and reverse primers, 6% DMSO, 1x PCR buffer, 1 ng/uL DNA, and 0.05 unit/uL Platinum Taq (Invitrogen/Life Technologies). A touchdown PCR program was used for PCR amplification (Peltier Thermocycler, PTC-200, MJ Research, Bio-Rad Laboratories, Inc., Italy). PCR conditions were as follows: 94°C for 2 min; three

10

cycles of 94°C for 15 s, 64°C for 30 s, 70°C for 30 s; three cycles of 94°C for 15 s, 61°C for 30 s, 70°C for 30 s; three cycles of 94°C for 15 s, 58°C for 30 s, 70°C for 30 s; and 35 cycles of 94°C for 15 s, 57°C for 30 s, and 70°C for 30 s, followed by 70°C for 5 min and 12°C thereafter. PCR products were purified using AMPure (Agencourt Bioscience Corp., Beckman Coulter S.p.A, Milan, Italy). Cycle sequencing was carried out using BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA) with an initial denaturation at 97°C for 3 min, followed by 28 cycles of 97°C for 10 s, 50°C for 20 s, and 60°C for 2 min.

Sequencing products were purified using CleanSeq (Agencourt Bioscience, Beckman Coulter) and analyzed on a 3730 DNA Analyzer, ABI capillary electrophoresis system (Applied Biosystems). Sequence traces were analyzed using the Mutation Surveyor software package (SoftGenetics, State College, PA).

15

#### Immunohistochemical analysis

Paraffin embedded sections were dewaxed, hydrated, treated with proteinase K (DAKO, Glostrup, Denmark) and immunostained using a labelled polymer detection system (Bond Polymer Define Detection, Vision Biosystem, Mount Waverley, Australia) and automated stainer (BOND-maX, Vision BioSystem). The primary polyclonal antibody was used at a dilution of 1:1000. Negative controls were obtained by replacement of primary antiserum with buffer. Only tumors which exhibited intense membrane were categorized as overexpressing MLK4.

25

#### DNA Constructs and Mutagenesis

The cDNA encoding for MLK4 set forth in SEQ ID NO.:46 was synthesized by RT-PCR by first inserting into the TopoTA vector. Full-length MLK4 cDNA was subcloned into the pCEV29.1 (10) or into pRRL plasmid (11). Mutants of MLK4 containing point mutations were constructed using the QuikChange II XL Site-directed mutagenesis kit (Stratagene) with MLK4 wt plasmid as the template DNA. The following oligonucleotides, and their reverse complements, were used as mutagenesis primers:

G291E,  
5'-CTTTGAAGATTACAGATTTTGAGTTGGCGAGGGAATGGCAC-3'

35

(SEQ ID No:47);

R470C,

5'-TGAGCAGCAGCTGGCAGAGTGCAGATCGACGTGCTGGAG-3'

(SEQ ID No.:48).

5 The presence of the appropriate mutations was confirmed by DNA sequencing.

Sequences of G291E and R470C mutants of MLK4 are set forth in SEQ ID No: 49 and 50, respectively.

10 In the following table a list of the MLK4 mutation identified in the present study are reported with the corresponding SEQ ID No.

**Table 3.**

<b>MLK4 mutations Amino acid level</b>	<b>MLK4 mutations Nucleotide level</b>	<b>SEQ ID No.</b>
H261Y	C781T	51
H261Q	C783G	52
G291E	G872A	49
A293E	C878A	53
W296Stp	G888A	54
R470C	C1408T	50
R553Stp	C1657T	55
R555Stp	C1663T	56
N596I	A1787T	57
K629E	A1885G	58
P843S	C2788T	59
H890P	A2669C	60
M894T	T2681C	61

15 *Antibodies*

MLK4 specific polyclonal antibodies were developed using a standard immunization protocol (two rabbits immunized for each epitope), we generated polyclonal antisera and antibodies directed against either the N- or the C-terminus of MLK4. Both antibodies were tested in immunoprecipitation and Western blot  
20 experiment, showing that these antibodies can be used to detect a 114 kDA protein (predicted molecular weight of MLK4) in total cell lysates. The primary antibodies used for immunoblotting were: anti-Vinculin (Sigma-Aldrich), anti-Actin (S.Cruz), anti-PERK1/2, anti-ERK1/2, anti-P-SAPK/JNK, anti-SAPK/JNK, anti-P-AKT, anti-AKT, anti-P-p38 and anti-P-38, (Cell Signaling).

### Cell culture

A-549, Colo-201 and HCT116 were cultured in RPMI-1640 medium (Invitrogen), NIH3T3, DLD-1 and HCT116-HKE-3 derived clones were grown in DMEM (Invitrogen). HTERT-HME1 were cultured in growth medium containing DMEM/F-12 (Invitrogen) supplemented with 20 ng/mL EGF, 10 ug/ml insulin, and 100 ug/ml hydrocortisone, while DiFi cells were grown in F-12 (Invitrogen). Cultures of mammalian cells were maintained in the appropriate medium supplemented with 10% fetal bovine serum (Sigma-Aldrich), 2 mM Lglutamine, 100 units/ml penicillin/streptomycin at 37°C in humidified air with 5% CO<sub>2</sub>. The mouse NIH3T3 fibroblasts are available by ATCC (CRL-1658) and were cultured in DMEM supplemented with 10% bovine serum (heat-inactivated at 56°C for 30 min) (Invitrogen) and 2 mM L-glutamine. All the cell lines are publicly available by American Type Culture Collection (ATCC), Manassas, USA.

15

### Transfections and transformation assays

NIH3T3 fibroblasts were seeded at  $2 \times 10^5$  cells in 100 mm plate and grown for 24 h before transfection using high-efficiency liposome transfection method (Lipofectamine 2000 and Plus Reagent; Invitrogen). Each plate was transfected with 4 µg of the pCEV29.1 vector containing either no insert or a cDNA encoding for a MLK4 protein (wild-type or mutant). For co-operation experiments, cells were transiently cotransfected with 0.1 µg of RasV12 plasmid and 4µg of pCEV containing the different MLK4 wt and mutants. Two days after transfection, cells were split into three plates cultured in DMEM containing 5% Bovine Serum and used for measuring focus formation. Foci were scored two weeks after transfection following fixation with glutaraldehyde 11% and Giemsa staining.

25

### Protein analysis

Total cellular proteins were extracted by solubilizing the cells in boiling SDS buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 1% SDS). The amount of proteins was quantified by the BCA Protein Assay Reagent kit (Pierce Chemical Co.) Extracts were electrophoresed on SDS-polyacrylamide gels, transferred onto nitrocellulose membranes (Hybond; GE Healthcare) and blocked in phosphate buffered saline, 0.1% Tween 20, 5% BSA. The membrane was then

35

incubated with appropriate dilutions of primary antibody (either MLK4, vinculin, actin, pERK, ERK, pJNK, JNK, pAKT, AKT, pP38, P38), followed by the appropriate peroxidase conjugated secondary antibody (Donkey anti Rabbit Bio-Rad Laboratories). Final detection was done by enhanced chemiluminescence (GE Healthcare). For immunoprecipitations, cells were lysed at 4°C with 800µl of a buffer containing 20 mM Tris-HCl [pH 7.4], 150 mM NaCl, 10% glycerol, and 1% Triton X-100 in the presence of protease and phosphatase inhibitors (1 mM Na<sub>3</sub>VO<sub>4</sub>, 100 mM NaF, 1 mM PMSF, 10 µg/ml leupeptin, 10 µg/ml aprotinin, and 1 µg/ml pepstatin). Extracts were clarified at 12,000 rpm for 20 min and quantified by bicinchoninic assay. Protein A-Sepharose beads were pre-incubated with primary antibody (1h at +4°C) and then washed three times with lysis buffer before incubation with total cellular proteins. Extracts were incubated with primary antibody for 2 h at +4°C. Immune complexes were collected with protein A-Sepharose, washed three times in lysis buffer. SDS PAGE western blotting was performed as previously described.

#### Ras pull-down assay

Cell lysates were incubated with GST-RBD on glutathione-agarose beads at 4°C for 1h. The beads were then collected by centrifugation, resuspended in Laemmli reducing sample buffer and boiled for 8 minutes. Western blot analysis was performed using anti-Ras antibody (Calbiochem).

#### Kinase assays

Immunoprecipitated MLK4 was subjected to two independent strategies. For radioactive kinase assay, immunoprecipitates were washed three times with kinase assay buffer [50 mM Hepes (pH 7.5), 150 mM NaCl, 12.5 mM MgCl<sub>2</sub>, 2mM EGTA 1mM dithiothreitol and 1 mM sodium orthovanadate] and the incubated with kinase buffer in the presence of 15 µM ATP including [ $\gamma$ -<sup>32</sup>P]ATP for 30 min at 37°C. Reactions were stopped by the addition of 30 µl of Laemmli buffer and boiling for 1 min. Samples were analyzed by 15% SDS/PAGE followed by autoradiography.

#### Suppression of gene expression by RNAi

MLK4 expression was suppressed in tumor cells by lentiviral-mediated expression of shRNAs specifically targeting the MLK4 transcript. To rule out the

variability of biological responses, we used four different shRNA directed against MLK4 sequence. ShRNA sequences were derived from the Sigma-Aldrich Mission library under accession no. NM\_032435 (sequence 1, TRCN0000003209; sequence 2, TRCN0000003210; sequence 3, TRCN0000003211; sequence 4, TRCN0000003213). All sequences were able to suppress MLK4 levels while control ShCTRL (SHC002V) was not. The two most efficient ShRNA, among the different recipient cell lines, were then selected to be used in the subsequent experiments. Western blotting analysis of shRNA infected cells showed a reduction by over 80% of MLK4 protein. Results shown in the figures refer to sequence 2 and 4, renamed as Sh1 and Sh2 respectively.

#### Targeted deletion of the MLK4 locus in human cancer cells

Disruption of the MLK4 exon 1 in the CRC HCT116 cells was performed as previously described (12). The targeting vector pAAV-Neo-MLK4 was constructed by PCR; DLD1 genomic was used as template for both homology arms. Constructs and primer sequences are publicly available at Università degli Studi di Torino - Dipartimento Scienze Oncologiche, strada Provinciale 142, Candiolo (TO). Clones were selected after 2 weeks of growth under 0.4 mg/ml for HCT116 cells geneticin (Invitrogen, Carlsbad, CA) selection and then propagated in the absence of selective agents. Homologous recombination events were identified by locus-specific PCR screening.

#### Anchorage-independent growth assays

Cells were diluted to a concentration of 500-1000 cells/ml in appropriate medium containing 5% FBS, 0.5% Seaplaque agar. Cells were seeded in 24-well plates (1 ml/well) containing a 1 % agar underlay and supplemented three times a week with corresponding medium. Colonies were photographed and scored two weeks after seeding.

#### Invasion Assays

Cell migration and invasion assays were performed using Costar 24-well plate transwell either 8  $\mu$ m pore size transwell migration plates or Matrigel matrix-coated polycarbonate filters, respectively. Lower wells containing 500  $\mu$ l of media to which either 5% serum (to measure basal condition) or 10ng/ml HGF

(to measure induced invasion) had been added, whereas upper wells contained 100  $\mu$ l of cells, suspended to a concentration of  $10^5$  for HCT116 or  $5 \times 10^4$  for A549 and DLD1, in media containing the same serum concentration of lower wells. After incubation for 20 hr, cells/media in the upper wells were discarded, and the migrated cells on the underside of the transwell membranes were fixed with glutaraldehyde 11% in PBS, stained with Crystal Violet and counted with Metamorph Image Analysis Software.

#### Drug viability assay

Cetuximab was obtained from the Hospital Pharmacy. Cell lines were seeded in 100 $\mu$ l medium at appropriate density in 96-well plastic culture plates. Plates were incubated at 37°C in 5% CO<sub>2</sub> for 72h, after which cell viability was assessed by ATP content using the CellTiter-Glo® Luminescent Assay (Promega). Luminescence was detected using a DTX-880 plate reader (Beckman-Coulter).

#### Animal studies

All animal procedures were approved by the Ethical Commission of the University of Turin and by the Italian Ministry of Health. Six-week-old immunocompromised CD1-/- nude athymic female mice (Charles River Laboratories, Lecco, Italy) were injected subcutaneously in right posterior flanks. Tumor appearance was evaluate every 2d using a caliper. Tumor volume was calculated using the formula  $V = 4/3 \times (d/2)^2 \times (D/2)$ , where d is the minor tumor axis and D is the major tumor axis. Superficial pulmonary metastases were contrasted by black India ink air way infusion before excision, and were counted on dissected lung lobes under a stereoscopic microscope.

Cell lines were seeded in 100 $\mu$ l medium at appropriate density in 96-well plastic culture plates. Plates were incubated with cetuximab (Erbix) at 37°C in 5% CO<sub>2</sub> for 72h, after which cell viability was assessed by ATP content using the CellTiter-Glo® Luminescent Assay (Promega). Luminescence was detected using a DTX-880 plate reader (Beckman-Coulter).

**SEQUENCE LISTING**

<110> Bardelli, Alberto  
 Martini, Miriam  
 5  
 <120> MLK4 gene, a new diagnostic and prognostic marker in cancers  
 <130> BW012029-CF  
 10 <160> 61  
 <170> PatentIn version 3.5  
 15 <210> 1  
 <211> 19  
 <212> DNA  
 <213> artificial  
 20 <220>  
 <223> forward primer exon 1\_1  
 <400> 1  
 caccgatogc gattoctac 19  
 25 <210> 2  
 <211> 18  
 <212> DNA  
 <213> artificial  
 30 <220>  
 <223> reverse primer exon 1\_1  
 <400> 2  
 35 togaaggcga cgtgtacc 18  
 <210> 3  
 <211> 18  
 <212> DNA  
 <213> artificial  
 40 <220>  
 <223> sequence primer exon 1\_1  
 45 <400> 3  
 TccTaccccc Tggccttc 18  
 50 <210> 4  
 <211> 18  
 <212> DNA  
 <213> artificial  
 55 <220>  
 <223> forward primer exon 1\_2  
 <400> 4  
 60 cTcagcgTcc TggTgTc 18  
 <210> 5  
 <211> 18  
 <212> DNA  
 65 <213> artificial  
 <220>  
 <223> reverse primer exon 1\_2

<400> 5  
ggcagctcga tgatggtg 18

5 <210> 6  
<211> 18  
<212> DNA  
<213> artificial

10 <220>  
<223> sequence primer exon 1\_2

<400> 6  
ggcagctcga tgatggtg 18

15 <210> 7  
<211> 19  
<212> DNA  
<213> artificial

20 <220>  
<223> forward primer exon 1\_3

25 <400> 7  
catcttcccc gccaaactac 19

30 <210> 8  
<211> 18  
<212> DNA  
<213> artificial

35 <220>  
<223> reverse primer exon 1\_3

<400> 8  
ggcccagttg accagcac 18

40 <210> 9  
<211> 18  
<212> DNA  
<213> artificial

45 <220>  
<223> sequence primer exon 1\_3

50 <400> 9  
tcttccccgc caactacg 18

55 <210> 10  
<211> 18  
<212> DNA  
<213> artificial

60 <220>  
<223> forward primer exon 1\_4

<400> 10  
caggaggtgg ccgtgaag 18

65 <210> 11  
<211> 18  
<212> DNA  
<213> artificial

70 <220>  
<223> reverse primer exon 1\_4

<400> 11  
 cggcctaagg ctgacctcc 18

5  
 <210> 12  
 <211> 20  
 <212> DNA  
 <213> artificial

10  
 <220>  
 <223> sequence primer exon 1\_4

15  
 <400> 12  
 ctccacactc ttccatctg 20

20  
 <210> 13  
 <211> 18  
 <212> DNA  
 <213> artificial

25  
 <220>  
 <223> forward primer exon 1\_5

<400> 13  
 gcaactgctg gtcaactg 18

30  
 <210> 14  
 <211> 18  
 <212> DNA  
 <213> artificial

35  
 <220>  
 <223> reverse primer exon 1\_5

40  
 <400> 14  
 cggcctaagg ctgacctcc 18

45  
 <210> 15  
 <211> 18  
 <212> DNA  
 <213> artificial

<220>  
 <223> sequence primer exon 1\_5

50  
 <400> 15  
 actgggcccgt gcagctag 18

55  
 <210> 16  
 <211> 23  
 <212> DNA  
 <213> artificial

60  
 <220>  
 <223> forward primer exon 2

<400> 16  
 caqgcacaqa taaccacta aag 23

65  
 <210> 17  
 <211> 26  
 <212> DNA  
 <213> artificial

70  
 <220>

<223> reverse primer exon 2  
 <400> 17  
 5 aaaccataaa aattcattca aaaagg 26  
  
 <210> 18  
 <211> 20  
 10 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> sequence primer exon 2  
 15 <400> 18  
 catgccacag aaaaattgtg 20  
  
 <210> 19  
 20 <211> 21  
 <212> DNA  
 <213> artificial  
  
 <220>  
 25 <223> forward primer exon 5  
  
 <400> 19  
 30 ggggaatttca gotttcctt g 21  
  
 <210> 20  
 <211> 22  
 <212> DNA  
 35 <213> artificial  
  
 <220>  
 <223> reverse primer exon 5  
 40 <400> 20  
 ccctatctct ttcctggttc tg 22  
  
 <210> 21  
 <211> 18  
 45 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> sequence primer exon 5  
 50 <400> 21  
 ggggtcgcca gtgtagtg 18  
  
 <210> 22  
 <211> 21  
 <212> DNA  
 55 <213> artificial  
  
 <220>  
 <223> forward primer exon 6\_1  
 60 <400> 22  
 cgatcotgca atattctttg c 21  
 65  
  
 <210> 23  
 <211> 20  
 <212> DNA  
 70 <213> artificial

<220>  
 <223> reverse primer exon 6\_1  
 5 <400> 23  
 gtggcaccta cgaaatgtcc 20

<210> 24  
 <211> 22  
 10 <212> DNA  
 <213> artificial

<220>  
 <223> sequence primer exon 6\_1  
 15 <400> 24  
 llgclglglg claaalacal gg 22

<210> 25  
 <211> 25  
 <212> DNA  
 <213> artificial  
 20

<220>  
 <223> forward primer exon 7\_2  
 25 <400> 25  
 llcclcatlg llllcatgll lggag 25

<210> 26  
 <211> 27  
 <212> DNA  
 35 <213> artificial

<220>  
 <223> reverse primer exon 7\_2  
 40 <400> 26  
 gcatgataaa ttcctcctag aaaattg 27

<210> 27  
 <211> 26  
 <212> DNA  
 <213> artificial  
 45

<220>  
 <223> sequence primer exon 7\_2  
 50 <400> 27  
 ttttgagtea gtagttca caactg 26

<210> 28  
 <211> 20  
 <212> DNA  
 <213> artificial  
 55

<220>  
 <223> forward primer exon 8  
 60 <400> 28  
 ttgtcagccc ttcaaaactg 20

<210> 29  
 <211> 20  
 <212> DNA  
 <213> artificial  
 70

<220>  
 <223> reverse primer exon 8  
 5 <400> 29  
 ctgctaata tccccctgac 20

<210> 30  
 <211> 21  
 <212> DNA  
 <213> artificial  
 10

<220>  
 <223> sequence primer exon 8  
 <400> 30  
 tggctgtat ccaccaaac c 21

<210> 31  
 <211> 37  
 <212> DNA  
 <213> artificial  
 20

<220>  
 <223> forward primer exon 9\_1  
 <400> 31  
 30 gtaaaaacgac ggccagtcaa saggcctgca tgtgttt 37

<210> 32  
 <211> 21  
 <212> DNA  
 <213> artificial  
 35

<220>  
 <223> reverse primer exon 9\_1  
 <400> 32  
 40 ctctctgagc atcttccca a 21

<210> 33  
 <211> 17  
 <212> DNA  
 <213> artificial  
 45

<220>  
 <223> sequence primer exon 9\_1  
 <400> 33  
 50 gtaaaaacgac ggccagt 17

<210> 34  
 <211> 23  
 <212> DNA  
 <213> artificial  
 60

<220>  
 <223> forward primer exon 9\_2  
 <400> 34  
 65 gactacattg atctccctct tgg 23

<210> 35  
 <211> 37  
 <212> DNA  
 70

<213> artificial  
 <220>  
 <223> reverse primer exon 9\_2  
 5 <400> 35  
 gtaaaaacgac ggccagttct tctcttctt eggcaac 37

10 <210> 36  
 <211> 17  
 <212> DNA  
 <213> artificial

15 <220>  
 <223> sequence primer exon 9\_2  
 <400> 36  
 20 gtaaaaacgac ggccagt 17

<210> 37  
 <211> 20  
 <212> DNA  
 25 <213> artificial

<220>  
 <223> forward primer exon 9\_3  
 30 <400> 37  
 cataaaacgac aggctgctga 20

35 <210> 38  
 <211> 37  
 <212> DNA  
 <213> artificial

40 <220>  
 <223> reverse primer exon 9\_3  
 <400> 38  
 gtaaaaacgac ggccagtagc aggcactttg tggagaa 37

45 <210> 39  
 <211> 17  
 <212> DNA  
 <213> artificial

50 <220>  
 <223> sequence primer exon 9\_3  
 <400> 39  
 55 gtaaaaacgac ggccagt 17

<210> 40  
 <211> 38  
 <212> DNA  
 60 <213> artificial

<220>  
 <223> forward primer exon 9\_4  
 65 <400> 40  
 gtaaaaacgac ggccagtcca caccttcttt ctccacaa 38

70 <210> 41  
 <211> 20

	<212> DNA	
	<213> artificial	
5	<220> <223> reverse primer exon 9_4	
	<400> 41 caattcccaca tgtctgctgt	20
10	<210> 42 <211> 17 <212> DNA <213> artificial	
15	<220> <223> sequence primer exon 9_4	
20	<400> 42 gtaaaaacgac ggccagt	17
25	<210> 43 <211> 37 <212> DNA <213> artificial	
30	<220> <223> forward primer exon 9_5	
	<400> 43 gtaaaaacgac ggccagtcaa gaaacttgcc gtcttcc	37
35	<210> 44 <211> 26 <212> DNA <213> artificial	
40	<220> <223> reverse primer exon 9_5	
45	<400> 44 caaattatta aatgtcatca ccagga	26
50	<210> 45 <211> 17 <212> DNA <213> artificial	
55	<220> <223> sequence primer exon 9_5	
	<400> 45 gtaaaaacgac ggccagt	17
60	<210> 46 <211> 3111 <212> DNA <213> homo sapiens	
65	<400> 46 atggcctttgc gggggccgc gggagcacc gacaccocgg tgtcctcggc cgggggagcc	50
	ccggcgggc cagcgtcctc gtctccacc tctcggggc gctcggcctc gggggcgcg	120
	ggcctgtggg ccgcgctcta tgactacgag cctcggcggc aggacgagct gagcctcgg	180
70	ccggcccaagc tgggtggaggt gctgtcgcag gacgcgcgcg tctcggggcga ccggggctgg	240

	tgggcaggcc aggtgcagcg ggcctcggc atcttcccgg ccaactacgt ggtccctgc	300
5	cgcccgcccg ccagccccgc gccgcggccc tcggggccca gctccccggt acacgtcgcc	360
	ttcagagcgg tcgaactgaa gtagctcctc cgcctcggg gcttcgggca ggtgtacgc	420
	gccacctggc agggccaggc ggtggccgtg aaggcggcgc gccaggacc ggagcaggac	480
10	gcggcgcccg ctgcagagag cgtgcggcgc caggctcggc tcttcgccat gctgcggcgc	540
	ccccacatca tcgagctggc cggcgtgtgc ctgcagcgc cgcacctctg cctgggtctg	600
15	gagttcgccc gcggcgagc gctcaaccga cgcctcggcg ctgccaccgc cgcgcccgac	660
	cggcgccggc cgggcccccg ccggcgcggc cgcacctc cgcacctgct ggtcaactgg	720
	gccgtgcaga tagcggggg catgctctac ctgcatgagg aggccttctg gcccatctg	780
20	caccgggacc tcaagtcag caacatcttg ctacttgaga agatagaaca tgatgacatc	840
	tccaataaaa ctttgaagat tacacatctt cggttcggca ggaatgca caggaccacc	900
25	aaaatgagca cagcaggcac ctatgcctgg atggccccg aagtgatcaa gtcttccttg	960
	ttttctaagg gaagcgacat ctggagctat cggagtctgc tgtgggsact gctcacggca	1020
	gaagtcccc atcggggcat tgatggctc cgcgtggctt atggggtagc agtcaataaa	1080
30	ctcactttgc ccattccatc caactgcct cagccgtttg ccaagctcat gaagaaatgc	1140
	tggcaacaag accctcatat tcgtccatcg ttgccttaa tctcgaaca gttgactgct	1200
35	attgaagggg cagtgatgac tgagatgct caagaatctt ttcattocat gcaagatgac	1260
	tggaaactag aaattcaaca aatgtttgat cagttgagaa caaaggaaaa ggagctgca	1320
	tcccggaag aggagctgac tcggcggtc ctgcagcaga agtctcagga ggagctgcta	1380
40	aagcggcgtg agcagcagct ggcagcgc cagatcagc tgcggagcg ggaactaac	1440
	attctgata tccagctaaa ccagcagaag ccaaggtaa agaaggaa gggcaagttt	1500
45	aagagaagc gtttaagct caaagatga catcgaatca gttaccctc agatctccag	1560
	cacaagataa ccgtgcagc ctctcccaac ttggacaaac ggcggagcct gaacagcagc	1620
	agttccagtc ccccgagcag ccccaaatg atgccccgc tccgagccat acagttgact	1680
50	tcagatgaaa gcaataaaac tggggaagg aacacagtct ttcgacaaga agaatttgag	1740
	gatgtaaaa ggaattttaa gaaaaaagg tgtacctgg gaccaaatc cattcaaatg	1800
55	aaagatagaa cagattgcaa agaaaggata agacctctc ccgatggcaa cagtccttg	1860
	tcaactatc taataaaaa tcagaaaacc atgcccctgg cttcattggt tgtggaccag	1920
	ccagggtcc ctgaaagagc aaaactttc cctgatggat tagaacacag aaaccacaaa	1980
60	caataaaaa tgcctagtca ggcctacatt gatctacct tgggaaaga tgcctcagaga	2040
	gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgcacagtc	2100
65	tcatttgaga tgactcctac gaatagtctg agtagatcc ccagagaaa gaaaccggag	2160
	tcagctcctg atggg-gcac cgtcctcctg ccatcggcgg ctctgggacl ggaccctaga	2220
	gagcttcata aagccaggc tgcctgaagaa ccgttgcca aggaagagaa gaagaaaccg	2280
70	gagggaalc tccagcgggc tccaagctc cgcagaagc ccagctcc ccaagcctg	2340

	ccatccacct gtggggagge cagcagocca cctccctgc cactgtcaag tgcctcggc	2400
	atcctctcca caccctcttt cccacaaag tgcctgctgc agatggacag tgaagatcca	2450
5	ctgggtggaca gtgcacctgt cacttgtgac tctgagatgc tcaactcogga tttttgtccc	2520
	actgccccag gaagtggctg tgagccagcc ctcatgccaa gacttgacac tgattgtagt	2580
10	gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt accttactgt	2640
	gcttcttcaa aacatagacc gccacatcac agacggacca tgtctgatgg aatccgacc	2700
	ccaactggtg caactattat cccagccact ggagcctctg cactgccact ctgccctca	2750
15	ccgtctcttc acagtcctct gccaaaggag gtctcaccsa agaagccag cactgtccac	2820
	alcglycclc agcglcgccl tgcclccclg agaagccgcl cagalclgcl ccaggcllac	2880
20	ccacagacag cagtgtctca gctggcacag actgctctg tagtgggtcg cccaggaca	2940
	catccacccc aattcctcgc tgccaaggag agaactaaat cccatgtgcc ttcattactg	3000
	gatgctgacg tggaaaggta gagcagggac tacaactgtgc cactgtgcag aatgaggagc	3050
25	aaaaccagcc ggccactat atatgaactg gagaaagaat cctgtctta a	3111
	<210> 47	
30	<211> 41	
	<212> DNA	
	<213> artificial	
	<220>	
35	<223> mutagenesis primer G291E	
	<400> 47	
	ctttgaagaa tacagatctt gagttggcga gggaaatggca c	41
40	<210> 48	
	<211> 40	
	<212> DNA	
	<213> artificial	
45	<220>	
	<223> mutagenesis primer R470C	
	<400> 48	
50	tgagcagcag ctggcagagt gcgagatcga cgtgctggag	40
	<210> 49	
	<211> 304*	
55	<212> DNA	
	<213> homo sapiens	
	<400> 49	
	atggctttgc ggggcgcgcg gggagcagcc gacaccccgg tgcctcggc cgggggagcc	50
60	cccggcggcc cagcgcctc gtcgtccacc tctcggggcg gctcggcctc ggcgggagcg	120
	gggctgtggg ccgcgctcta tgactacgag gctcggggcg agaacagact gagcctcggg	180
	ccggccagca tgggtggaggt gctgtcgcag gacgcgcgcg tgcggggcga ccagggctgg	240
65	lgggcaggcc agglccagcg ccgcclggc alclccccg ccaactacgl ggcclccclg	300
	cgcccgcccg ccagccccg gcgcgcgcgc tccgggccca gctccccggt acacgtcggc	350
70	llcgagcggc lggagclgaa ggagclcalc ggcgclggg gcllccggca gglglaccg	420

gocacctggc agggocagga ggtggccgtg aaggcggcgc gccaggacc gccagcaggac 480  
 gggcggggcg ctgocgagag cgtgogggcg gaggctcggc tcttcgccat gctgcccgcac 540  
 5 cccaacatca togagctgcg cggcgtgtgc ctgcagcagc cgcacctctg cctgggtctg 600  
 gagttcgccc gggcgggagc gctcaaccga cgcctggcgg ctgccaccgc cgcgccggac 660  
 10 cggcggcgcg cgggcccccg ccgcccggcg cgcctccctc cgcacgtgct ggtcaactgg 720  
 gocgtgcaga tagcggggg catgctctac ctgcatgagg aggccttcgt gccatcctg 780  
 caccgggacc tcaagccag caacattttg ctacttgaga agatagsaca tgatgacatc 840  
 15 tgcataaca ctttgaagat tacagatttt gaggttggcg gggatggca caggaccacc 900  
 aaaalgagca cagcaggcac ctalccclgg alggccccg aalgalcaa glcllclllg 960  
 20 ttttctaagg gaagcgacat ctggagctat gaggctgtgc tgtgggact gctcacggga 1020  
 gaagtcccc atcggggcat tgatggctc gccgtggctt atgggtagc agtcaataaa 1080  
 ctacttttg ccattccatc cactgcccct gaggcgtttg ccaagctcat gaagaaatgc 1140  
 25 tggcaacaag acctcctat tctccatcg tttgcttaa ctctgacaa gttgactgct 1200  
 atgaagggg cagtgatgac tgagatgct caagaattt ctattccat gcaagatgac 1260  
 30 lggaaactag aaallcaaca salglllgal gagllgagaa caaaggaaa gggcctcgcg 1320  
 tcccggaag agggagctgac tggcgggct ctgcagcaga agtctcagga gggcctgcta 1380  
 aagcggcglg agcagcagcl ggcaagcgc gagalogacg lglggagcg ggaacllaac 1440  
 35 attctgatac tccagctaaa ccaggagag cccaaggtaa agaagaggaa gggcaagttt 1500  
 aagagaagtc gtttaagct caaagatgga catcgaatca gtttaccctc agatttccag 1560  
 40 cacaagatac ccgtgcaggc ctctcccac ttggacaaac gggggagcct gaccagcagc 1620  
 agttccagtc ccccgagcag ccccaaatg atgcccgcac tccgagccat acagttgact 1680  
 toagatgaaa gcaataaaa tggggagag aacacagctc ctgacacaga agaatttgag 1740  
 45 gatgtaaaa ggaattttaa gaaaaaggt tgtacctggg gaccaaatc cattcaaatg 1800  
 aaagatagaa cagattgcaa sqaaaaggata agacctctc ccgatggcaa caqtcttgg 1860  
 50 tcaactatcc taataaaaaa tcagaaaaac atgcccctgg ctctattggt tgggagccag 1920  
 ccagggtccg gtgacagagc aaaaatttcc cctgatgat tagaacccag aaacccaaa 1980  
 caaataaaa tgcctagtca ggctacatt gatctaccctc ttgggaaaga tgcctagaga 2040  
 55 gagaatcctg cagaagctga aagctgggag caggcagcct ctgcgaatgc tcccacagtc 2100  
 tccattgaga tgactcctac gaatagtctg agtagatcc ccagagaaa gaaacggag 2160  
 60 tcaqctctgc atgggagcag cgtccttctg ccacccgtg ctctgggact ggaacctcag 2220  
 gagcttcata aagcaacaggc tgcgaagaa ccgttgcca aggaagagaa gaagaaacga 2280  
 gagggaatcc tccagcggc tcccaagtc ccagaaagcg ccagtcctcc cacaagcctg 2340  
 65 ccctccacc gtggggaggc cagcagccca cctccctgc cactgtcaag tgcctcggc 2400  
 atcctctcca cactctctt ctccacaag tgcctgctgc agatggcag tgaagatca 2460  
 70 ctgggtggca gtgcacctg cactgtgac tctgagatgc tcaaccogga tttttgtccc 2520  
 actgccccag gaagtggctg tgagccagcc ctcatgcaa gacttgacac tgattgtagt 2580

	gtatcaagca acttgccgtc ttccttecta cagcagacat gtgggactgt accttactgt	2640
5	gcttcttcaa aacatagacc gtcacatcac agacggacca tgtctgatgg aatccgacc	2700
	coaactggtg caactattat ctcaaccact cgaacctctg cactgccact ctgccoctca	2760
	ccctgctcctc acagtcactc gccaaaggag gctcacccca agaagcacag cactgtccac	2820
10	atcgtgcctc agcgtcgcct tcctccctc agaaqccctc cagatctgcc tcaggcttac	2880
	ccacagacag cagtgtctca gctggcacag actgctctgt tagtgggtcg ccaggacca	2940
15	catcccacc aattcctcgc tgcacaagg agaaactaaat cccatgtgcc ttcattactg	3000
	gatgctgacg tggaaaggtca gagcagggac tacactgtgc c	3041
20	<210> 50 <211> 3041 <212> DNA <213> homo sapiens	
25	<400> 50 atggctttgc gggggccgc gggagcgacc gacaccccgg tgcctcggc cgggggagcc	60
	cccggggct cagcgtcctc gtcgtccacc tctcggggc gctcggcctc gggggggcg	120
30	ggcclglgg ccgcgclcla lgaclacgag cclcgcgcg aggacggcl gagcclcgcg	180
	cccgccacg tgggtggagt gctgtcgcag gacgcgcgc tctcggggca cgggggtgg	240
	lgggcaggcc agglgcagcg gcgcclcgcc alcllcccg ccaaclacgl ggcclccclgc	300
35	cgcccccg ccagccccg gcgcgcgcc tcggggcca gctccccgt acccgtcgc	360
	ttcggcggc tggagctgaa ggagctcctc ggcgtgggg gcttcgggca ggtgtaccgc	420
40	gccacctggc agggccaggc ggtggccgtg aaggcggcgc gccaggccc ggcgcaggac	480
	ggggggcg ctcgcagagc cgtgcggcgc gaggtcggc tcttcgccat gctgcggcac	540
	cccaacatca tcgagctcgc cggcgtgtgc ctgcagcag ccacacctg cctggtgtg	600
45	gagttcggc gcgggggagc gctcaaccga gcgctggcgc ctgccaccgc cgcctccgac	660
	ccgcgcgcgc ccggcccccg ccgcgcgcgc ccgcatcctc ccacagtct ggtcaactgg	720
50	gcgctgcaga tagcggggg catgctctac ctgcctgagg aggccttcgt gcccatcctg	780
	ccccgggacc tcaagcccag caacatcttg ctacttgaga agatagsaca tgatgacatc	840
	tgcataaaa ctttgaagat tacagatctt cggttggcga gggaatggca caggaccacc	900
55	aaaatgagca cagcaggcac ctatcctcgc atggcccccg aagtqatcaa gctctccttg	960
	ttttctaagg gaagcgacat ctggagctat cggagtctgc tctgggact gctcacggc	1020
60	gaaqtcccc atcggggcat tgatgcctc ccctgqctt atgggqtagc agtcaataaa	1080
	ctcactttgc ccattccatc caoctgcctc gaggcgtttg ccaagctcat gaagaatgc	1140
	tgcacaacag accctcatat tcgtccatcg tttcccttaa ctctcgsaca gttgactct	1200
65	atggaagggc cagtgatgac tgagatgctc caagaacttt ctcatccat gcaagatgac	1260
	tggaaactag aaattcaaca aatgtttgat gagttgagaa caaaggaaaa ggagctcga	1320
70	tcccggaag aggagctgac tcggcggct ctgcagcaga agtctcagga ggagctccta	1380
	aagcggcgtg agcagcagct ggcagagtgc gagatcgac tctgggagcg ggaacttaac	1440

attctgatac tccagctaaa ccaggagcag cccaaggtaa agaagaggaa gggcaagttt 1500  
 5 aagagaagtc gtttaagct caaagatgga catcgaatca gtttaccctc agatttccag 1560  
 cacaagataa ccgtccagc ctctcccaac ttggacaaac ggcggagcct qaacaqcagc 1620  
 agttccagtc ccccgagcag ccccaaatg atgcccgcac tccgagccat acsgttgact 1680  
 10 tcagatgaaa qcaataaaaac ttggcgasgg aacacagctct ttcgacsaga aqzatttqag 1740  
 gangtaaaaa ggaattttta gaaaaaaggf tgtacctggg gaccaaaatc cattcaaatg 1800  
 15 aaagatagaa caqatcgcaa agaaaaggata agacctctct ccgatggcaa cagtccctgg 1860  
 tcaactatct taataaaaaa tcagaaaaac atgcccctgg ctccattggt tgtggaccag 1920  
 ccagggtccz gtgaagagcc aaaaactttcc cctgatggat tagaacacag aaaccacaaa 1980  
 20 caaataaaaa tgcctagtea ggctacatt gatctacctc ttgggaaga tgcctagaga 2040  
 gagaatcctq cagaagctga aagctgggag caqgcagcct ctccgaatgc tcccacagtc 2100  
 25 tccattgaga tgactcctac gaatagtctg agtagatccc ccagagaaa gaaaccggag 2160  
 tcagctctgz atgggtgcac cgtccctctg gcacgggtg ctctgggact ggaacctaga 2220  
 gagcttcata aagcacaggc tgcctgaagaa ccgttgccc aggaagagaa gaagaaacga 2280  
 30 gagggaatct tccagcgggc tccaagtc ccagagaagcg ccagtcctcc cacaagcctg 2340  
 ccaccaccz gtggggaggc cagcagccca cctccctgc cactgtcaag tgcctcggc 2400  
 35 atcctctcca cacctcttt cccacaaag tgcctgctgc agatggacag tgaagatoca 2460  
 ctggggaca gtgcacctgt cacttgtagc tctgagatgc tccctcggga tttttgtccc 2520  
 accgccccag gaagtgtog tgagccagcc ctcatgcca gacttgacac tgattgtagt 2580  
 40 gtatcaagaa acttgccgtc tccctccta cagcagacat gtgggaatgt acctactgt 2640  
 gcttcttcaa aacatagacc gccacatcac agacggacca tgtctgatgg aatccgacc 2700  
 45 ccaactggtg caactattat cccagccact ggagcctctg cactgccaact ctgcccotca 2760  
 ccctgctctc acagtcact gccaaggag gctcaccca agaagcacag cactgtccac 2820  
 atcgtgcctc agcgtgcctc tgcctcctg agaagccct cagatctgcc tccggttac 2880  
 50 ccacagacag cagtgtctca gctgcccag actgctctg tagtgggtcg cccagacca 2940  
 catcccacc aattctctgc tgccaaggag agaactaaat cccatgtgcc ttcattactg 3000  
 55 gatgctgacg tggaaggtea gagcagggac tacaactgtgc c 3041

<210> 51  
 <211> 3041  
 <212> DNA  
 60 <213> homo sapiens

<400> 51  
 atggctttgc ggggcgccgc gggaccgacc cacaccccgg tgcctcggc cgggggagcc 60  
 65 cccggggcct cagcgtctct gctgtccacc tccctggggc gctcggcctc gggggcggc 120  
 gggctgtggg ccgctctcta tgactacgag gctcggggc aggaacgagct gagcctcggg 180  
 70 ccgggccagc tgggtggaggt gctgtcgcag gacgcccgc tgcggggcga ccagggtgg 240  
 tgggcaggcc aggtgcagcg ggcctcggc atcttcccgc ccaactacgt ggtccctgc 300

	cgcccgcccg ccagcccccg gccgcggccc tcgcggccca gctccccggt acccgctggc	360
5	ttcagagcggc tggagctgaa ggagctcctc ggcgctgggg gcttcgggca ggtgtaaccg	420
	gccacctggc agggccagga agtgcacctg aagccggcgc gccagacc ccagccagac	480
	ggggcgccgg ctgccagagc cgtgcggcgc gaggctcggc tcttcgccat gctgcggcac	540
10	cccaacatca tcgaactgag cggcctgtgc ctgcagcagc ccacctctg cctggtctctg	600
	gagttcgccc gggcgggagc gctcaaccga ggcctggcgc ctgccaccgc cgcgccggac	660
15	ccgcgcgggc ccggcccccg ccgcgcgggc ccgctccctc ccgacctgct ggtcaaccgg	720
	gcccgtcaga tagccggggc catgctctac ctgcatgagg aggccttcgt gcccatcctg	780
	taccgggacc tcaagtcagc caacattttg ctacttgaga agatagsaca tgatgacatc	840
20	tgcataaaaa ctttgaagat tacagatttt gggttggcga gggaatggca caggaccacc	900
	aaaatgagca cagcagccac ctatccctgg atgcccccg aagtgatcaa gtcttccttg	960
25	ttttctaagg gaagcgacat ctggagctat ggagtgcctc tgtgggaact gctcacggga	1020
	gaagtccccg atcggggcat tgatggcctc gccgtggcct atggggtagc agtcaataaa	1080
	ctcactttgc ccattccatc cacctgccct gagccgtttg ccaagctcat gaaagaatgc	1140
30	tggcaacaag accctcatat tcgtccatcg tttgccttaa tctctgsaca gttgactgct	1200
	attgaagggg cagtgatgac tgagatgcct caagaatctt tcattccat gccagatgac	1260
35	tggaaaactg aaattcaaca aatgtttgat gaggttgaga caaaggaaaa ggagctgoga	1320
	tcccgggagc aggagctgac tcggcggcct ctgcagcaga agtctcagga ggagctgcta	1380
	aagcggcgtg agcagcagct ggcagagcgc gagatcgacg tgcctggagcg ggaacttaac	1440
40	attctgatac tcagcctaaa ccagcagcag cccaaggtaa agaagaggaa gggcaagttt	1500
	aagagaagtc gtttaagct caaacatgga catcgaaatc gtttacctc agatttcagc	1560
45	cacaagataa ccgtgcaggc cctcccccac ttggacaaac ggcggagcct gaccagcagc	1620
	agttccagtc ccccgagcag ccccaaatg atgccccgac tccgagccat acagttgact	1680
	tcagatgaaa gcaataaaac ttggggaagg aacacagctc ttgacacaga agaatttgag	1740
50	gatgtaaaaa ggaattttta gaaaaaaggc tgtacctggg gaccacatc cattcaaatg	1800
	aaagatagaa cagatcgcaa agaaaggata agacctctct ccgatggcaa cagtccctgg	1860
55	tcaactatcc taataaaaaa tcagaaaaac atgcccctgg cttcattggt tgtggaccag	1920
	ccagggtccc gtgaagagcc aaaactttcc cctgatggat tagaacscag aaaaacaaa	1980
	caataaaaac tgcctagtca ggcctacatt gatctacctc ttgggaasaga tgctcagaga	2040
60	gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgccacagtc	2100
	tcatttgaga tgactcctac gaatagtctg agtagatccc ccagagaaa gaaaaccggag	2160
65	tcagctctgt atgggtgcac cgtccttctg gcctcgggtg ctctgggact ggaactcaga	2220
	gagcllcala aagccacagg cglclgaagaa ccgllgcca aggaagagaa gaagaaacga	2280
	gagggaatcc tcacagcggc tccaagtcc ccgagaagcg ccagcccccc caaagcctg	2340
70	ccalccaccg glggggaggc cagcagccca ccccccctgc caclglcaag lgccclyggc	2400

atcctctcca cacctctttt cccacaaag tgctgtctgc agatggacag tgaagatcca 2450  
 ctgggtggaca gtgcacctgt cacttgtgac totgagatgc cactccgga tttttgtccc 2520  
 5 actgccccag gaagtggctg tgagccagcc ctcatgccaa gacttgacac tgattgtagt 2580  
 gtatcaagaa acttgccgtc ttccttcta cagcagacat gtgggaagt accttaactgt 2640  
 10 gcttcttcaa aacatagacc gtcacatcac agacggacca tgtctgctgg aatccgacc 2700  
 ccaactggtg caactattat cccagccact ggagcctctg cactgccact ctgccctca 2750  
 cctgctccto acagtcactt gccaaaggag gctcaccca agaagcacag cactgtccac 2820  
 15 atcgtgccto agcgtgcgcc tgctcctctg agaagccgt cagatctgcc tcaggcttac 2880  
 ccacagacag caglgllclca gclggcacag actgclglg laglggg\_lcg ccaggacca 2940  
 20 catccaccoc aattcctcgc tgccaaggag agaactaaat cccatgtgcc ttcattaactg 3000  
 gatgctgacg tggaaaggta gagcagggac tacactgtgc c 3041  
  
 <210> 52  
 25 <211> 3041  
 <212> DNA  
 <213> homo sapiens  
  
 <400> 52  
 30 atggcttttg gggggccgc gggagccac gacaccccgg tgcctcggc cgggggagcc 60  
 cccggggcc cagcgcctc gctgtccac tctcgggag gctcggctc ggcggggcg 120  
 35 gggctgtggg ccgctctcta tgactacgag gctcgggag aggaacagct gagcctcggg 180  
 ccggccacg tgggtggagt gctgtcgcag gacgcccgc tgtcgggca caggggctgg 240  
 tgggcaggcc aggtgcagcg ggcctcggc atcttcccg ccaactacgt ggtccctgc 300  
 40 ccccccggc ccagccccg gccgcgcc tgcggccca gctcccggg acacgtcgc 360  
 ttcagagcgc tggagctgaa ggagctcctc ggcgtgggg gcttcgggca ggtgtacgc 420  
 45 gccacctgg agggccagga ggtgcctgt aaggcgggc gccaggacc ggagcaggac 480  
 gggcggcg ctgcagagag cgtgcgggc gaggtcggc tcttcgccat gctgcggcac 540  
 cccaacatca tcagctcgc cggcgtgtc ctgcagcag ccacactctg cctggtctg 600  
 50 gagttcggc cgggggagc gctcaaccga gcgctggcg ctgccacgc cggcccggac 660  
 ccgcccgc ccggccccg ccgcccgc cgcctcctc ccacgtgct ggtcaactgg 720  
 55 ggcgtgcaga tagcggggg catgctctac ctgcatgag aggccttctg gccatcctg 780  
 cagcgggacc tcaagccag caacattttg ctacttgaga agatagsaca tgatgacatc 840  
 tcgaataaaa ctttgaagat tacagatttt cggttggcga ggaatggca caggaccacc 900  
 60 aaaatgagca cagcaggcac ctatgcctgg atggccccg aagtgatcaa gcttctctg 960  
 tttctaaag gaagagacat cggagctat ggagcctgc tggggact gctcacogga 1020  
 65 gaagtccc atcggggcat tgatgcctc gccgtggtt atgggtagc agtcaataaa 1080  
 ctcaclllgc ccallccalc cccclgccc gagcgglllg ccaagclcal gaaagaalgc 1140  
 tggcaacaag accctcatat tcgtccatcg ttgccttaa ctctcgsaca gttgactgct 1200  
 70 allgaagggg caglyalgac lgagalgcc caagaalcll llcallecal gcaagalgac 1260

	tggaaactag aaattcaaca aatgtttgat gagttgagaa caaaggaaaa ggagctgoga	1320
	tcccggaag aggagctgac tccggcggct ctgcagcaga agtctcagga ggagctgcta	1380
5	aagcggcgtg agcagcagct ggcagagcgc gagaacgacg tgcctggagcg ggaacttaac	1440
	attctgatac tccagctaaa ccaggagaag cccaaggtaa agaagaggaa gggcaagttt	1500
10	aagagaagct gtttaagct caaagatgga catcgaatca gtttaccttc agatttccag	1560
	cacaagataa ccgtgcaggc ctctcccaac ttggacaaac ggcggagcct gaacagcagc	1620
	agttccagtc ccccgagcag ccccaaatg atgcccagc tccgagccat acagttgact	1680
15	tcagatgaca gcaataaaac ttggggacgg aacacagtct ctgcacaga agcatttgag	1740
	gaglaaaaa ggaalullaa gaaaaaaggl lglaccclgg gaccaaalcc callcaaalg	1800
20	aaagatagaa cagatcgcaa agaaaggata agacctctct ccgatggcaa cagtccttgg	1860
	tcaactatct taataaaaaa tcagaaaaac atgcccctgg cttcattggt tgtggaaccag	1920
	ccagggtccc gtgaagagcc aaaaactttc cctgatggat tagaacacag aaacccaaaa	1980
25	caataaaaaa tgcttagtca ggctacatt gatctacctc ttgggaaaga tgcctcagaga	2040
	gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgccacagtc	2100
30	lccallgaga lgacclclac gaalagclcg aglagalccc ccagagaaa gazaaccgag	2160
	tcagctctgc atgggtgcac cgtcctctc gcatcggtag ctctgggact ggcctcaga	2220
	gagcllcala aagcacaggc lgcgaagaa ccgllgcaca aggaagsgaa gaagaaacga	2280
35	gagggaatcc tccagcgggc tccaagtc ccagagaagcg ccagtcctcc cacaagcctg	2340
	ccatccacct gtggggaggc cagcagccca cctccctgc cactgtcaag tgccctgggc	2400
40	atcctctcca cactctctt ctccacacag tgctctctgc agatggacag tgaagatccc	2460
	ctgggtgaca gtgcacctgt cacttctgac tctgagatgc tcactccgga tttttgtccc	2520
	actgcccacg gaagtggctg tgagccagcc ctcatgcca gacttgscac tgattgtagt	2580
45	gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt accttactgt	2640
	gcttcttcaa aacataaac gtcacatcac agacggacca tctctgctgg aatcccacc	2700
50	ccaactgggt caactattat ctccagccct ggagcctctg cactgccaact ctgccctca	2760
	ccctgctctc acagtcctct gccaaaggag gtctcaccca agaagccag cactgtccac	2820
	atcgtgcctc agcgtgcgcc tgctccctg agaagccct cagatctgcc tcaggcttac	2880
55	ccacagacag cagtctctca gctgcacacg actgcctctg tagtgggtcg ccacagacca	2940
	caccccacc aattctctgc tgccaaggag agaactaaat cccatgtgcc ttcattactg	3000
60	gatctgacg tggaaagtca ggcagggac tacactgtgc c	3041
	<210> 53	
	<211> 3041	
	<212> DNA	
65	<213> homo sapiens	
	<400> 53	
	atggctttgc gggggcggc gggagcagc gacaccccgg tgcctcggc cgggggagcc	50
70	cccgccggc cagcgcctc g_cglccacc lccclgggc gclgggcclc ggcggggcgc	120

gggctgtggg ccgocctcta tgactacgag gctcggcgcg aggaocgagct gagcctcggg 180  
 ccggccagc tggtggaggt gctgtcgcag gacgcccgcg tgcggggcga cgaaggctgg 240  
 5 tgggcaggcc aggtgcagcg ggcctcggc atcttcccgc ccaactacgt ggtccctgc 300  
 cgcggcgccg ccagccccgc gcgcgcgcc tgcgggcccga gctcccgggt acacgtcggc 360  
 10 ttcagcggg tggagctgaa ggagctcctc ggcgctgggg gcttcgggca ggtgtaccgc 420  
 gccacctggc agggccagga ggtgcocgtg aaggcggcgc gccaggacc gccagcggac 480  
 gggggggcg ctgcagagag cgtgcggcgc gaggctcggc tcttcgccat gctgcggcac 540  
 15 cccaacatca tcgagctgcg cggcgtgtgc ctgcagcagc ccacacctg cctgggtctg 600  
 gagllcggcc gggggggagc gclcaaccga cgcclggcgc clgcacacgc cgcggcgac 660  
 20 ccgocgcgc ccggccccgc ccgocgcgc cgcctccctc ccacagtgct ggtcaactgg 720  
 gccgtgcaga tagcggggg catgctctac ctgcatgagg aggccttcgt gcccatcctg 780  
 caccgggacc tcaagccag caacattttg ctacttgaga agatagsaca tgatgacatc 840  
 25 tgcaataaaa ctttgaagat tacagatttt gggttggaga ggaatggca caggaccacc 900  
 aaaatgagca cagcaggcac ctatgcctgg atggccccg aagtgatcaa gtcttcttg 960  
 30 lllcltaag gaagcgacal clggagclal ggaglgclgc glggggacl gclcacggg 1020  
 gaagtcccc atcggggcat tgatgcctc gccgtggctt atggggtagc agtcaataaa 1080  
 clcaclllgc ccclccalc cccclgccl gagcgglllg ccaagclcal gaaagaa gc 1140  
 35 tggcaacacg accctcatat tcgtccatcg tttgccttaa tctcgcaca gttgactgct 1200  
 attgaaaggg cagtgatgac tgagatgcct caagaatctt ttcattccat gcaagatgac 1260  
 40 tggaaactag aaattcaaca catgtttgat gagttagaa caaaggcaaa ggagctcgc 1320  
 tcccggggag aggagctgac tggggcggt ctgcagcaga agtctcagga ggagctcctc 1380  
 aagcggcgtg agcagcagct ggcagagcgc gagatcgaag tgcgggagc ggaactaac 1440  
 45 atctcgatac tccagctaaa ccaggagag ccaaggtaa agaagaggaa gggcaagttt 1500  
 aaggaagtc gtttaagct caaacatgga catcgaatca gtttaccctc agatctccag 1560  
 50 cacaagataa ccgtgcaggc ctctcccac ttggacaaac ggcggagcct gaacagcagc 1620  
 agttccagtc ccccgagcag cccacactg atgccccgc tccgagccat acagttgact 1680  
 tcagatgaaa gcaataaaac tgggggaggg aacacagctc ttcgacaaga agaatttgag 1740  
 55 gatgtaaaaa ggaattttaa gaaaaaaggt tctacctcgg gaccaasctc cattcaaatg 1800  
 aaagatagaa cagattgcaa agaaaggata agacctctct ccgatggcaa cagtccttgg 1860  
 60 tcaactatct taataaaaaa tcagaaaaac atgccttgg ctctattggt tctggaaccg 1920  
 ccagggtccc gtgaagagcc aaaactttcc cctgatggat tagaacacag aaaaccaaaa 1980  
 caaataaaaa tgcctagtca ggcctacatt catctacctc tgggagaga tctcagaga 2040  
 65 gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgcacagtc 2100  
 tcatttgaga tgactcctac gaatagtctg agtagatcc ccagagaaa gaaaccggag 2160  
 70 tcagctctgc atgggtgcac cgtccttctg ccatcgggtg ctctgggact ggaacctcaga 2220  
 gagcttcata aagcaccaggc tgcgtaagaa ccgttgcccga aggaagagaa gaagaaacga 2280

gagggaaatc tccagcgggc tccaagtc ccgagaagcg ccagtcctcc cacaagcctg 2340  
 ccateccacc gtggggaggc cagcagccca cctccctgc cactgtcaag tgcctcgggc 2400  
 5 atcctctcca cactctctt cccacasaq tccctgctgc agatqgscag tgaagatcca 2460  
 ctgggtggaca gtgcacctgt cacttgtgac tctgagatgc tcactccgga tttttgtccc 2520  
 10 actgccccag aaagtqgtc tgaqccagcc ctcatgccaa gacttqscac tqtattgagt 2580  
 gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt accttactgt 2640  
 15 gcttcttcaa aacatagacc tccacatcac agacggacca tqtctgatgg aatcccacc 2700  
 ccaactggtg caactattat ctccagccact ggagcctctg cactgccact ctgccctca 2760  
 ccctgctctc acagtcctct gcccaaggag gtctcaccca agaagcacag cactgtccac 2820  
 20 atcgtgcctc agcgtgcctc tgcctccctg agaagccctc cagatctgcc tcaggcttac 2880  
 ccacagacag cagtctctca cctgcacac actgctctct tagtggctcg cccagacca 2940  
 25 catcccacc aattcctcgc tgcgaaggag agaactaaat cccatgtgcc ttcattactg 3000  
 gatgctgacg tggaaaggtc gagcagggac tcaactgtgc c 3041  
  
 <210> 54  
 <211> 3041  
 <212> DNA  
 <213> homo sapiens  
  
 <400> 54  
 35 atggccttgc gggggcggc gggagcggc gacaccccgg tgcctcggc cgggggagcc 60  
 cccggcggc cagcgcctc gtctccacc tctcggggc gctcggcctc ggcggggcgg 120  
 40 gggctgtggg ccgctctcta tgactacgag cctcggggc aggaocgct gagcctcggg 180  
 ccggccagc tgggtggagt gctgtgcag gacgcggcg tgtggggcga cggggcctgg 240  
 tgggcaggcc aggtgcagc ggcctcggc atcttcccg ccaactcgt ggctcctgc 300  
 45 cgcggggcgg ccagcccccg gcgcggccc tgcgggcca gctcccggg acscgtcggc 360  
 ttccagcggc tggagctgaa qgagctcctc cgcctcggg ccttcgggca ggtgtaaccg 420  
 50 gccacctggc agggccagga ggtggcctg aaggcgggc gccaggacc ggagcaggac 480  
 gggcggcgg ctcgcagag cgtgggggc gaggtcggc tcttcggcat gctggggcgc 540  
 cccaacatca tccagctgcg cggcgtgtg ctgcagcagc cgcacctctg cctgggtctg 600  
 55 gagttcggc ccggcggag cctcaaccga ccctcggcg ctgccaccg cggccccgac 660  
 ccgcgcggc ccggccccg ccgcgcggc cgcctcctc cgcacgtgt ggtcaactgg 720  
 60 gccgtgcaqa tagcgcggg catgctctac ctgcatgag aggccttctg gcccatcctg 780  
 caccgggacc tcaagccag caacattttg ctacttgaga agatagaaca tcatgacatc 840  
 tccaataaaa ctttgaagat tacacatttt cggttgcca qggaatgaca caqgaccacc 900  
 65 aaaatgagca cagcaggcac ctatccttg atggccccg aagtgatcaa gtcttctctg 960  
 ttttctaagg gaagcgacat ctggagctat caggtctgc tgtgggaact gctcacggga 1020  
 70 gaagtcccc atcggggcat tgatgcctc gccgtggctt atgggtagc agtcaataaa 1080  
 ctactttgc ccattccatc cactgcctc gagcgtttg ccaagctcat gasagaatgc 1140

5            tggcaacaag accctcatat tegtccatcg tttgccttaa ctctcgsaca gttgactgct       1200  
           atggaagggg cagtgatgac tgagatgcct caagaatctt ttcaatccat gcaagatgac       1260  
           tcgaaaactag aaattcaaca aatgtttgat caqttgacaa caaaggsaaa ggaqctcoga       1320  
 10           tcccgggaag aggagctgac tcggccggct ctgcagcaga agtctcagga ggagctgcta       1380  
           aaqcqccgtg agcaqcaqct gqcaeqccq cagaacqacg tcgtqgsqcg ggaacttaac       1440  
           atcttgataa tccagctaaa ccaggagaag cccaaggtaa agaagaggaa gggcaagttt       1500  
 15           aagagaagtc qtttaaaqct caaacatgga catcqaatca gtttaccttc agatttccag       1560  
           cacaagataa ccgtgcaggc cctcccac ttggacaaac ggcggagcct gaacagcagc       1620  
           agttccagtc ccccgagcag ccccaaatg atgccccgac tccgagccat acagttgact       1680  
 20           tcagatgaaa gcaataaaac ttggggaagg aacacagtct ttcgacaaga agaatttgag       1740  
           gatgtaaaaa qgaattttaa gaaaaaaggc tqtacctcgg gaccaaaatc cattcaaatg       1800  
           aaagatagaa cagatcgcaa agaaaggata agacctctct ccgatggcaa cagtcccttg       1860  
 25           toaactatct taataaaaaa tcagaaaaac atgcccttgg cttcattggt tgtggaaccag       1920  
           ccagggtcct gtgaagagcc aaaactttcc cctgatggat tagaacacag aaaaccaaaa       1980  
 30           caaataaaat tgccatagtc ggctacatt gatcaccctc ttgggaasaga tgcacagaga       2040  
           gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgccacagtc       2100  
           tcatttgaga tgactcctac gaatagtctg agtagatcc ccagagaaa gasaacggag       2160  
 35           tcagctctgt atgggtgcac cgtcctctg ccacccgttg ctctgggact ggacctcaga       2220  
           gagcttcata aagcacaggc tgcgaagaa ccgttgcca aggaagsgaa gaagaaaoga       2280  
 40           gaggaatcc tccagcgggc tccaagtcc ccgagaagcg ccagtcctcc cacaagcctg       2340  
           ccatccact gtggggaggc cagcagccca cctccctgc cactgtcaag tgccctgggc       2400  
           atcctctcca cacctcttt ctocacaaag tgccctgctc agatggacag tgaagatcca       2460  
 45           ctggtggaca gtgcacctgt cacttgtagc tctgagatgc tcaactccga tttttgtccc       2520  
           actgccccag gaagtgtgog tgagccagcc ctcatgcca gacttgacac tgattgtagt       2580  
 50           gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt acctactgt       2640  
           gctcttcaa aacatagacc gtcacatcac agacggacca tgtctgatgg aatccgacc       2700  
           ccaactggtg caactattat cccagccact ggagcctctg cactgccact ctgccctca       2760  
 55           ccgtctctc acagtcact gccaaaggag gtctcaccca agaagcscag cactgtccac       2820  
           atcgtgcctc agcgtgcgcc tgccctcctg agaagccgct cagatctgcc tcaggcttac       2880  
 60           ccacagacag cagtgtctca gctggcacag actgcctctg tagtgggtcg cccaggacca       2940  
           catcccacc aattcctcgc tgccaaggag agaactaaat cccatgtgcc ttcattactg       3000  
 65           gatgctgacg tggaaagtca gacagggac tacaactgtc c                               3041

70           <210> 55  
           <211> 3041  
           <212> DNA  
           <213> homo sapiens

<400> 55  
 atggctttgc gggggccgc gggagcacc cacaccccg tgcctcggc cgggggagcc 60  
 5 cccggcggcc cagcgcctc gtcgtccacc tccctcggcg gctcggcctc ggcgggagcg 120  
 ggcctgtggc ccgcctcta tgactacgag cctcgcgcg agaacgagc gacctcggc 180  
 ccggccacc tggtagaggt gctgtcgcg gacgcgcgc tgcgggcga cggggctgg 240  
 10 tggcagagc aggtcagag cgcctcggc atcttcccg ccaactacgt gctccctgc 300  
 cggccggccg ccagccccc gccgcgcgc tgcggccca gctcccccgt acacgtcgc 360  
 15 ttcagagcgc tggagctgaa gtagctcctc ccgcctggg gcttcggca ggtgacgc 420  
 gccacctggc agggccaggc ggtggccgtg aaggcgggc gccaggacc ggcagagc 480  
 ggggggggg ctcgcagag cgtgcccgc gaggctcggc tcttcgccat gctgcggc 540  
 20 cccaacatca tggagctgag cggcctgtgc ctgcagcgc ccacctctg cctgggtgctg 600  
 gagtccccc ccgcggagc gctcaaccg ccgcctggc ctcaccacc cggcccccgc 660  
 25 ccggcgcgc ccggccccc ccgcgcgcgc ccgacctc ccgacctgct ggtcaactgg 720  
 gccgtgcaga tagcggggg catgctctac ctgcctgagc aggccttctg gccatcctg 780  
 cccggggacc tcaagccag caacatctt cctctgaga agatagaaca tgatgacac 840  
 30 tccaataaaa ctttgaagat tacagatctt gggttggga ggaatggca caggaccacc 900  
 aaaatgagca cagcagcac ctatgcctg atggccccc aagtgatcaa gctctcctt 960  
 35 tttctaaag gaagcagac ctggagctat gtagtctgc tgggggaact gctcacggc 1020  
 gaagtcccc atcggggcat tgatgcctc ccgctggctt atggggtagc agtcaataa 1080  
 ctactcttgc ccattccac cactgcctc ggcgctttg ccaagctcat gaaagaatgc 1140  
 40 tggcaacaag accctcatat tctccatcg ttgccttaa tctcgaaca gttgactgct 1200  
 attgaaaggc cagtgatgac tgagatgct caagaatct tctattccat gcaagatgac 1260  
 45 tggaaactag aaattcaaca aatgcttgat gagttagaa caaaggaaa ggcgctcga 1320  
 tccgggagc aggagctgac tgggcggct ctgcagcaga agtctcaggc ggcgctcga 1380  
 aagcggcgtg agcagcagc ggcagcgc gtagatgagc tgcggagcg ggaactaac 1440  
 50 attctgata tccagctaaa ccagcagag ccaaggtaa agaagaggaa gggcaagttt 1500  
 aagagaagtc gtttaagct caaagatgga catcgatca gtttacctc agatttcag 1560  
 55 cacaagataa ccgtgcagc cctcccccac ttggacaaac ggcggagcct gaacagcgc 1620  
 agttccagtc ccccgagcag cccacactg atgcctgac tccgagccat acagttgact 1680  
 tcagatgaaa gcaataaaac tggggaagg aacacagctc ttcgacaaga agaatttgag 1740  
 60 gatgtaaaa ggaatttaa gaaaaaggc tgtacctgg gaccaaaac cattcaaatg 1800  
 aaagatagca cagattgca agaaaggata agacctctc ccgatggca cagtccttg 1860  
 65 toaactatc taataaaaa tcagaaaacc atgccttgg ctctattgtt tgtggaccg 1920  
 ccaggglcc glgaagagc aaaaclllc cclgalgal gagaacacag aazaccaaa 1980  
 caataaaaa tgcctagta ggcctacatt gatctacct tgggagaga tgcagaga 2040  
 70 gagaalccc cagaagcga agcggggg gaggcagcc clggagagc tgcacaglc 2100

	tocattgaga tgactcctac gaatagtctg agtagatccc cccagagaaa gaaaacggag	2160
	tcagctctgt atgggtgcac cgtccttctg ccatcgggtg cctcgggact ggacctcaga	2220
5	gagcttcata aagcacaggc tgcctgaagaa ccgttgccca aggaagsgaa gaagaaacga	2280
	gagggaatcc tccagcgggc tccaagtcc cgcagaagcg ccagtccccc cacaagcctg	2340
10	ccatccacct gtggggaggc cagcagccca cctccctgc cactgtcaag tgcctcggc	2400
	atcctctcca cacctcttt cccacaaag tgcctgctgc agatggacag tgaagatcca	2460
	ctgggtgaca gtgcacctgt cacttgtagc tctgagatgc tcaactcggga tttttgtccc	2520
15	actgccccag gaagtggctg tgagccagcc ctcctgcca gacttgccac tgattgtagt	2580
	glalcaagaa acllgcgcgc lccclccca cagcagacal glgggaalgl accllacldl	2640
20	gcttcttcaa aacatagacc gtcacatcac agacggacca tctctgatgg aatccgacc	2700
	ccaactggtg caactattat cccagccact ggagcctctg cactgccact ctgccccca	2760
	ccctgctcctc acagtcctct gccaaaggag gctccacca agaagcacag cactgtccac	2820
25	atcgtgcctc agcgtgcctc tgcctcctg agaagccctc cagatctgcc tcaggcttac	2880
	ccacagacag cagtgtctca gctggcacag actgctctgt tagtgggtcg cccaggaoca	2940
30	caccccccc aallcctcgc tgcgaaggag agaactaaaal cccalglgc llcallacld	3000
	gatgctgacg tggaaaggta gacagggac tacactgtgc c	3041
35	<210> 56 <211> 3041 <212> DNA <213> homo sapiens	
40	<400> 56 atggctttgc ggggcgcccg cggagcgacc gacaccccgg tgcctcggc cgggggagcc	60
	cccgcgcgcc cagcgcctc gtcgtccacc tctcggcg gctcggctc ggcggcgcg	120
45	ggcgtgtggg ccgctctcta tgactacgag cctcggcg aggaacagct gagcctcgg	180
	cccgccagc tgggtggagt gctgtcgcag cagcccgcc tgcggggca cagggctgg	240
	tggcagggc aggtgcagcg gcgcctcggc atcttcccag ccaactacgt ggtccctgc	300
50	cgcccgcccg ccagcccccg gcgcgcgcc tgcggcca gctcccggg acacgtcgc	360
	ttcagcggc tggagctgaa ggagctcctc cgcctcggg gcttcggga ggtgtaccg	420
	gccacctggc agggccagga ggtgcctgt aaggcggcg gccaggacc ggagcaggc	480
55	gggcggcg ctgcccagag cgtgcccgc cagcctcggc tcttcgcat gctcggcac	540
	cccaacatca tcagctcgc cgcctgtgc ctgcagcag ccacacctc cctggtctg	600
60	gagttcgc ccggcgagc gctcaaccga cgcctggcg ctgccaacgc cgcggcgac	660
	ccgcgcgc ccggccccg ccgcgcgc cgcacctc ccacgtgct ggtcaactgg	720
	gcgtgcaga tagcgggg catgctctac ctgcatgag aggccttct gccatcctg	780
65	caccggacc lcaagccag caacallllg clacllgaga agalagsaca lgalgacalc	840
	tcgaataaaa ctttgaagat tacagatctt cggttggcg ggaatggca caggaccacc	900
70	aaaalgagca cagcaggcac clalgccgg alggccccg aaglgalcaa glcllclll	960

ttttctaagg gaagagacat ctggagctat ggagtgcctgc tgtgggaact gctcacogga 1020  
 gaagtcccc atcggggcat tgatgcctc gccgtggctt atggggtagc agtcaataaa 1080  
 5 ctcactttgc ccattccatc cacctgcctt gaggcgtttg ccaagctcat gasagaatgc 1140  
 tggcaacaag accctcatat tcttccatcg tttgccttaa ttctcgsaca gttgactgct 1200  
 10 attgaagggg cagtgatgac tgagatgcct caagaatctt ttcaattccat gcaagatgac 1260  
 tggaaactag aaattcaaca aatgtttgat gagttagaaa caaaggaaaa ggagctcoga 1320  
 tcccggaag aggagctgac tgggcggct ctgcagcaga agtctcagga ggagctccta 1380  
 15 aagggcgctg agcagcagct ggcagcagcc gagatcagc tggcggcggc ggactttaac 1440  
 alclclgala lccagclaaa ccagcagcag cccaagglaa agaagaggaa ggcacaglll 1500  
 aagagaagtc gtttaagct caaagatgga catcgaatca gtttacctc agatttccag 1560  
 20 cacaagataa ccgtgcaggc ctctcccaac ttggacaaac ggcggagcct gascagcagc 1620  
 agttccagtc ccccgagcag ccccaaatg atgccccgac tctgagccat acagttgact 1680  
 25 tcagatgaaa gcaataaaac tggggaagg aacacagctc ttgcacsaga agaatttgag 1740  
 gatgtaaaaa ggaattttaa gaaaaaaggc tgtacctggg gaccaaatc cattcaaatg 1800  
 aaagalagaa cagalgcaa agaaaggala agaccclclcl ccgalggcaa caglclclgg 1860  
 30 tcaactatct taataaaaaa tcagaaaaac atgcccttgg ctccattggt tgtggaccag 1920  
 ccagggccc glgaagagcc aaaaclllcc cclgalggal lagaaccag aaaaaccaaa 1980  
 35 caaataaaat tgcctagtcg ggcctacatt gatctacctc ttgggacaga tgcctcagagc 2040  
 gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgactgc tgcacagtc 2100  
 40 tccattgaga tgactcctac gaatagtctg agtagatccc cccagagaaa gacaaccggag 2160  
 tcagctctg atggggcagc cgtcctctg gcatcggctg ctctgggact ggacctcaga 2220  
 gagcttcata aagcacaggc tctgaagaa ccgttgcca aggaagagaa gaagaaaaga 2280  
 45 gagggaatc tccagcgggc tccaagtc ccagagaagcc ccagtcctcc cacaagcctg 2340  
 ccctccacc ctggggaggc cagcagccca cctccctgc cactgtcaag tgcctcggc 2400  
 50 atctctcca cactctctt ccccaaaag tgcctgctgc agatggacag tgaagatcca 2460  
 ctgggtgaca gtgcacctg cacttgtagc tctgagatgc tcactccgga tttttgtccc 2520  
 actgcccag gaagtggctg tgagccagcc ctcatgcca gacttgacc tgattgtagt 2580  
 55 gtatcaagaa acttgcctc tccctccta cagcagacat gtcggastgt acctactgt 2640  
 gcttctca aacatagacc gtcacatcac agacggacca tctctgctgg aatccgacc 2700  
 60 ccaactggtg caactattat ccaaccact agagcctctg cactgccact ctgcccctca 2760  
 ccctgctctc atagtcctct gccaaaggag gctctaccca agaagcag cactgtccac 2820  
 atcgtcctc agcctcctc tgcctccctg agaaagcct cagactctcc tcaggttac 2880  
 65 ccacagacag cagtgtctca gctggcacag actgctctg tagtggctc cccaggacca 2940  
 catcccacc aattcctcgc tgccaaggag agaactaaat cccatgtgcc ttcattactg 3000  
 70 gatgctgacg tggaaagtca gacagggac tacactgtgc c 3041

```

<210> 57
<211> 3041
<212> DNA
<213> homo sapiens

5
<400> 57
atggcctttgc gggggcgcgc gggagcgcac gacaccccgg tgcctcgggc cgggggagcc      50
cccgggggcct cagcgcctctc gtcgtccacc tctcggggcg gctcggcctc ggggggagcg      120
10 gggctgtggg ccgcgcctcta tgactacgag gctcggggcg aggacgagct gagcctcggg      180
cggcggccagc tggtgagggt gctgtcgcag gacgcgcgcg tgcggggcga cgggggctgg      240
15 tgggcaggcc aggtgcagcg ggcctcggc atcttcccgg ccaactacgt ggcctcctgc      300
cggccgggcg ccagccccgc gccgcgcgcc tgcgggcaca gclccccggL acacglgccc      360
20 ttcagagcggc tggagctgaa ggagctcacc ggcgctgggg gcttggggca ggtgtaocgc      420
gccacctggc agggccaggga ggtgcgcgtg aaggcggcgc gccaggaccc ggagcaggac      480
gcgggcggcgg ctgcgcgagag cgtgcggcgc gaggctcggc tcttcgccat gctgcggcac      540
25 cccaacatca tcgagctgog cggcgtgtgc ctgcagcagc cgcacctctg cctgggtctg      600
gagttgcgcc gggggggagc gctcaaccca ggcctggcgc ctgccaacgc cggccccggc      660
30 cggcgcgcgc cggccccccg ccgcgcgcgc cgcaccccgc cgcacglgcl gglcaaclgg      720
gcccgtgcaga tagcgcgggg catgctctac ctgcatgagg aggccttctg gccatcctg      780
caccgggacc tcaagccag caacalLLLg clacLLgaga agalagsaca lgalgacalc      840
35 tgcataaaca ctttgaagat tacagatctt gggttggcga gggaatggca caggaccacc      900
aaaatgagca cagcaggcac ctatgcctgg atggcccccg aagtgatcaa gtcttctctg      960
40 ttttctaagg gaagcgacat ctggagctat cggagtctgc tctgggcaact gctcacccgc      1020
gaagtcctct atcggggcat tgatggctc gccgtggctt atggggtagc agtcaataaa      1080
ctcactttgc ccattccacc caactgcctt ggcgcgtttg ccaagctcat gaaagaatgc      1140
45 tggcaacaag accctcatat tcgtccatcg tttgccttaa ttctcgaaca gttgactgct      1200
atcgaaggcg cagtqatgac tqagatgctt caaqaatctt ttcattccat gcaagatgac      1260
50 tggaaactag aaatcaaca aatgthttgat gagttgagaa caaaggaaaa ggagctcoga      1320
tcccgggacg aggagctgac tggggcggct ctgcagcaga agtctcagga ggagctgcta      1380
aagcggcgtg agcagcagct ggcagagcgc gagatcgacg tctgagagcg ggaactaac      1440
55 attctgatac tcacgctaaa ccagcagcag cccaagqtaa aqaagagqaa ggcacaactt      1500
aagagaagtc gtttaagct caaagatgga catcgaaatca gtttacctc agatttccag      1560
60 cacaagataa ccgtcagcgc ctctcccaac ttggacaaac ggcggagcct gaaacagcgc      1620
agttccagtc ccccgagcag ccccaaatg atgccccgac tccgagccat acagttgact      1680
tcagatgaaa gcaataaaaac tggcgaaag aacacagctct ttcgacsaga aqsatttgaq      1740
65 gatgtaaaaa ggaattttaa gaaaaaaggf tgtacctggg gaccaatctc cattcaaatg      1800
aaagatagaa cagattgcaa agaaaaggata agacctctct ccgatggcaa cagtccttgg      1860
70 tcaactatcc taataaaaaa tcagaaaacc atgcccttgg cttcattggt tgtggaccag      1920
ccagggtccc gtgaagagcc aaaactttcc cctgatggat tagaacacag aaaaccsaaa      1980

```

	caaataaaac	tgcctagtea	ggcctacatt	gatctacctc	ttgggacaga	tgotcagaga	2040		
5	gagaatcctg	cagaagctga	aagctgggag	gaggcagcct	ctcgaaatgc	tgocacagtc	2100		
	tocattgaga	tgaactctac	qaatagtctg	agtatgctcc	cccagagaaa	gaaaaccag	2160		
	tcagctctg	atggg-gcac	cgctcctctg	gcctcgggtg	ctctgggact	ggcctcaga	2220		
10	gaqcttcata	aaqccacagg	tqctcaagaa	ccgttgccca	aggaagsgaa	gaaqaaacga	2280		
	gaggaatcc	tccagcgggc	tcccaggtcc	ccgagaagcg	ccagtccctc	cacaagcctg	2340		
	ccatccacc	qtgggagagc	cagcagccca	ccctccctgc	cactgtcaag	tgcctcggc	2400		
15	atcctctcca	cacctctctt	ctccacaag	tgctcctctg	agatggacag	tgaagatcca	2460		
	ctgggtgaca	gtgcacctgt	cacttgtgac	tctgagatgc	ccactccgga	tttttgtccc	2520		
20	actgccccag	gaagtggctg	tgagccagcc	ctcatgccaa	gactcgacac	tgattgtagt	2580		
	qtatcaagaa	acttccctc	tcccttcta	cagcagacat	qtgggaaagt	accttaactgt	2640		
	gcttcttcaa	aacatagacc	gtccatcac	agacggacca	tgctctgctg	aaatccgacc	2700		
25	ccaactggtg	caactattat	ccagccact	ggagcctctg	cactgccact	ctgcccctca	2760		
	ccctgctctc	acagtcctct	gccaaaggag	gtctcaccca	agaagcacag	cactgtccac	2820		
30	atcgtgcctc	agcgtccccc	tgctccctg	agaagccct	cagatctgcc	tcaggcttac	2880		
	ccacagacag	cagtgctcca	gctggcacag	actgctctg	tagtggctcg	cccaggacca	2940		
	catcccacc	aattcctcgc	tgccaaggag	agaactaaat	cccatgtgcc	ttcattactg	3000		
35	gatgctgacg	tggaagggtca	gagcagggac	tacactgtgc	c		3041		
40	<210>	58	<211>	3041					
	<212>	DNA	<213>	homo sapiens					
45	<400>	58	atggcttttg	ggggcgccgc	gggagcgcac	gacaccccgg	tgctcctggc	cgggggagcc	60
			cccgcggcc	cagcctctc	gcctccacc	tctcggccg	gctcggctc	ggcgggagc	120
50			gggctgtgg	ccgcctcta	tgactacag	gctcgggg	aggaagagct	gagcctcgg	180
			ccgcggcag	tggtggaggt	gctgtcag	gacgcggcc	tgctggggca	cgagggtgg	240
			tgggcaggcc	aggtgcagcg	gcgcctcggc	atcttcccgg	ccaactacgt	ggctccctgc	300
55			ccccggccg	ccagccccgc	gcgcggccc	tcgcggccc	gctccccgt	accgctcgc	360
			ttcgagcggc	tgagctgaa	ggagctcctc	ggcctgggg	gcttcgggca	ggtgtaacc	420
60			gccacctggc	agggccaqga	qgtgcctctg	aaqgcggcc	gccagqacc	ggaqcaqqac	480
			ggggcgggg	ctgcagagag	cgctggggc	gaggctcggc	ctctcggcat	gctgcggcac	540
			ccccacatca	tcgaactcgc	cgccctctgc	ctgcagcagc	ccacactctg	cctggtctctg	600
65			gagttcggcc	gggggggagc	gctcaaccga	ggcctggccg	ctgccaacgc	cgccccggac	660
			ccgcggcgc	ccggcccccg	ccgcggcgc	cgcatccctc	cgcaactgct	ggtcaactgg	720
70			gocgtgcaga	tagcggggg	catgctctac	ctgcatgagg	aggccttcgt	gcccactcctg	780
			caccgggacc	tcaagccag	caacatcttg	ctacttgaga	agatagsaca	tgatgacatc	840

	tgcaataaca ctttgaagat tacagatfff eggttggcga gggaatggca caggaccacc	900
5	aaaatgagca cagcaggcac ctatgcctgg atggccccg aagtgatcaa gtcttccttg	960
	ttttctaagq gaagcagcat ctggagctat cgaatgctgc tctgggaact gctcacccga	1020
	gaagtcccc atcggggcat tgatggcctc gccgtggctt atggggtagc agtcaataaa	1080
10	ctcactttgc ccattccatc cacctgccct cagccgtttg ccaagctcat gasagaatgc	1140
	tggcaacaag accctcatat tctgcatcgt tttgccttaa ctctogaaca gttgactgct	1200
15	atggaaggq cagtqatqac tgagatgct caagaatctt ctattccat gcaagatqac	1260
	tggaaactag aaattcaaca aatgtttgat cagttgagaa caaaggaaaa ggagctcga	1320
	tcccggaag aggagctgac tggcgctgt ctgcagcaga agtctcagga ggagctgcta	1380
20	aagcggcgtg agcagcagct ggcagagcgc cagatcgacg tctggagcgc ggaactaac	1440
	attctgatac tccactaaa ccagcagaag cccaagataa aqaagagaa ggcacaagttt	1500
25	aagagaagtc gtttaagct caaagatgga catcgaatca gtttacctc agatttccag	1560
	cacaagataa ccgtgcaggc ctctccsac ttggcaaac ggcggagcct gascagcagc	1620
	agttccagtc ccccgagcag cccacaatg atgcccagc tccgagccat acagttgact	1680
30	tcagatgaaa gcaataaaac tggggaagg aacacagtct ctgcacsaga agaatttgag	1740
	gatgtaaaa ggaatttta gaaaaaggc tgtacctgg gaccaaatc cattcaaatg	1800
35	aaagatagaa cagattgcaa agaaaggata agacctctct ccgatggcaa cagtccttg	1860
	tcaactatct taataaaaa tcaggaaacc atgcccctgg ctctattggt tgggaccag	1920
	ccagggtccg gtgaagagcc aaaactttcc cctgatgat tagaaccag azaacaaaa	1980
40	caataaaac tgcctagtca ggcctacatt gatctacctc ttgggaaga tgcacagaga	2040
	gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgcacagtc	2100
45	tcatttgaga tgactctac gaatagtctg agtagatcc ccagagaaa gaaacccgag	2160
	tcagctctgc atgggtgac cgtcctctg ccatcgggtg ctctgggact ggacctcaga	2220
	gagcttcata aagccaggc tgotgaagaa ccgttgcca aggaagagaa gaagaaacga	2280
50	gagggaatct tcacggggc tccaagtc ccagaaagc ccagtcctcc cacaagcctg	2340
	ccatccaccg ttggggaggc cagcagccca cctccctgc cactgtcaag tgcctgggc	2400
55	atcctctcca cacctcttt cccacaaag tgctgctgc agatggacag tgaagatcca	2460
	ctgggggaca gtgcacctgt caottgtgac totgagatgc tcactccgga tttttgtcc	2520
	actgccccag gaagtggctg tgagccagcc ctcatgcaa gacttgacac tgattgtagt	2580
60	gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt acctactgt	2640
	gcttcttcaa aacatagacc gtcacatcac agacggacca tgcctgatgg aatccgacc	2700
65	caaactggtg caactattat cccagccact ggagcctctg cactgccaact ctgcccctca	2760
	ccgctccctc acaglcactc gccaaaggag gctccacca agaagccag cactgcccac	2820
	atcgtgcctc agcgtgcctc tgcctccctg agaagccct cagatctgcc tcaggcttac	2880
70	ccacagacag cagly.ctca gctggcacag actgccclyg laglygglyg cccaggacca	2940

catcccacc aattcctcgc tgccaaggag agaactaaat cccatgtgcc ttcattactg 3000  
 gatgctgacg tggaaaggtca gagcagggac tacactgtgc c 3041  
 5  
 <210> 59  
 <211> 3041  
 <212> DNA  
 <213> homo sapiens  
 10  
 <400> 59  
 atggctttgc ggggagccgc gggagcgacc gacaccccgg tgtcctcggc cgggggagcc 60  
 15  
 cccggcggcc cagcgccttc ttcgtccacc tctcggggcg gctcggcctc ggcgggcggc  
 gggctgtggg ccgctctcta tgactacgag gctcggcgcg aggacgagct gagcctcggg 180  
 cggcggccagc tgggtggaggt gctgtcgcag gacgcgcgcg tgtcgggcga cgggggtgg 240  
 20  
 tgggcaggcc aggtgcagcg ggcctcggc atcttcccgc ccaactacgt ggctccctgc 300  
 cccccggccg ccagcccccg cccgcgcgcc tcgcggccca gctccccggt acacgtcggc 360  
 25  
 ttccgagcggc tggagctgaa ggagctcacc ggcgctgggg gcttcgggca ggtgtaaccg 420  
 gccacctggc agggccaggc ggtggccgtg aaggcggcgc gccaggaccc ggagcaggac 480  
 gggcggcgcg ctgcgcgagc cgtgcggcgc gaggctcggc tcttcgccat gctgcggcac 540  
 30  
 cccaacatca tcgagctgcg cggcgtgtgc ctgcagcagc cgcacctctg cctgggtgctg 600  
 gagttcggcc cggcgggagc gctcaaccga cgcctggcgc ctgccaacgc cgcgccggac 660  
 35  
 ccgcgcgccg ccggcccccg ccgcgcgcgc cgcaccctc ccacagtgt ggtcaactgg 720  
 gccgtgcaga tagcgcgggg catgctctac ctgcatgagg aggccttctg gcccatcctg 780  
 caccgggacc tcaagccag caacatcttg ctacttgaga agatagsaca tgatgacatc 840  
 40  
 tcgaataaaa ctttgaagat tacagatctt cggttggcga ggaatggca caggaccacc 900  
 aaaatgagca cagcaggcac ctatgcctgg atggcccccg aagtgatcaa gtcttccttg 960  
 45  
 ttttctaagc gaagcgacat cggagctat cggagtctgc tgtgggaact gctcacggga 1020  
 gaagtcccc atcggggcat tgatggcctc gccgtggctt atggggtagc agtcaataaa 1080  
 ctcactttgc ccattccatc caoctgcctc gaggcgtttg ccaagctcat gaagaaatgc 1140  
 50  
 tggcaacaag accctcatal tctccatcg tttgcctaa ttctogaaca gttgactgct 1200  
 atcgaagggc cagtgatgac tgagatgcct caagaatctt ctattccat gcaagatgac 1260  
 55  
 tggaaactag aaattcaaca aatgttgat cagttgagaa caaaggaaaa ggagctcga 1320  
 tcccgggaag aggagctgac tcggcggcct ctgcagcaga agtctcagga ggagctcota 1380  
 aagcggcgtg agcagcagct ggcagcgcgc gagatcgacg tgcgggagcg ggaacttaac 1440  
 60  
 atctgatac tccagctaaa ccagcagaag cccaaggtaa agaagaggaa gggcaagttt 1500  
 aagagaagtc gtttaagct caaagatgga catcgatca gtttacctc agatttcag 1560  
 65  
 cacaagataa ccgtgcagcc ctctcccaac ttggacaaac ggcggagcct gaacagcagc 1620  
 agctccaglc ccccgagcag cccacacagc atgccccgac cccgagccal acagllgac 1680  
 tcagatgaaa gcaataaaa tgggggaagg aacacagctc ctcgacsaga agaatttgag 1740  
 70  
 gaglaaaaa ggaalctaa gaaaaaaggc tglacclygg gaccaaaalc callcaaal 1800

	aaagatagaa cagatcgcaa agaaaggata agacctctct ccgatggcaa cagtccttgg	1850
	tcaactatcc taataaaaaa tcagaaaaacc atgcccttgg cttcattggt tgtggaccag	1920
5	ccagggtccc gtgaagagcc aaaactttcc cctgatggat tagaacacag aasacccaaa	1980
	caataaaaaa tgcctagtca ggctacatt gatctacctc ttgggaasaga tgcacagaga	2040
10	gagaatcctg cagaagctga aagctgggag gaggcagcct ctgcgaatgc tgcacagtc	2100
	tcattgaga tgactcctac gaatagtctg agtagatccc ccagagaaa gaaaccggag	2160
	tcagctctgt atgggtgcac cgtccttctg gcacgggtgg ctctgggact ggacctcaga	2220
15	gagcttcata aagccacaggc tgcctgaaga ccgttgccca aggaagcaga gacgaaacgc	2280
	gagggaaacc tccagcgggc tccaagacc ccagaccacc caaagccctg	2340
20	ccatccaccg gtggggaggc cagcagacca cctccctgc cactgtcaag tgcctcgggc	2400
	atcctctcca cacctctttt ctccacaag tgctctctgc agatggacag tgaagatcca	2460
	ctgggtgaca gtgcacctgt cacttctgac tctgagatgc tccactcggga tttttgtccc	2520
25	actgcctcag gaagtgtctg tgagccagcc ctcatgccaa gacttgacac tgattgtagt	2580
	gtatcaagaa acttgccgtc ttccttcta cagcagacat gtgggaatgt acctaatgt	2640
30	gctccttcaa aacalagacc gtcacalcac agacggacca gctcctcagg aaalccgacc	2700
	ccaactggtg caactattat ctccagccact ggagcctctg cactgccact ctgccctcca	2760
	ccctcctccc acaglcaccl tccaaggag gctcacccca agaagccacg caacgcccac	2820
35	atcgtcctcc agcgtccccc tgcctcctg agaagccctc cagatctgcc tcaggcttac	2880
	ccacagacag cagtgctcca gctggcacag actgcctctg tagtgggtcg ccacagacca	2940
40	catccaccac aattcctcgc tgcacaaggag agaactaaat cccatgtgcc ttcattactg	3000
	gatgctgacg tggaaaggtca gagcagggac tacactgtgc c	3041
45	<210> 60 <211> 3041 <212> DNA <213> homo sapiens	
50	<400> 60 atggctttgc gggggccgc gggagcagc gacaccccgg tgcctcggc cgggggagcc	60
	cccggcggcc cagcgtcctc gtcgtccacc tctcggggcg gctcggcctc ggcgggogcg	120
55	ggcgtgtggg ccgcgctcta tgactacgag gctcggggcg aggaocagct gagcctcggg	180
	ccggccagc tgggtggagt gctgtcagc gacgccggcg tgcggggcga cgaaggctgg	240
	tgggcaggcc aggtgcagcg ggcctcggc atcttcccgg ccaactacgt ggtcccctgc	300
60	cgcccgcccg ccagccccgc gccgcggccc tgcgggccca gctcccgggt acacgtcggc	360
	ttcagcgggc tggagctgaa ggagctcctc ggcgctgggg gcttcgggca ggtgtaaccg	420
	gocacotggc agggccagga ggtggccgtg aaggcggggc gccagaccac ggagcaggac	480
65	gcgccggcgg clgcccagag cglcggggcg caggcctcgg ccllccgccal gclgcccac	540
	cccaacatca tcagactcgg cggcgtgtgc ctgcagcagc ccacacctc cctgggtctg	600
70	gagllccccc gggggggagc gclcaaccga gcgclggggc clgcccacgc cggcccggac	660

cccgcgcgcg cccgcccccg cccgcgcgcg cgcacccctc cgcacgtgct ggtcaactgg 720  
 gccgtgcaga tagcgcgggg catgctctac ctgcatgagg aggccttcgt gcccatcctg 780  
 5 caccgggacc tcaagccag caacattttg ctacttgaga agatagsaca tgatgacatc 840  
 tcgaataaaa ctttgaagat tacagatttt cgggtggcga ggaatggca caggaccacc 900  
 10 aaaatgagca cagcaggcac ctatgccttg atggccccg aagtgatcaa gtcttccttg 960  
 tttcttaagg gaagcagat ctggagctat cggagtctgc tgtgggaact gctcacogga 1020  
 gaagtcccc atcggggcat tgatgcctc cccgtggctt atggggtagc agtcaataaa 1080  
 15 ctcactttgc ccattccatc cccctgccct cagccgtttg ccaagctcat gacagaatgc 1140  
 tggcaacaag accclcalal cglccalcg lllgccllaa llclcgaaaca gllgactgcl 1200  
 20 attgaagggg cagtgatgac tgagatgctt caagaatttt ctattccat gcsagatgac 1260  
 tggaaactag saattcaaca aatgtttgat gagtggagaa caaaggaaaa ggagctgoga 1320  
 tcccggggag aggagctgac tccggcggct ctgcagcaga agtctcagga ggagctgcta 1380  
 25 aagcggcgtg agcagcagct ggcagagcgc gagatcgacg tctggagcg ggaacttaac 1440  
 attctgatac tccagctaaa ccaggagaag cccaaggtaa agaagaggaa gggcaagttt 1500  
 30 aagagaaglc glllaaagcl caaagalgga calcgaaalca glllaccllc agalllccag 1560  
 cacaagataa cccgtgcaggc ctctcccaac ttggacaaaac ggcggagcct gascagcagc 1620  
 agllccaglc ccccgagcag ccccaaalg algccccgac ccgagccal acagllgact 1680  
 35 tcagatgaca gcaataaaac tggggagag aacacagctt ctgaccaga agcatttgag 1740  
 gatgtaaaaa ggaattttta gaaaaaagg tgtacctggg gaccaaaatc cattcaaatg 1800  
 40 aaagatagca cagatcgcaa cgaaggata agacctctct ccgatggcaa cagtccttgg 1860  
 tcaactatct taataaaaaa tcagaaaaac atgccttgg ctctattggt tgtggaacag 1920  
 ccagggtccc gtgaagagcc aaaactttcc cctgatgat tagaacscag aaaacccaaa 1980  
 45 caaataaaa tgcctagtca ggcctacatt gatctacctc ttgggaaga tgctcagaga 2040  
 gagaatcctg cagaagctga saqctgggag cagcagcct ctgcgaatgc tgcacacgct 2100  
 50 tccattgaga tgaactctac gaatagtctg agtagatccc ccagagaaa gaaaacggag 2160  
 tcagctctgt atggggcac cgtccttctg ccatcgggtg ctctgggact ggacctcaga 2220  
 gagcttcata aagccacaggc tgotgaaga ccgttgccc aaggaagaga gaaagaaaga 2280  
 55 gagqqaatc tccagcggc tccaaqtcc ccagaagcc ccagtcctcc cacaagcctg 2340  
 ccctccacc gtggggaggc cagcagccca cctccctgc cactgtcaag tgcctgggc 2400  
 60 atcctctcca cacctcttt ctccacaaag tccctgctgc agatqgscag tgaagatcca 2460  
 ctgggtgaca gtgcaactgt cacttgtgac tctgagatgc tcaatccgga tttttgtccc 2520  
 actgccccag gaagtgtctg tgaqccagcc ctcatgcca qacttgscac tgaattgaqt 2580  
 65 gtatcaagaa acttgccgtc tcccttcta cagcagacat gtgggaatgt acctactgt 2640  
 gctcttcaa aacatagacc gccacatccc agacggacca tgtctgctgg aatccgacc 2700  
 70 ccaactgggt caactattat ctccagccact cggacctctg cactgccact ctgcccctca 2760  
 cctgctcctc acagtcactt gccaaaggag gctcaccoca agaagcacag cactgtccac 2820

atcgtgcctc agcgtcgcgc tgccctccctg agaagccgct cagatctgcc tcaggcttac 2880  
 ccacagacag cagtgctca gctggcacag actgacctgt tagtgggtcg cccaggacca 2940  
 5 catcccacc aattctctgc tgccaaggag aqaactaaat cccatgtgcc ttcattactg 3000  
 gatgctgacg tggaaaggtca gagcagggac tacactgtgc c 3041  
 10  
 <210> 61  
 <211> 3041  
 <212> DNA  
 <213> homo sapiens  
 15  
 <400> 61  
 atggclllge ggggggcgc gggagcgc caccaccgg lglcclcgge cgggggagcc 60  
 ccggcgccgc cagcgtccctc gctgtccacc toctcggcg gctcggctc ggccggcgcg 120  
 20 gggctgtggg ccgcgctcta tgactacgag cctcggcg aggaocgagct gagcctcggg 180  
 cgcggccagc tggtagaggt gctgtgcag gacgcgcgc cgtcggcgga cgggggtgg 240  
 25 tgggcaggcc aggtgcagcg gcgcctcggc atcttcccgc ccaactacgt ggtccctgc 300  
 cgcggcgccg ccagccccc gcgcgcgcgc tocgggcca gctcccggt acacgtcgc 360  
 llogagcggc lggagclgaa ggagclcalc ggcgclggg gcllcgggca gglglacgc 420  
 30 gccacctggc agggccagga ggtgcgcgtg aaggcggcgc gccaggacc ggagcaggac 480  
 gcggcgccg clgcagagag cglgcggcgc gaggclcgge cllogccal gclgcggcac 540  
 35 cccaacatca togagctcg cggcgtgtgc ctgcagcagc ccacactctg cctggctctg 600  
 gagttccccc gcggcgagc gctcaaccga gcgctggcgc ctgccaccgc cgcggcgac 660  
 ccgcgcgcgc ccggccccc gcgcgcgcgc ccacaccctc ccacactgct ggtcaactgg 720  
 40 gcgctgcaga tagcggggg catgctctac ctgcatgagg aggccttctg gccacactc 780  
 cacggggacc tcaagccag caacatcttg ctacttgaga agatagsaca tgatgacac 840  
 45 tgcaataaaa ctttgaagat tacagatctt gggttggcga ggaatggca caggaccacc 900  
 aaaatgagca cagcagcac ctatccctg atggccccc aagtqatcaa gttctcctg 960  
 50 tttctaaagg gaagcgacat cggagctat ggagtgtgc tgtgggaact gctcacggga 1020  
 gaagtcccc atcggggcat tgatcgctc gcgctggctt atggggtagc agtcaataaa 1080  
 ctcactttgc ccattccatc cccctgcctt gagccgtttg ccaagctcat gasagaatgc 1140  
 55 toqcaacaag accctcatal togtccatcg tttgcctaa ctctcgsaca gttgactgct 1200  
 atgaagggg cagtgatgac tgagatgct caagaatctt ctcatccat gcaagatgac 1260  
 60 toqaaaacta aaattcaaca satgtttgat caqttgaca caaaqgaaa ggaqctcga 1320  
 tccggggaag aggagctgac tggggcgct ctgcagcaga agtctcagga ggagctgcta 1380  
 aaqcgcgctg agcagcagct qccacagcgc cagatcagc cgtcggscgc ggaacttaac 1440  
 65 atctgatac tccagctaaa ccaggagaag cccaaggtaa agaagaggaa gggcaagttt 1500  
 aagagaagct gtttaagct caaagatgga catcgaatca gtttaccctc agatttccag 1560  
 70 cacaagataa ccgtgcagc ctctccaac ttggacaaac ggcggagcct gaacagcagc 1620  
 agttccagtc ccccgagcag cccacaatg atgccccgc cccgagccat acagttgact 1680

	tcagatgaca gcaataaaaac ttggggaagg aacacagttc ttogacsaga agsatttgag	1740
5	gaggtaaaaa ggaattttta gaaaaaagggt tgtacctggg gaccaaaatc cattcaaatg	1800
	aaagatagaa cagatcgcaa agaaaaggata agacctctct ccgatggcaa cagtccctgg	1860
	tcaactatcc taataaaaaa tcagaaaacc atgcccttgg ctccattggt tgtggaccag	1920
10	ccagggtccc qtgaagagcc aaaaactttcc cctgatgcat tagaacscag aasacccaaa	1980
	caaaataaaa ttgctagttca ggctacatt gatctacctc ttgggaaaga tgcctagaga	2040
15	gagaatcctg cagaaqctga aagctgggag cagcagcct ctgcgaatgc tgcacagtc	2100
	tccattgaga tgactcctac gaatagttct agtagatccc ccagagaaa gaaaacggag	2160
	tcagctctgt atgggtgcac cgtcctctg gcctcgggtg ctctgggact ggacctcaga	2220
20	gagcttcata aagcaccaggc tgcctgaaga ccgttgcaca aggaagagaa gaagaaacga	2280
	gagqqaatcc tccacgggc tccaaqtc ccacagaacc ccagtcctcc cacaagcctg	2340
25	ccatccacc ttggggaggc cagcagccc cctccctgc cactgtcaag tgcctcggc	2400
	atcctctcca cactctctt cccacacag tgcctgctgc agatggscag tgsagatcca	2460
	ctgggtggaca gtgcacctgt cacttgtgac tctgagatgc tccctcggga tttttgtccc	2520
30	accgccccag gaagtggctg tgagccagcc ctcatgccc gaacttgscac tgattgtagt	2580
	gtatcaagaa acttgccgtc ttccttccca cagcagacat gtgggaatgt accctactgt	2640
35	gcttcttcaa aacatagacc gtcacatcac agacggacca cgtctgatgg aatccgacc	2700
	ccaactggtg caactattat cccagccact ggagcctctg cactgccact ctgcccctca	2760
	ccgctcctc acagtcctct gccaaaggag gtctcaccca agaagcacag cactgtccac	2820
40	atcgtgcctc agcgtcgccc tgcctccctg agaagccgct cagatctgcc tcaggcttac	2880
	ccacagacag cagtgtctca gctggcacag actgcctctg tagtgggtcg cccaggacca	2940
45	catcccacc aattcctgct tgcgaaggag agaactaaat cccatgtgcc ttcattactg	3000
	gagctgacg tggaaagtca gagcaggac tacactgtgc c	3041

**Bibliografy**

1. A. Bardelli *et al.*, *Science (New York, N.Y)* **300**, 949 (May 9, 2003).
2. H. Davies *et al.*, *Nature* **417**, 949 (Jun 27, 2002).
- 5 3. T. Maeda *et al.*, *Nature* **433**, 278 (Jan 20, 2005).
4. S. Shirasawa, M. Furuse, N. Yokoyama, T. Sasazuki, *Science (New York, N.Y)* **260**, 85 (Apr 2, 1993).
5. P. M. Comoglio, L. Trusolino, *The Journal of clinical investigation* **109**, 857 (Apr, 2002).
- 10 6. S. Velho *et al.*, *Human molecular genetics* **19**, 697 (Feb 15).
7. D. Brancho *et al.*, *Molecular and cellular biology* **25**, 3670 (May, 2005).
8. S. Arena, A. Pisacane, M. Mazzone, P. M. Comoglio, A. Bardelli, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11412 (Jul 3, 2007).
- 15 9. Y. Samuels *et al.*, *Cancer cell* **7**, 561 (Jun, 2005).
10. P. Michieli *et al.*, *Oncogene* **12**, 775 (Feb 15, 1996).
11. A. Follenzi, L. E. Ailles, S. Bakovic, M. Geuna, L. Naldini, *Nature genetics* **25**, 217 (Jun, 2000).
- 20 12. M. Kohli, C. Rago, C. Lengauer, K. W. Kinzler, B. Vogelstein, *Nucleic acids research* **32**, c3 (2004).

**Claims**

1. An *in vitro* diagnostic method comprising measuring MLK4 gene expression in a tumor cell sample material of a patient, wherein measuring MLK4 gene overexpression in said tumor cell sample material is indicative of an invasive potential of said tumor.
2. The *in vitro* diagnostic method according to claim 1, wherein said measuring MLK4 gene expression comprises measuring MLK4 protein expression and/or MLK4 coding nucleic acid, preferably RNA, expression.
3. The *in vitro* diagnostic method according to claim 1 or claim 2, wherein said tumor cell sample material is a bioptic sample.
4. The *in vitro* diagnostic method according to any one of the preceding claims, wherein said bioptic sample corresponds to an invasive portion of said tumor cell sample material.
5. The *in vitro* diagnostic method according to any one of the preceding claims, wherein said measuring MLK4 gene expression comprises measuring wild-type and/or mutated MLK4 gene expression.
6. The *in vitro* diagnostic method according to claim 5, wherein said mutated MLK4 gene contains at least one somatic mutation at the amino acid level selected from H261Y, H261Q, G291E, A293E, W296Stp, R470C, R553Stp, R555Stp, N596I, K629E, P843S, H890P, and M894T mutations.
7. The *in vitro* diagnostic method according to claim 5, wherein said mutated MLK4 gene contains at least one somatic mutation at the nucleic acid level selected from C781T, C783G, G872A, C878A, G888A, C1408T, C1657T, C1663T, A1787T, A1885G, C2788T, A2669C, T2681C mutation.
8. The *in vitro* diagnostic method according to any one of the preceding claims, wherein said measuring MLK4 gene expression is performed using a reagent able to selectively bind to MLK4 protein and/or MLK4 coding nucleic acid and wherein said reagent is directly or indirectly labeled with a detectable substance.

9. The *in vitro* diagnostic method according to claim 9, wherein said reagent is an antibody specific for MLK4 protein and/or portions thereof or a nucleic acid probe specific for MLK4 nucleic acid and/or fragments thereof.
10. The *in vitro* diagnostic method according to any one of the preceding claims, wherein said method further comprises measuring at least one KRAS gene molecular alteration in said tumor cell sample material, wherein said measuring at least one KRAS gene alteration comprises measuring KRAS protein and/or RNA expression and/or mutations in the nucleotide sequence of the KRAS gene.
11. The *in vitro* diagnostic method according to any one of the preceding claims, wherein said tumor is a colorectal, bladder, breast, gastric, melanoma, lung, ovary or glioblastoma.
12. The *in vitro* diagnostic method according to any one of claims 5 to 11, wherein measuring mutated MLK4 gene overexpression allows to predict whether said patient will be nonresponsive to treatment with EGFR specific binding agents, preferably EGFR specific antibodies.
13. A method of treating cancer in a patient whose tumor bears a mutations in the nucleotide sequence of the gene KRAS, said method comprising administering an inhibitor of MLK4 enzymatic activity to said patient to reduce or abrogate MLK4 RNA or protein expression.
14. A method of treating cancer in a patient having a tumor, wherein said tumor bears at least one mutation in the KRAS gene encoded protein, and wherein said at least one mutation of the KRAS gene encoded protein is selected among G12S, G12V, G12D, G12A, G12C, G13A and G13D, said method comprising administering an antagonist and/or an inhibitor of MLK4 expression and/or MLK4 enzymatic activity to said patient.
15. Use of an inhibitor of MLK4 enzymatic activity in the manufacture of a medicament for reducing or abrogating MLK4 RNA or protein expression in a patient whose tumor bears a mutation in the nucleotide sequence of the gene KRAS.
16. Use of an antagonist and/or inhibitor of MLK4 expression and/or MLK4 enzymatic activity in the manufacture of a medicament for the treatment of cancer in a

patient whose tumor bears at least one mutation in the KRAS gene encoded protein, and wherein said at least one mutation of the KRAS gene encoded protein is selected among G12S, G12V, G12D, G12A, G12C, G13A and G13D.

17. A method according to any one of claims 1, 13 or 14, substantially as herein described with reference to any one or more of the examples but excluding comparative examples.

18. Use according to claim 15 or claim 16, substantially as herein described with reference to any one or more of the examples but excluding comparative examples.

Figure 1

a

Tumor type	Nt change (cDNA)	Aa change (Protein)	Exon	Mutation type	Conserved residue	Ref
CRC	c. C781T	p. H261Y	1	Missense	x	5
CRC	c. C783G	p. H261Q	1	Missense	x	5
CRC	c. G872A	p. G291E	2	Missense	x	5
CRC	c. C878A	p. A293E	2	Missense	x	5
CRC	c. G888A	p. W296Stp	2	Nonsense	x	5
CRC	c. C1408T	p. R470C	5	Missense	x	5
CRC	c. C1408T	p. R470C	5	Missense	x	5
CRC	c. C1657T	p. R553Stp	6	Nonsense	x	5
GBM	c. C1663T	p. R555Stp	6	Nonsense	x	(This study)
CRC	c. A1787T	p. N596I	7	Missense	x	5
CRC	c. A1885G	p. K629E	8	Missense		5
GBM	c. C2788T	p. P843S	9	Missense		(This study)
GBM	c. A2669C	p. H890P	9	Missense		6
GBM	c. T2681C	p. M894T	9	Missense		6

b

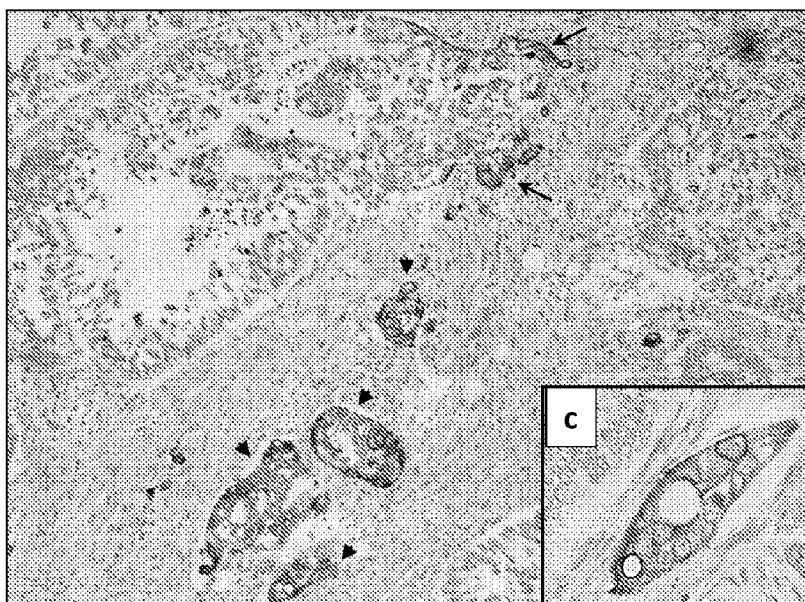


Figure 1/cont.

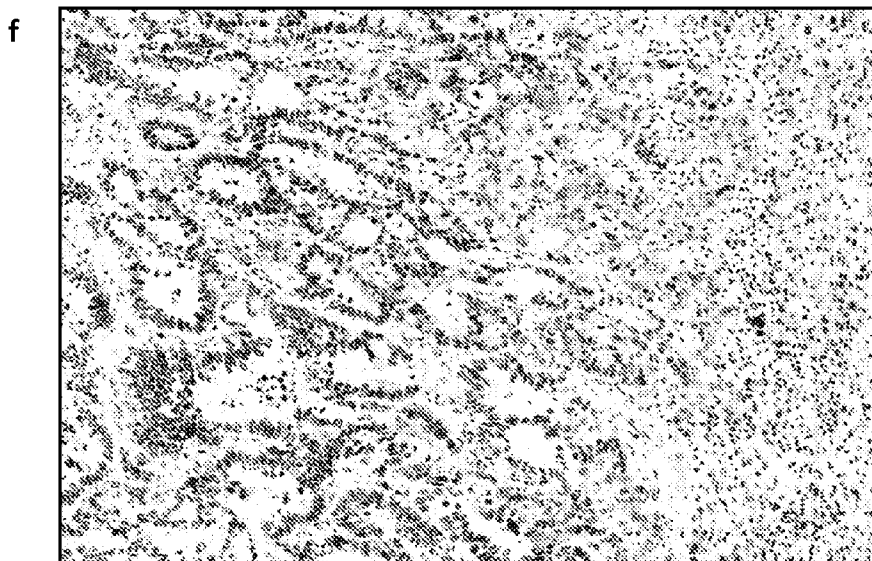
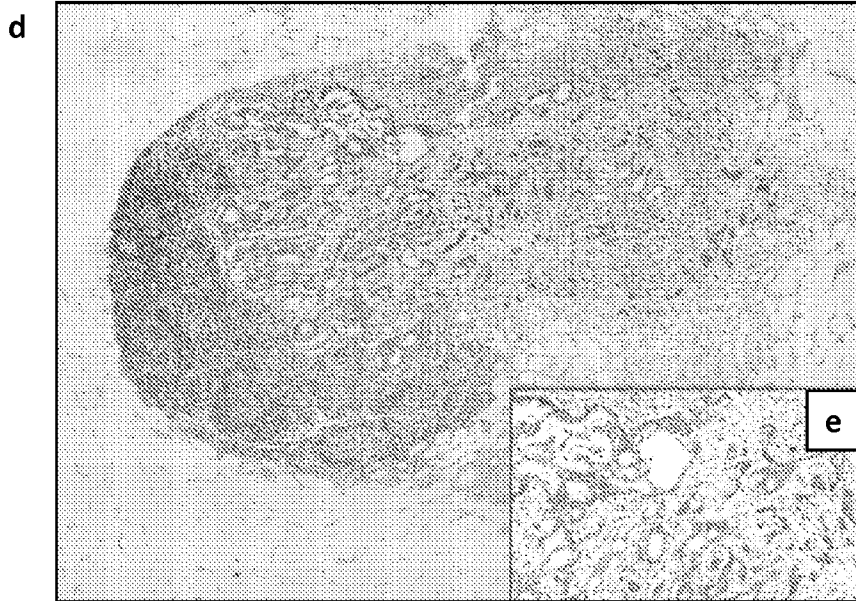


Figure 2

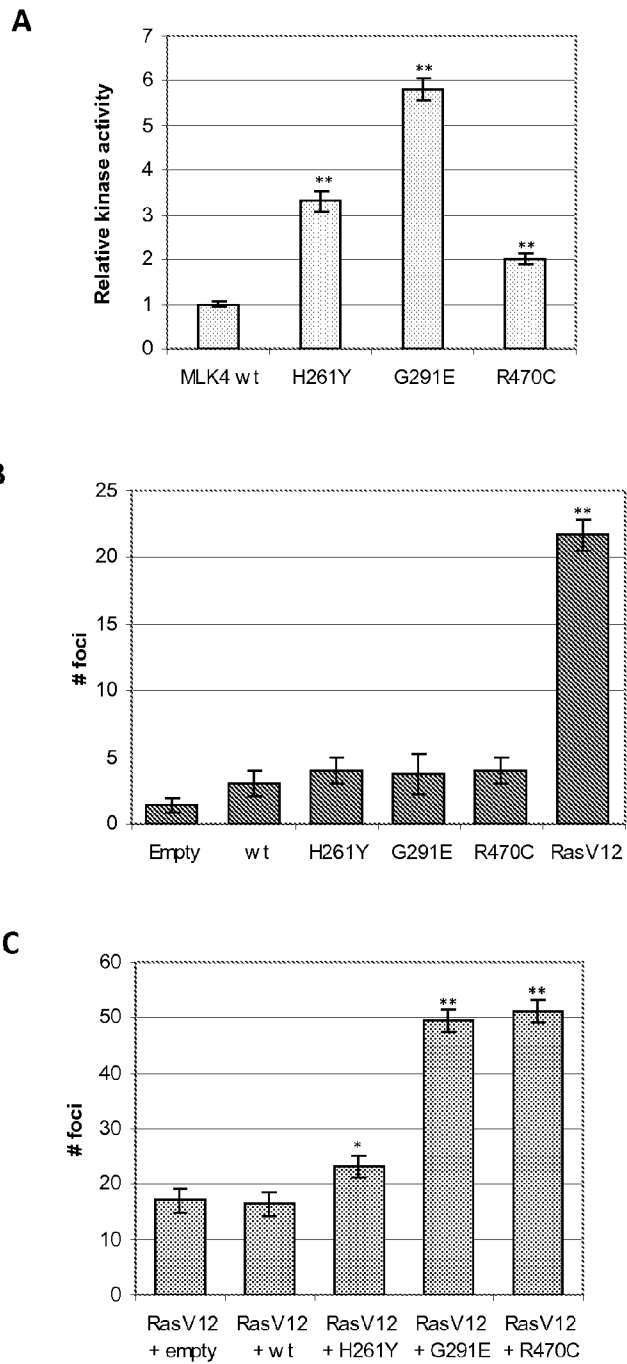
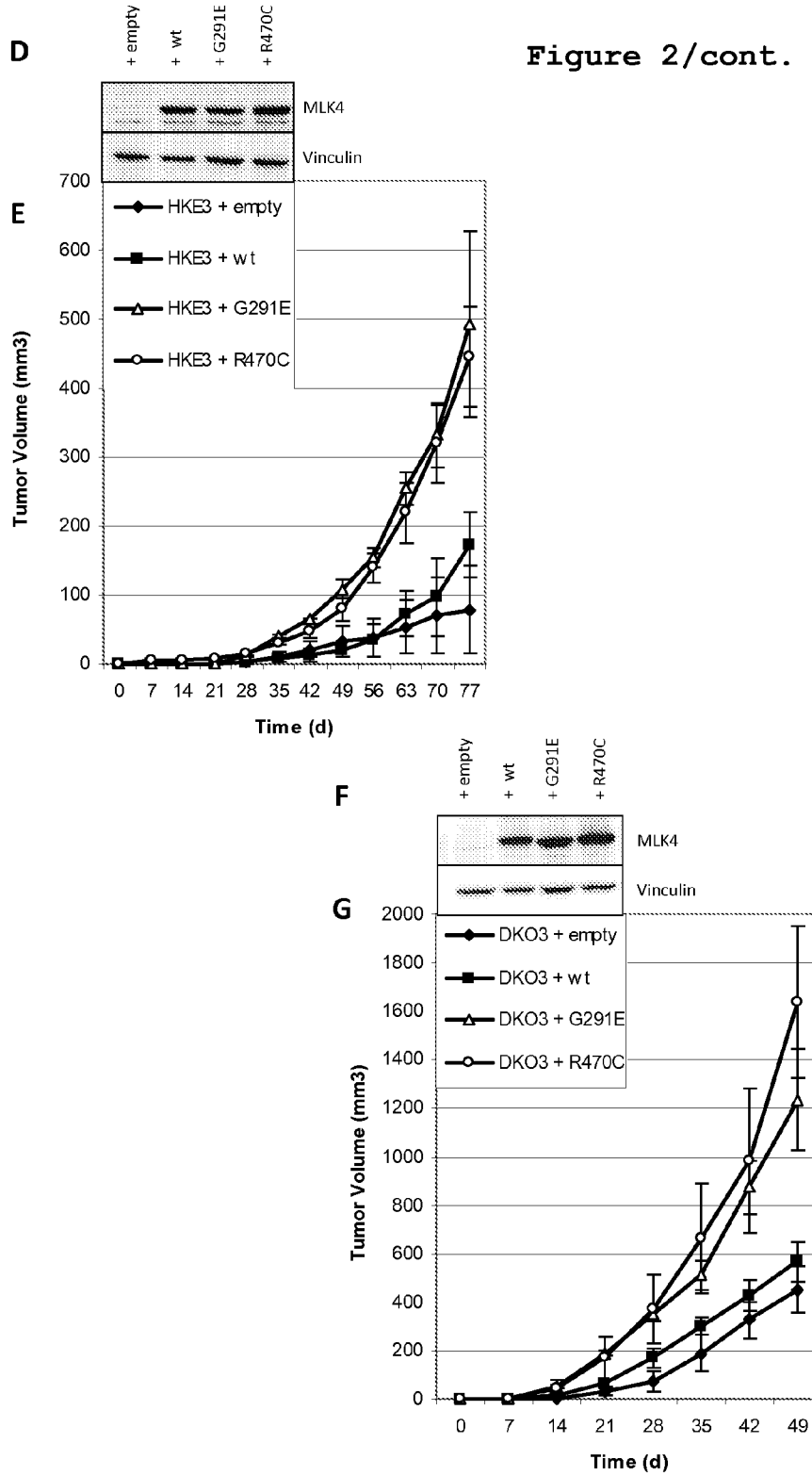
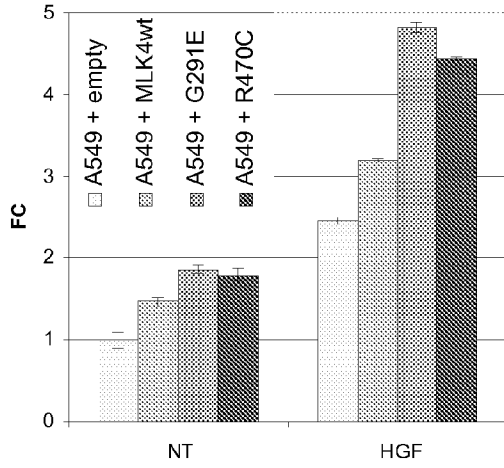


Figure 2/cont.

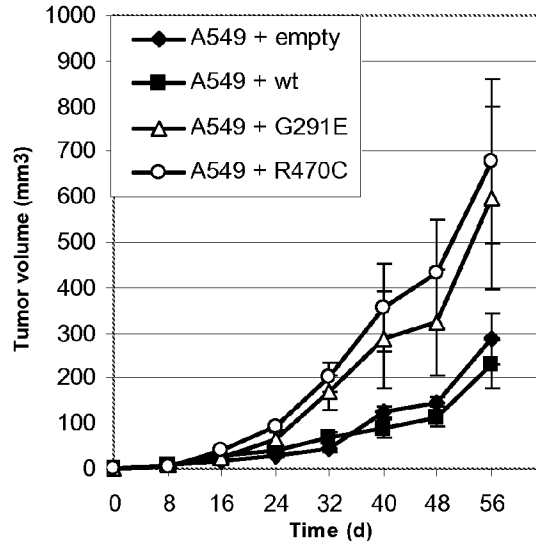


**A**

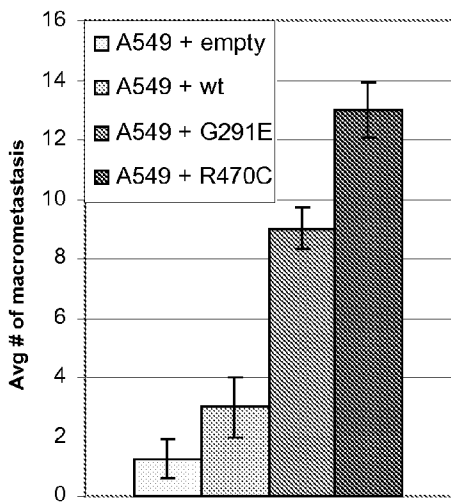


**Figure 3**

**B**



**C**



**D**

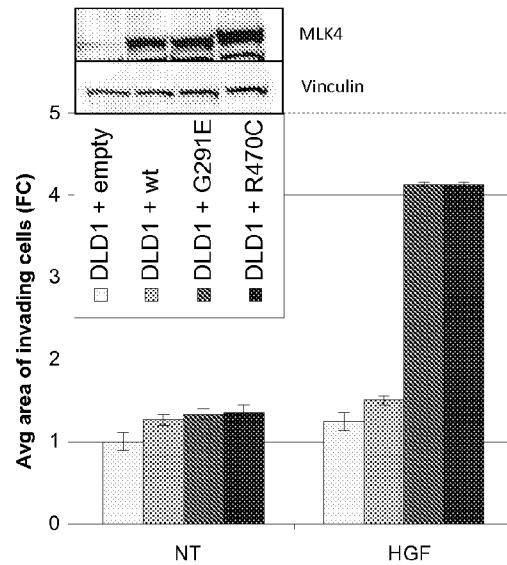
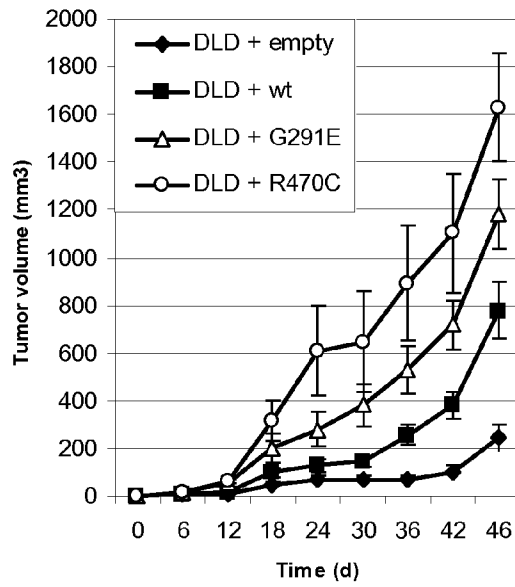


Figure 3/cont.

E



F

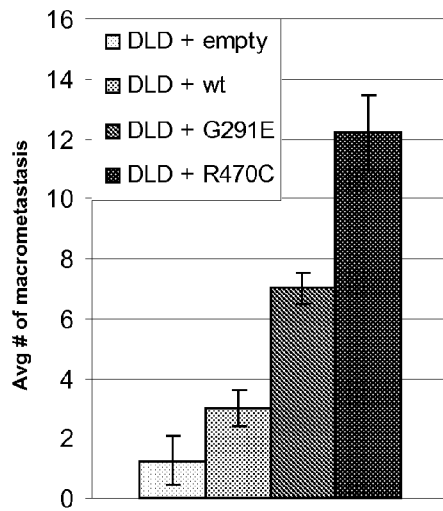


Figure 4

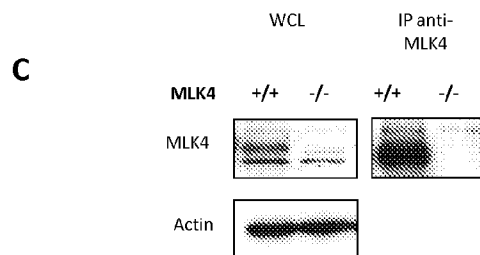
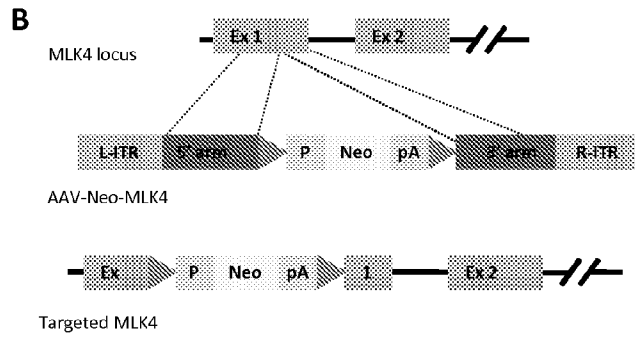
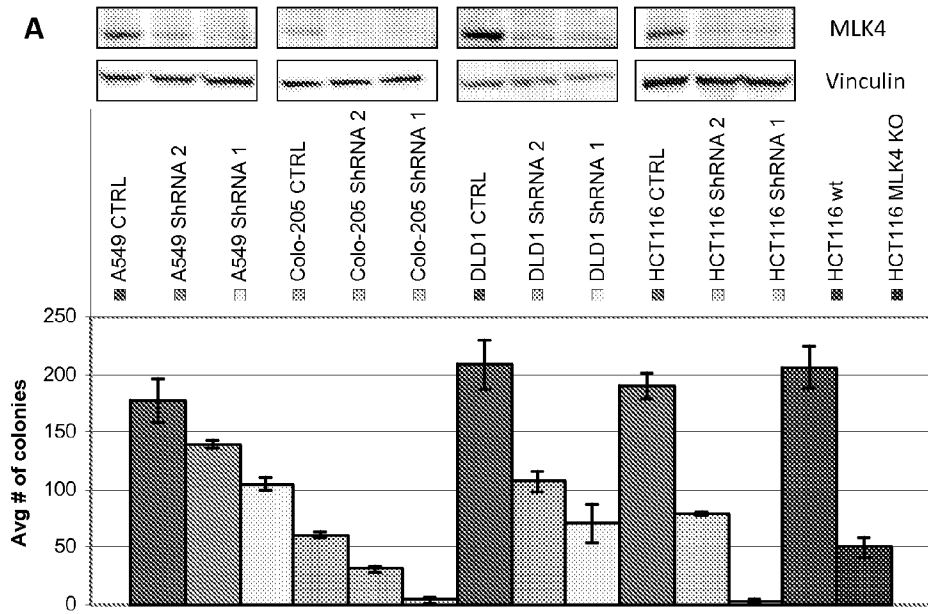


Figure 4/cont.

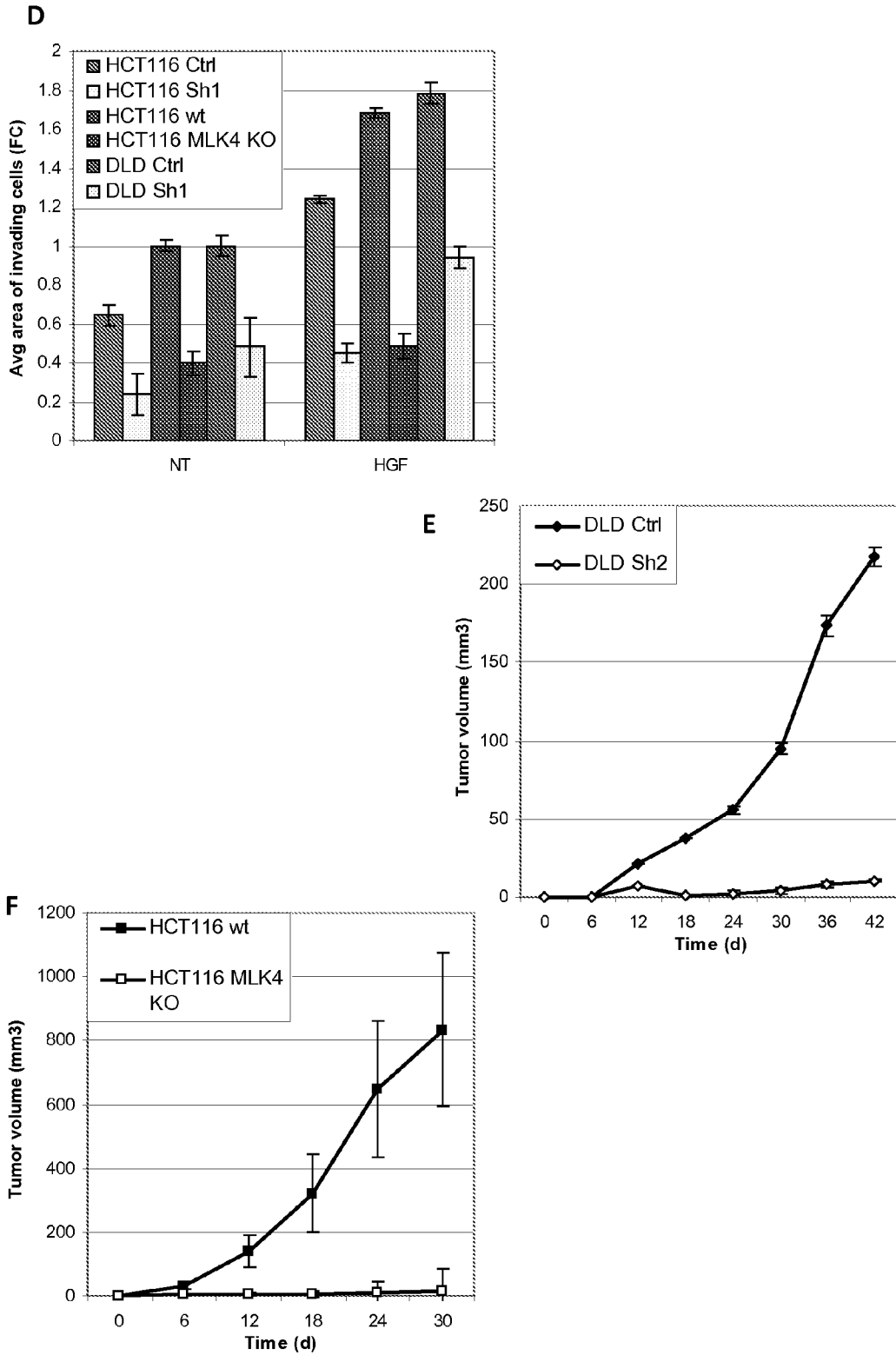
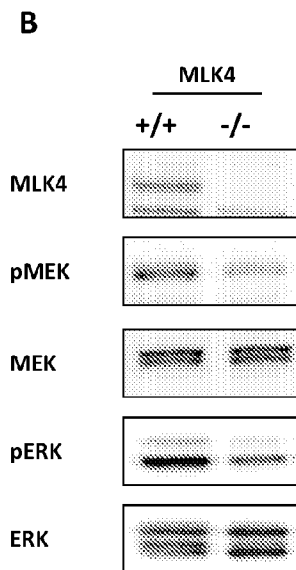
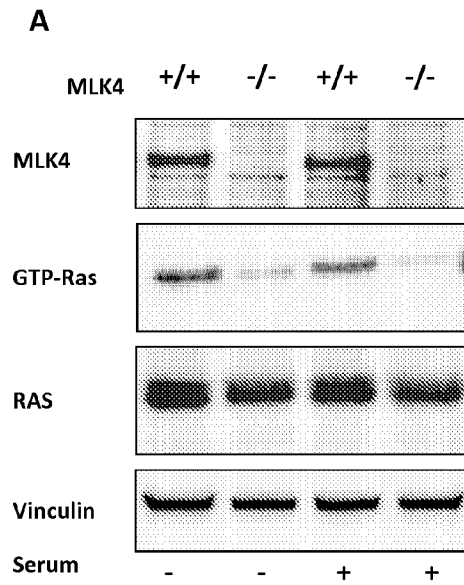


Figure 5



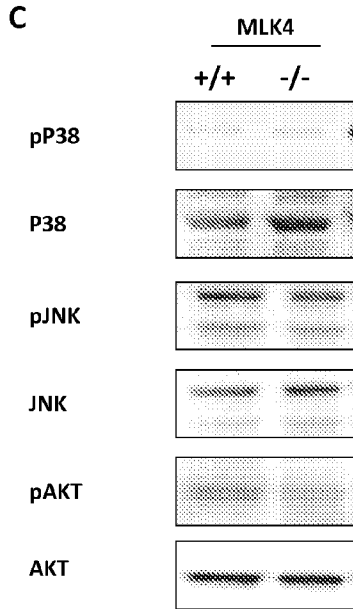
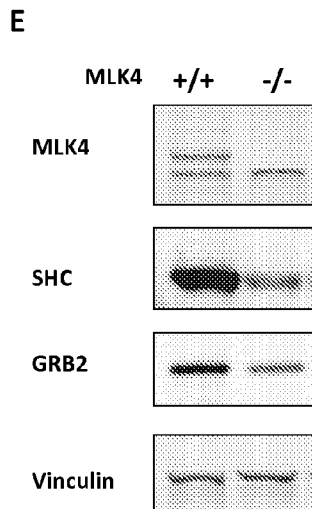
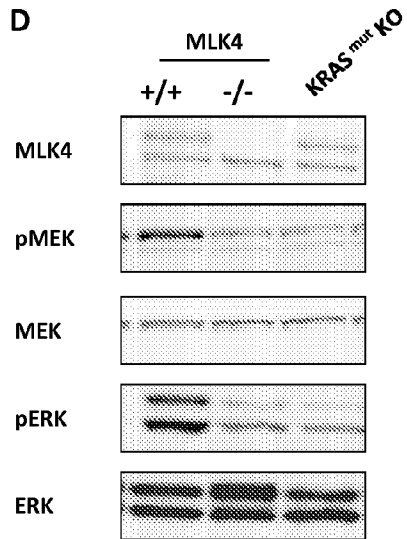
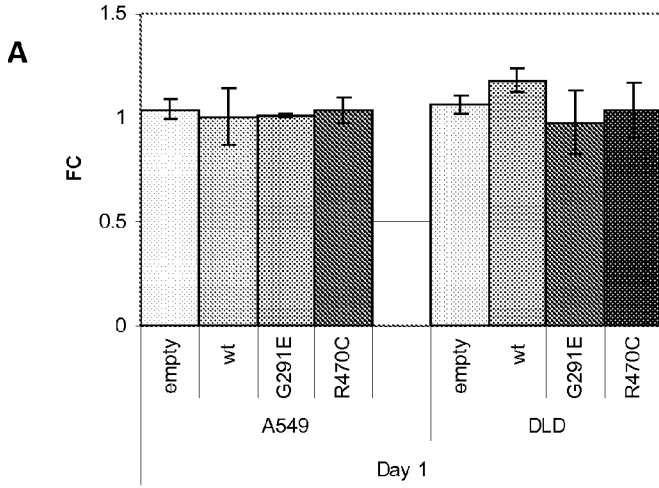


Figure 5/cont.



11/12



**Figure 6**

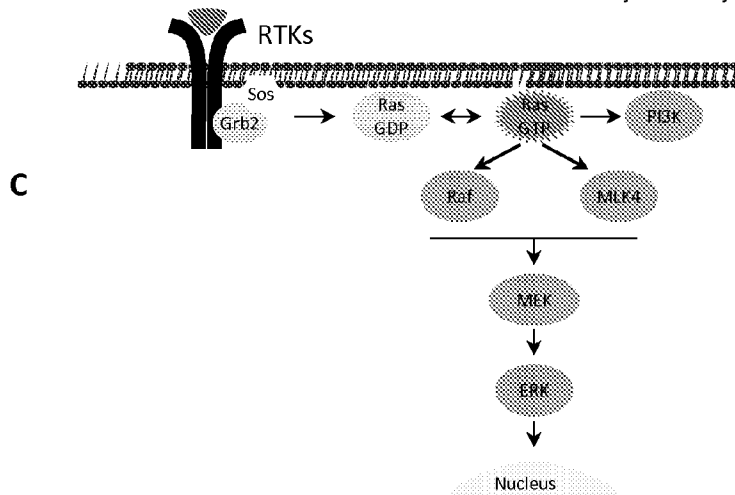
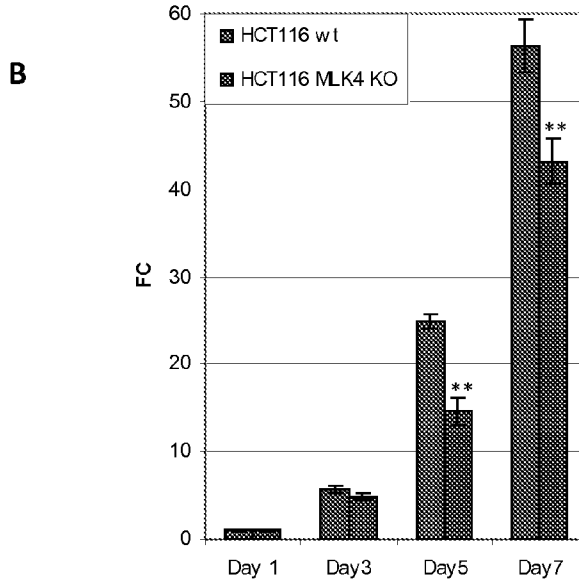
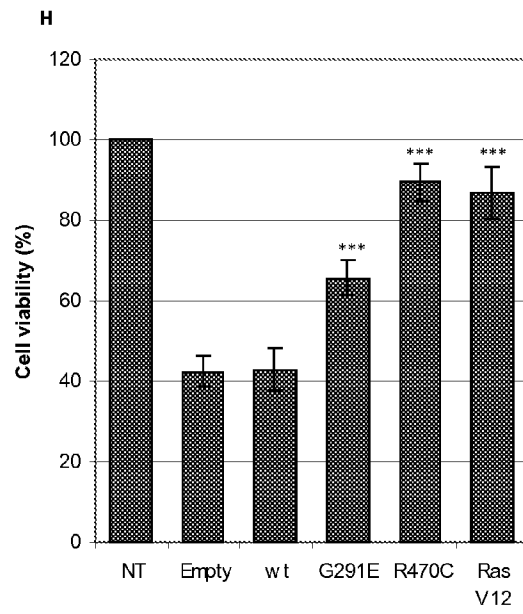


Figure 7



pct i b2010052684- seq1 .txt  
SEQUENCE LISTING

<110> Bardelli, Alberto  
Martini, Miriam

<120> MLK4 gene, a new diagnostic and prognostic marker in cancers

<130> BWO12029-CF

<160> 61

<170> Patent In version 3.5

<210> 1  
<211> 19  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 1\_1

<400> 1  
caccgatcgcgattcctac 19

<210> 2  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 1\_1

<400> 2  
tcgaaggcgcgtgtacc 18

<210> 3  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 1\_1

<400> 3  
tcctacccc tgccttc 18

<210> 4  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 1\_2

<400> 4  
ctcagcgtcc tctgcgtc 18

<210> 5  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 1\_2

<400> 5  
cgcagctcga t gat gtt g 18

<210> 6  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 1\_2

<400> 6  
gcagctcgat gat gtt gg 18

<210> 7  
<211> 19  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 1\_3

<400> 7  
catcttcccc gccaaactac 19

<210> 8  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 1\_3

<400> 8  
ggcccagttg accagcac 18

<210> 9  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 1\_3

<400> 9  
tcttccccgc caactacg 18

<210> 10  
<211> 18  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 1\_4

<400> 10  
caggaggtgg ccgtgaag 18

<210> 11  
<211> 18  
<212> DNA

<213> artificial  
 <220>  
 <223> reverse primer exon 1\_4  
 <400> 11  
 cgccctaagg ctgactcc 18

<210> 12  
 <211> 20  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> sequence primer exon 1\_4  
 <400> 12  
 ctcccacctc tttccatctg 20

<210> 13  
 <211> 18  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> forward primer exon 1\_5  
 <400> 13  
 gcacgtgctg gtcaactg 18

<210> 14  
 <211> 18  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> reverse primer exon 1\_5  
 <400> 14  
 cgccctaagg ctgactcc 18

<210> 15  
 <211> 18  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> sequence primer exon 1\_5  
 <400> 15  
 actgggccgt gcagatag 18

<210> 16  
 <211> 23  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> forward primer exon 2  
 <400> 16  
 caggcacaga taaccacta aag 23

<210> 17  
 <211> 26  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> reverse primer exon 2  
  
 <400> 17  
 aaaccat aaa aattcattca aaaagg 26  
  
 <210> 18  
 <211> 20  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> sequence primer exon 2  
  
 <400> 18  
 cat gccacag aaaaattgtg 20  
  
 <210> 19  
 <211> 21  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> forward primer exon 5  
  
 <400> 19  
 gggaatttca gcttctcctt g 21  
  
 <210> 20  
 <211> 22  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> reverse primer exon 5  
  
 <400> 20  
 ccctatctct ttctctggtc tg 22  
  
 <210> 21  
 <211> 18  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> sequence primer exon 5  
  
 <400> 21  
 ggggtcgcca gtgt agt g 18  
  
 <210> 22  
 <211> 21  
 <212> DNA  
 <213> artificial  
  
 <220>  
 <223> forward primer exon 6\_1

<400> 22  
cgatcctgca atattctttg c 21

<210> 23  
<211> 20  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 6\_1

<400> 23  
gtggcaccta cgaaatgtcc 20

<210> 24  
<211> 22  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 6\_1

<400> 24  
ttgctgtgtg ctaaatacat gg 22

<210> 25  
<211> 25  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 7\_2

<400> 25  
tttctcattg ttttcatggt ttgag 25

<210> 26  
<211> 27  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 7\_2

<400> 26  
gcatgataaa ttcatcctag aaaattg 27

<210> 27  
<211> 26  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 7\_2

<400> 27  
ttttgagtaa ggtagtttca caactg 26

<210> 28  
<211> 20  
<212> DNA

<213> artificial  
 <220>  
 <223> forward primer exon 8  
 <400> 28  
 ttgtcagccc ttcaaaactg 20

<210> 29  
 <211> 20  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> reverse primer exon 8  
 <400> 29  
 cttgctaatc tcccctctgc 20

<210> 30  
 <211> 21  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> sequence primer exon 8  
 <400> 30  
 tggctgtat ccacaaaac c 21

<210> 31  
 <211> 37  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> forward primer exon 9\_1  
 <400> 31  
 gtaaaacgac ggccagt gaa aagggt gca tgtgttt 37

<210> 32  
 <211> 21  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> reverse primer exon 9\_1  
 <400> 32  
 ctctctgagc atctttcca a 21

<210> 33  
 <211> 17  
 <212> DNA  
 <213> artificial  
 <220>  
 <223> sequence primer exon 9\_1  
 <400> 33  
 gtaaaacgac ggccagt 17

<210> 34  
<211> 23  
<212> DNA  
<213> artificial  
  
<220>  
<223> forward primer exon 9\_2  
  
<400> 34  
gcctacattg atctacct tgg 23

<210> 35  
<211> 37  
<212> DNA  
<213> artificial  
  
<220>  
<223> reverse primer exon 9\_2  
  
<400> 35  
gtaaaacgac ggccagtct tctcttcctt gggcaac 37

<210> 36  
<211> 17  
<212> DNA  
<213> artificial  
  
<220>  
<223> sequence primer exon 9\_2  
  
<400> 36  
gtaaaacgac ggccagt 17

<210> 37  
<211> 20  
<212> DNA  
<213> artificial  
  
<220>  
<223> forward primer exon 9\_3  
  
<400> 37  
cataaagcac aggctgctga 20

<210> 38  
<211> 37  
<212> DNA  
<213> artificial  
  
<220>  
<223> reverse primer exon 9\_3  
  
<400> 38  
gtaaaacgac ggccagtagc aggcacttg tggagaa 37

<210> 39  
<211> 17  
<212> DNA  
<213> artificial  
  
<220>  
<223> sequence primer exon 9\_3

<400> 39  
gt aaaacgac ggccagt 17

<210> 40  
<211> 38  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 9\_4

<400> 40  
gt aaaacgac ggccagtcca caccttcttt ctccacaa 38

<210> 41  
<211> 20  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 9\_4

<400> 41  
cat tcccaca t g t c t g c t g t 20

<210> 42  
<211> 17  
<212> DNA  
<213> artificial

<220>  
<223> sequence primer exon 9\_4

<400> 42  
gt aaaacgac ggccagt 17

<210> 43  
<211> 37  
<212> DNA  
<213> artificial

<220>  
<223> forward primer exon 9\_5

<400> 43  
gt aaaacgac ggccagtcaa gaaacttgcc gtcttcc 37

<210> 44  
<211> 26  
<212> DNA  
<213> artificial

<220>  
<223> reverse primer exon 9\_5

<400> 44  
caaattatta aatgtcatca ccagga 26

<210> 45  
<211> 17  
<212> DNA

&lt;213&gt; artificial

&lt;220&gt;

&lt;223&gt; sequence primer exon 9\_5

&lt;400&gt; 45

gt aaaacgac ggccagt

17

&lt;210&gt; 46

&lt;211&gt; 3111

&lt;212&gt; DNA

&lt;213&gt; homo sapiens

&lt;400&gt; 46

at ggct t t gc ggggcgccgc gggagcgacc gacaccccg t gt cct cggc cgggggagcc 60

cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120

gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180

cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggca cgagggct gg 240

t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300

cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccggg acacgt cgcc 360

t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420

gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480

gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540

cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600

gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660

ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacacct c cgcacct gct ggt caact gg 720

gccgt gcaga tagcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780

caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840

t gcaat aaaa ct t t gaagat t acagat t t t ggg t ggca gggat ggca caggaccacc 900

aaaat gagca cagcaggcac ct at gcct gg at ggcccccg aagt gat caa gt ct t cct t g 960

t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020

gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080

ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140

t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200

at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260

t ggaaact ag aat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320

t cccgggaag aggagct gac t cgggaggct ct gcagcaga agt ct cagga ggagct gct a 1380

aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440

at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500

aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560

cacaagat aa ccggt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620

pct i b2010052684- seq1 . t xt

agt t ccagt c ccccgagcag ccccacaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaa t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
 caaat aaaat t gcct agt ca ggct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at gggg gcac cgt cct t ct g gcat cggt gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gcca aggaagagaa gaagaaacga 2280  
 gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
 at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
 ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
 act gccccag gaagt ggt cg t gagccagcc ct cat gcca gact t gacac t gat t gt agt 2580  
 gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaat gt acct t act gt 2640  
 gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
 ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
 cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
 at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cg t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcaggac t aact gt gc cact gt gcag aat gaggagc 3060  
 aaaaccagcc ggccat ct at at at gaact g gagaaagaat t cct gt ct t a a 3111

<210> 47  
 <211> 41  
 <212> DNA  
 <213> artificial

<220>  
 <223> mutagenesis primer G291E

<400> 47  
 ct t t gaagat t acagat t t t gagt t ggcga gggaaat ggca c 41

<210> 48  
 <211> 40  
 <212> DNA  
 <213> artificial

&lt;220&gt;

&lt;223&gt; mutagenesi s pri mer R470C

&lt;400&gt; 48

t g a g c a g c a g c t g g c a g a g t g c g a g a t c g a c g t g c t g g a g 40

&lt;210&gt; 49

&lt;211&gt; 3041

&lt;212&gt; DNA

&lt;213&gt; homo sapi ens

&lt;400&gt; 49

a t g g c t t t g c g g g g c g c c g c g g g a g c g a c c g a c a c c c c c g g t g t c c t c g g c c g g g g g a g a c c 60

c c c g g c g g c t c a g c g t c c t c g t c g t c c a c c t c c t c g g g c g g c t c g g c g g g c g c g 120

g g g c t g t g g g c c g c g c t c t a t g a c t a c g a g g c t c g c g g c g a g g a c g a g c t g a g c c t g c g g 180

c g c g g c c a g c t g g t g g a g g t g c t g t c g c a g a c g c c c g c c g t g t c g g g c g a c g a g g g c t g g 240

t g g g c a g g c c a g g t g c a g c g g c g c g c c t c g g c a t c t t c c c c g c c a a c t a c g t g g c t c c c t g c 300

c g c c c g g c c g c c a g c c c c g c g c g c c g c c t c g c g g c c c a g c t c c c c g g t a c a c g t c g c c 360

t t c g a g c g g c t g g a g c t g a a g g a g c t c a t c g g c g t g g g g g c t t c g g g c a g g t g t a c c g c 420

g c c a c c t g g c a g g g c c a g g a g g t g g c c g t g a a g g c g g c g c g c c a g g a c c c g g a g c a g g a c 480

g c g g c g g c g g c t g c c g a g a g c g t g c g g c g c g a g g c t c g g c t c t t c g c c a t g c t g c g g c a c 540

c c c a a c a t c a t c g a g c t g c g c g g c g t g t g c t g c a g c a g c g c a c c t c t g c c t g g t g c t g 600

g a g t t c g c c c g c g g c g g a g c g c t c a a c c g a g c g c t g g c c g c t g c c a a c g c c g c c c c g g a c 660

c c g c g c g c g c c c g g c c c c c g c c g c g c g c g c g c a t c c c t c c g c a c g t g c t g g t c a a c t g g 720

g c c g t g c a g a t a g c g c g g g g c a t g c t c t a c c t g c a t g a g g a g g c c t c g t g c c c a t c c t g 780

c a c c g g g a c c t c a a g t c c a g c a a c a t t t t g c t a c t t g a g a a g a t a g a a c a t g a t g a c a t c 840

t g c a a t a a a a c t t t g a a g a t t a c a g a t t t t g a g t t g g c g a g g g a a t g g c a c a g g a c c a c c 900

a a a a t g a g c a c a g c a g g c a c c t a t g c c t g g a t g g c c c c c g a a g t g a t c a a g t c t t c c t t g 960

t t t t c t a a g g a a g c g a c a t c t g g a g c t a t g g a g t g c t g c t g t g g g a a c t g c t c a c c g g a 1020

g a a g t c c c c t a t c g g g g c a t t g a t g g c c t c g c c g t g g c t t a t g g g g t a g c a g t c a a t a a a 1080

c t c a c t t t g c c c a t t c c a t c c a c c t g c c c t g a g c c g t t t g c c a a g c t c a t g a a a g a a t g c 1140

t g g c a a c a a g a c c c t c a t a t t c g t c c a t c g t t t g c c t t a a t t c t c g a a c a g t t g a c t g c t 1200

a t t g a a g g g g c a g t g a t g a c t g a g a t g c c t c a a g a a t c t t t t c a t t c c a t g c a a g a t g a c 1260

t g g a a a c t a g a a a t t c a a c a a a t g t t t g a t g a g t t g a g a a c a a a g g a a a a g g a g c t g c g a 1320

t c c c g g g a a g a g g a g c t g a c t c g g g c g g c t c t g c a g c a g a a g t c t c a g g a g g a g c t g c t a 1380

a a g c g g c g t g a g c a g c a g c t g g c a g a g c g c g a g a t c g a c g t g c t g g a g c g g g a a c t t a a c 1440

a t t c t g a t a t t c c a g c t a a a c c a g g a g a a g c c c a a g g t a a a g a a g a g g a a g g g c a a g t t t 1500

a a g a g a a g t c g t t t a a a g c t c a a a g a t g g a c a t c g a a t c a g t t t a c c t t c a g a t t t c c a g 1560

c a c a a g a t a a c c g t g c a g g c c t c t c c c a a c t t g g a c a a a c g g c g g a g c c t g a a c a g c a g c 1620

agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaa t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaaccaaaa 1980  
 caaat aaaat t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
 gaggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
 at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
 ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
 act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
 gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt ggggat gt acct t act gt 2640  
 gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
 ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
 cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
 at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcaggac t acact gt gc c 3041

<210> 50  
 <211> 3041  
 <212> DNA  
 <213> homo sapiens

<400> 50  
 at ggct t t gc ggggcgccc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcgccagc t ggt ggaggt gct gt cgcag gacgccgccc t gt cgggcga cgagggtt gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t cccc ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccccc t cgcggccca gct ccccgtt acacgt cgcc 360  
 t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420

pct i b2010052684- seq1 . t xt

gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggaccc ggagcaggac 480  
 gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcct ccct c cgcacct gct ggt caact gg 720  
 gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t ggggt t ggca ggggat ggca caggaccacc 900  
 aaaat gagca cagcaggcac ct at gcct gg at ggccccg aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
 at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
 t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
 t cccgggaag aggagct gac t cgggcggct ct gcagcaga agt ct cagga ggagct gct a 1380  
 aagcggcgt g agcagcagct ggcagagt gc gagat cgacg t gct ggagcg ggaact t aac 1440  
 at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
 aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
 cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaac ggcggagcct gaacagcagc 1620  
 agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaa t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaaccaaaa 1980  
 caaat aaaat t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at ggggt gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
 gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
 at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460

pct i b2010052684- seq1 . t xt

ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gcca gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaaat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcaggac t aact gt gc c 3041

<210> 51  
<211> 3041  
<212> DNA  
<213> homo sapi ens

<400> 51  
at ggct t t gc ggggcgccgc gggagcgacc gacaccccg t gt cct cggc cgggggagcc 60  
cccggcggt cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
gggct gt ggg ccgcgt ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
cgcgccagc t ggt ggaggt gct gt cgcag gacgccgccc t gt cgggca cgagggt gg 240  
t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
cgccccggccg ccagccccgc gccgcccgc t cgcggccca gct ccccgt acacgt cgcc 360  
t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
gcggcgggcg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
ccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
ccgcgcgcgc ccggccccgc ccgcgcgcgc cgcacct cct c cgcacgt gct ggt caact gg 720  
gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
t accgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
t gcaat aaaa ct t t gaagat t acagat t t t gggg t ggcga gggaaat ggca caggaccacc 900  
aaaat gagca cagcaggcac ct at gcct gg at ggcccccg aagt gat caa gt ct t cct t g 960  
t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaaact gct caccgga 1020  
gaagt cccct at cggggcat t gat ggct c gccgt ggct t at ggggt agc agt caat aaa 1080  
ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200

at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
t cccgggaag aggagct gac t cgggcggt ct gcagcaga agt ct cagga ggagct gct a 1380  
aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
gat gt aaaaa ggaat t t t aa gaaaaaagg t gt acct ggg gaccaa t t c cat t caaat g 1800  
aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
caaat aaaa t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaaaggag agaact aat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcaggac t acact gt gc c 3041

<210> 52  
<211> 3041  
<212> DNA  
<213> homo sapi ens

pct i b2010052684- seq1 . t xt

<400> 52  
at ggct t t gc ggggcgccgc gggagcgacc gacaccccg t gt cct cggc cgggggagcc 60  
cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcga cgagggt gg 240  
t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgt acacgt cgcc 360  
t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacat ccct c cgcacgt gct ggt caact gg 720  
gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
cagcgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
t gcaat aaaa ct t t gaagat t acagat t t t ggg t ggcga gggaat ggca caggaccacc 900  
aaaat gagca cagcaggcac ct at gcct gg at ggccccg aagt gat caa gt ct t cct t g 960  
t t t t ct aagg gaagcgacat ct ggagct at ggag t gct gc t gt gggaact gct caccgga 1020  
gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
t cccgggaag aggagct gac t cgggcggt ct gcagcaga agt ct cagga ggagct gct a 1380  
aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
gat gt aaaaa ggaat t t t aa gaaaaaagg t gt acct ggg gaccaa t t c cat t caaat g 1800  
aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
caaat aaaa t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040

pct i b2010052684- seq1 . t xt

gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcagggac t aact gt gc c 3041

<210> 53  
<211> 3041  
<212> DNA  
<213> homo sapi ens

<400> 53  
at ggct t t gc ggggcccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
cccggcggt cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcccgc 120  
gggct gt ggg ccgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccc t gt cgggcca cgagggt gg 240  
t gggcagggc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
cgcccggccc ccagccccgc gccgccccc t cgcggccca gct ccccggg acacgt cgcc 360  
t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacat ccct c cgcacgt gct ggt caact gg 720  
gccgt gcaga t agcgcgggg cat gct ct ac ct gcacat gagg aggcct t cgt gcccat cct g 780

caccgggacc	t caagt ccag	caacat t t t g	ct act t gaga	agat agaaca	t gat gacat c	840
t gcaat aaaa	ct t t gaagat	t acagat t t t	gggt t ggaga	gggaat ggca	caggaccacc	900
aaaat gagca	cagcaggcac	ct at gcct gg	at ggcccccg	aagt gat caa	gt ct t cct t g	960
t t t t ct aagg	gaagcgacat	ct ggagct at	ggagt gct gc	t gt gggaact	gct caccgga	1020
gaagt cccct	at cggggcat	t gat ggcct c	gccgt ggct t	at ggggt agc	agt caat aaa	1080
ct cact t t gc	ccat t ccat c	cacct gcct	gagccgt t t g	ccaagct cat	gaaagaat gc	1140
t ggcaacaag	accct cat at	t cgt ccat cg	t t t gcct t aa	t t ct cgaaca	gt t gact gct	1200
at t gaagggg	cagt gat gac	t gagat gcct	caagaat ct t	t t cat t ccat	gcaagat gac	1260
t ggaaact ag	aaat t caaca	aat gt t t gat	gagt t gagaa	caaagggaaa	ggagct gcga	1320
t cccgggaag	aggagct gac	t cgggcccgt	ct gcagcaga	agt ct cagga	ggagct gct a	1380
aagcggcgt g	agcagcagct	ggcagagcgc	gagat cgacg	t gct ggagcg	ggaact t aac	1440
at t ct gat at	t ccagct aaa	ccaggagaag	cccaaggt aa	agaagaggaa	gggcaagt t t	1500
aagagaagt c	gt t t aaagct	caaagat gga	cat cgaat ca	gt t t acct t c	agat t t ccag	1560
cacaagat aa	ccgt gcaggc	ct ct cccaac	t t ggacaaac	ggcggagcct	gaacagcagc	1620
agt t ccagt c	ccccgagcag	ccccacaat g	at gccccgac	t ccgagccat	acagt t gact	1680
t cagat gaaa	gcaat aaaac	t t ggggaagg	aacacagt ct	t t cgacaaga	agaat t t gag	1740
gat gt aaaaa	ggaat t t t aa	gaaaaaaggt	t gt acct ggg	gaccaaat t c	cat t caaat g	1800
aaagat agaa	cagat t gcaa	agaaaggat a	agacct ct ct	ccgat ggcaa	cagt cct t gg	1860
t caact at ct	t aat aaaaaa	t cagaaaacc	at gccct t gg	ct t cat t gt t	t gt ggaccag	1920
ccagggt cct	gt gaagagcc	aaaact t t cc	cct gat ggat	t agaacacag	aaaaccaaaa	1980
caaat aaaat	t gcct agt ca	ggcct acat t	gat ct acct c	t t gggaaaga	t gct cagaga	2040
gagaat cct g	cagaagct ga	aagct gggag	gaggcagcct	ct gcgaat gc	t gccacagt c	2100
t ccat t gaga	t gact cct ac	gaat agt ct g	agt agat ccc	cccagagaaa	gaaaacggag	2160
t cagct ct gt	at ggggt gcac	cgt cct t ct g	gcat cggg gg	ct ct gggact	ggacct caga	2220
gagct t cat a	aagcacaggc	t gct gaagaa	ccgt t gccca	aggaagagaa	gaagaaacga	2280
gagggaat ct	t ccagcgggc	t t ccaagt cc	cgcagaagcg	ccagt cct cc	cacaagcct g	2340
ccat ccacct	gt ggggaggc	cagcagccca	ccct ccct gc	cact gt caag	t gccct gggc	2400
at cct ct cca	cacct t ct t t	ct ccacaaag	t gcct gct gc	agat ggacag	t gaagat cca	2460
ct ggt ggaca	gt gcacct gt	cact t gt gac	t ct gagat gc	t cact ccgga	t t t t t gt ccc	2520
act gccccag	gaagt ggt cg	t gagccagcc	ct cat gccaa	gact t gacac	t gat t gt agt	2580
gt at caagaa	act t gccgt c	t t cct t cct a	cagcagacat	gt gggaat gt	acct t act gt	2640
gct t ct t caa	aacat agacc	gt cacat cac	agacggacca	t gt ct gat gg	aaat ccgacc	2700
ccaact ggt g	caact at t at	ct cagccact	ggagcct ct g	cact gccact	ct gccct ca	2760
cct gct cct c	acagt cat ct	gccaagggag	gt ct caccca	agaagcacag	cact gt ccac	2820

at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcagggac t aact gt gc c 3041

<210> 54  
 <211> 3041  
 <212> DNA  
 <213> homo sapi ens

<400> 54  
 at ggct t t gc ggggcgccgc gggagcgacc gacaccccg t gt cct cggc cgggggagcc 60  
 cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcca cgagggct gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t cccc ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgg acacgt cgcc 360  
 t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
 gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
 gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacat ccct c cgcacgt gct ggt caact gg 720  
 gccgt gcaga tagcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t gggg t ggcca gggaaat gaca caggaccacc 900  
 aaaat gagca cagcaggcac ct at gcct gg at ggcccccg aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
 at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
 t ggaaact ag aat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
 t cccgggaag aggagct gac t cgggaggct ct gcagcaga agt ct cagga ggagct gct a 1380  
 aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
 at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
 aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
 cacaagat aa cctg gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620

pct i b2010052684- seq1 . t xt

agt t ccagt c ccccgagcag ccccacaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaa t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
 caaat aaaat t gcct agt ca ggct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at gggg gcac cgt cct t ct g gcat cggt gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gcca aggaagagaa gaagaaacga 2280  
 gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
 at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
 ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
 act gccccag gaagt ggt cg t gagccagcc ct cat gcca gact t gacac t gat t gt agt 2580  
 gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaat gt acct t act gt 2640  
 gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
 ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
 cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
 at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cg t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcagggac t aact gt gc c 3041

<210> 55  
 <211> 3041  
 <212> DNA  
 <213> homo sapi ens

<400> 55  
 at ggct t t gc ggggcccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcccgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccc t gt cgggcca cgagggt gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t cccc ccaact acgt ggct ccct gc 300  
 cccccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgt acacgt cgcc 360

t t cgagcggc	t ggagct gaa	ggagct cat c	ggcgct gggg	gct t cgggca	ggt gt accgc	420
gccacct ggc	agggccagga	ggt ggccgt g	aaggcggcgc	gccaggaccc	ggagcaggac	480
gcggcggcgg	ct gccgagag	cgt gcggcgc	gaggct cggc	t ct t cgccat	gct gcggcac	540
cccaacat ca	t cgagct gcg	cggcgt gt gc	ct gcagcagc	cgcacct ct g	cct ggt gct g	600
gagt t cgccc	gcggcggagc	gct caaccga	gcgct ggccg	ct gccaacgc	cgccccggac	660
ccgcgcgcgc	ccggcccccg	ccgcgcgcgc	cgcat ccct c	cgcacgt gct	ggt caact gg	720
gccgt gcaga	t agcgcgggg	cat gct ct ac	ct gcat gagg	aggcct t cgt	gccccat cct g	780
caccgggacc	t caagt ccag	caacat t t t g	ct act t gaga	agat agaaca	t gat gacat c	840
t gcaat aaaa	ct t t gaagat	t acagat t t t	gggt t ggcga	gggaat ggca	caggaccacc	900
aaaat gagca	cagcaggcac	ct at gcct gg	at ggcccccg	aagt gat caa	gt ct t cct t g	960
t t t t ct aagg	gaagcgacat	ct ggagct at	ggagt gct gc	t gt gggaact	gct caccgga	1020
gaagt cccct	at cggggcat	t gat ggcct c	gccgt ggct t	at ggggt agc	agt caat aaa	1080
ct cact t t gc	ccat t ccat c	cacct gccct	gagccgt t t g	ccaagct cat	gaaagaat gc	1140
t ggcaacaag	accct cat at	t cgt ccat cg	t t t gcct t aa	t t ct cgaaca	gt t gact gct	1200
at t gaagggg	cagt gat gac	t gagat gcct	caagaat ct t	t t cat t ccat	gcaagat gac	1260
t ggaaact ag	aaat t caaca	aat gt t t gat	gagt t gagaa	caaaggaaaa	ggagct gcga	1320
t cccgggaag	aggagct gac	t cgggcggct	ct gcagcaga	agt ct cagga	ggagct gct a	1380
aagcggcgt g	agcagcagct	ggcagagcgc	gagat cgacg	t gct ggagcg	ggaact t aac	1440
at t ct gat at	t ccagct aaa	ccaggagaag	cccaaggt aa	agaagaggaa	gggcaagt t t	1500
aagagaagt c	gt t t aaagct	caaagat gga	cat cgaat ca	gt t t acct t c	agat t t ccag	1560
cacaagat aa	ccgt gcaggc	ct ct cccaac	t t ggacaaac	ggcggagcct	gaacagcagc	1620
agt t ccagt c	ccccgagcag	ccccacaat g	at gccct gac	t ccgagccat	acagt t gact	1680
t cagat gaaa	gcaat aaaac	t t ggggaagg	aacacagt ct	t t cgacaaga	agaat t t gag	1740
gat gt aaaaa	ggaat t t t aa	gaaaaaaggt	t gt acct ggg	gaccaaat t c	cat t caaat g	1800
aaagat agaa	cagat t gcaa	agaaaggat a	agacct ct ct	ccgat ggcaa	cagt cct t gg	1860
t caact at ct	t aat aaaaaa	t cagaaaacc	at gccct t gg	ct t cat t gt t	t gt ggaccag	1920
ccagggt cct	gt gaagagcc	aaaact t t cc	cct gat ggat	t agaacacag	aaaacaaaaa	1980
caaat aaaat	t gcct agt ca	ggcct acat t	gat ct acct c	t t gggaaaga	t gct cagaga	2040
gagaat cct g	cagaagct ga	aagct gggag	gaggcagcct	ct gcgaat gc	t gccacagt c	2100
t ccat t gaga	t gact cct ac	gaat agt ct g	agt agat ccc	cccagagaaa	gaaaacggag	2160
t cagct ct gt	at gggg gcac	cgt cct t ct g	gcat cggt gg	ct ct gggact	ggacct caga	2220
gagct t cat a	aagcacaggc	t gct gaagaa	ccgt t gccca	aggaagagaa	gaagaaacga	2280
gagggaat ct	t ccagcgggc	t t ccaagt cc	cgcagaagcg	ccagt cct cc	cacaagcct g	2340
ccat ccacct	gt ggggaggc	cagcagccca	ccct ccct gc	cact gt caag	t gccct gggc	2400

at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
 ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
 act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
 gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaat gt acct t act gt 2640  
 gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
 ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
 cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
 at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccaccc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcagggac t acact gt gc c 3041

<210> 56  
 <211> 3041  
 <212> DNA  
 <213> homo sapi ens

<400> 56  
 at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcga cgagggct gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgg acacgt cgcc 360  
 t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
 gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
 gcggcggcg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggccccgc ccgcgcgcgc cgcacat cct c cgcacgt gct ggt caact gg 720  
 gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t gggg t ggcga gggaaat ggca caggaccacc 900  
 aaaaat gagca cagcaggcac ct at gcct gg at ggccccgc aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggccct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag acct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200

pct i b2010052684- seq1 . t xt

at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
t cccgggaag aggagct gac t cgggctggct ct gcagcaga agt ct cagga ggagct gct a 1380  
aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ct gagccat acagt t gact 1680  
t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
gat gt aaaaa ggaat t t t aa gaaaaaagg t gt acct ggg gaccaa t t c cat t caaat g 1800  
aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaaccaaaa 1980  
caaat aaaa t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
gagggaaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaaat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaaaggag agaact aat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcagggac t acact gt gc c 3041

<210> 57  
<211> 3041  
<212> DNA  
<213> homo sapiens

<400> 57  
 at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggt cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcgccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggca cgagggct gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgt acacgt cgcc 360  
 t t cgagcggc t ggagct gaa ggagct cat c ggcgt gggg gct t cgggca ggt gt accgc 420  
 gccacct ggc agggccagga ggt ggccgt g aaggcggcg gccaggacc ggagcaggac 480  
 gcggcgggcg ct gccgagag cgt gcggcg gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggccccg ccgcgcgcgc cgcac cct c cgcacgt gct ggt caact gg 720  
 gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t ggg t ggcga gggaaat ggca caggaccacc 900  
 aaaat gagca cagcaggcac ct at gcct gg at ggccccg aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
 at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
 t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
 t cccgggaag aggagct gac t cgggcggt ct gcagcaga agt ct cagga ggagct gct a 1380  
 aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
 at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
 aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
 cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
 agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaggt t gt acct ggg gaccaat t t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980

caaat aaaat t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
gaggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcaggac t acact gt gc c 3041

<210> 58  
<211> 3041  
<212> DNA  
<213> homo sapi ens

<400> 58  
at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
cgcgccagc t ggt ggaggt gct gt cgcag gacgccgccc t gt cgggca cgagggct gg 240  
t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
cgcccggcgc ccagccccgc gccgcccgc t cgcggccca gct ccccgt acacgt cgcc 360  
t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
ccgcgcgcgc ccggccccgc ccgcgcgcgc cgcacat cct c cgcacgt gct ggt caact gg 720  
gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780

pct i b2010052684- seq1 . t xt

caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
t gcaat aaaa ct t t gaagat t acagat t t t ggg t ggcga gggaat ggca caggaccacc 900  
aaaat gagca cagcaggcac ct at gcct gg at ggcccccg aagt gat caa gt ct t cct t g 960  
t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
t cccgggaag aggagct gac t cgggaggct ct gcagcaga agt ct cagga ggagct gct a 1380  
aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaa t c cat t caaat g 1800  
aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
t caact at ct t aat aaaaa t caggaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaaccaaaa 1980  
caaat aaaat t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
t cagct ct gt at ggg t gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820

pct i b2010052684- seq1 . t xt

at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cgc t gccaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcaggac t aact gt gc c 3041

<210> 59  
 <211> 3041  
 <212> DNA  
 <213> homo sapi ens

<400> 59  
 at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggt cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcga cgagggt gg 240  
 t gggcagggc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgg acacgt cgcc 360  
 t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
 gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
 gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggccccgc ccgcgcgcgc cgcacct cct c cgcacgt gct ggt caact gg 720  
 gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t gggg t ggcga ggggat ggca caggaccacc 900  
 aaaat gagca cagcaggcac ct at gcct gg at ggccccg aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
 at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
 t ggaaact ag aaat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
 t cccgggaag aggagct gac t cgggcggt ct gcagcaga agt ct cagga ggagct gct a 1380  
 aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
 at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
 aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560

cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
 agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaaggt t gt acct ggg gaccaaat t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
 caaat aaaat t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at gggg gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
 gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400  
 at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
 ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
 act gcct cag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
 gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggaaat gt acct t act gt 2640  
 gct t ct t caa aacat agacc gt cacat cac agacggacca t gt ct gat gg aaat ccgacc 2700  
 ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
 cct gct cct c acagt cat ct gccaaaggag gt ct caccca agaagcacag cact gt ccac 2820  
 at cgt gcct c agcgt cgccc t gcct ccct g agaagccgct cagat ct gcc t caggct t ac 2880  
 ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggg cg cccaggacca 2940  
 cat cccacc aat t cct cgc t gccaaaggag agaact aaat cccat gt gcc t t cat t act g 3000  
 gat gct gacg t ggaaggt ca gagcaggac t aact gt gc c 3041

<210> 60  
 <211> 3041  
 <212> DNA  
 <213> homo sapi ens

<400> 60  
 at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
 cccggcggt c agcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
 gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gagcct gcgg 180  
 cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcga cgagggt gg 240  
 t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
 cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgg acacgt cgcc 360

pct i b2010052684- seq1 . t xt

t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
 gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc gccagcaggac 480  
 gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
 cccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
 gagt t cgccc gcggcgggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
 ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacat ccct c cgcacgt gct ggt caact gg 720  
 gccgt gcaga tagcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
 caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
 t gcaat aaaa ct t t gaagat t acagat t t t ggg t ggcga gggaat ggca caggaccacc 900  
 aaaaat gagca cagcaggcac ct at gcct gg at ggccccg aagt gat caa gt ct t cct t g 960  
 t t t t ct aagg gaagcgacat ct ggagct at ggagt gct gc t gt gggaact gct caccgga 1020  
 gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
 ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140  
 t ggcaacaag accct cat at t cgt ccat cg t t t gcct t aa t t ct cgaaca gt t gact gct 1200  
 at t gaagggg cagt gat gac t gagat gcct caagaat ct t t t cat t ccat gcaagat gac 1260  
 t ggaaact ag aat t caaca aat gt t t gat gagt t gagaa caaaggaaaa ggagct gcga 1320  
 t cccgggaag aggagct gac t cgggcggct ct gcagcaga agt ct cagga ggagct gct a 1380  
 aagcggcgt g agcagcagct ggcagagcgc gagat cgacg t gct ggagcg ggaact t aac 1440  
 at t ct gat at t ccagct aaa ccaggagaag cccaaggt aa agaagaggaa gggcaagt t t 1500  
 aagagaagt c gt t t aaagct caaagat gga cat cgaat ca gt t t acct t c agat t t ccag 1560  
 cacaagat aa ccgt gcaggc ct ct cccaac t t ggacaaac ggcggagcct gaacagcagc 1620  
 agt t ccagt c ccccgagcag ccccaaat g at gccccgac t ccgagccat acagt t gact 1680  
 t cagat gaaa gcaat aaaac t t ggggaagg aacacagt ct t t cgacaaga agaat t t gag 1740  
 gat gt aaaaa ggaat t t t aa gaaaaaagg t gt acct ggg gaccaa t t c cat t caaat g 1800  
 aaagat agaa cagat t gcaa agaaaggat a agacct ct ct ccgat ggcaa cagt cct t gg 1860  
 t caact at ct t aat aaaaa t cagaaaacc at gccct t gg ct t cat t gt t t gt ggaccag 1920  
 ccagggt cct gt gaagagcc aaaact t t cc cct gat ggat t agaacacag aaaacaaaa 1980  
 caaat aaaa t gcct agt ca ggcct acat t gat ct acct c t t gggaaaga t gct cagaga 2040  
 gagaat cct g cagaagct ga aagct gggag gaggcagcct ct gcgaat gc t gccacagt c 2100  
 t ccat t gaga t gact cct ac gaat agt ct g agt agat ccc cccagagaaa gaaaacggag 2160  
 t cagct ct gt at ggg t gcac cgt cct t ct g gcat cggg gg ct ct gggact ggacct caga 2220  
 gagct t cat a aagcacaggc t gct gaagaa ccgt t gccca aggaagagaa gaagaaacga 2280  
 gagggaat ct t ccagcgggc t t ccaagt cc cgcagaagcg ccagt cct cc cacaagcct g 2340  
 ccat ccacct gt ggggaggc cagcagccca ccct ccct gc cact gt caag t gccct gggc 2400

pct i b2010052684- seq1 . t xt

at cct ct cca cacct t ct t t ct ccacaaag t gcct gct gc agat ggacag t gaagat cca 2460  
ct ggt ggaca gt gcacct gt cact t gt gac t ct gagat gc t cact ccgga t t t t t gt ccc 2520  
act gccccag gaagt ggt cg t gagccagcc ct cat gccaa gact t gacac t gat t gt agt 2580  
gt at caagaa act t gccgt c t t cct t cct a cagcagacat gt gggat gt acct t act gt 2640  
gct t ct t caa aacat agacc gt cacat ccc agacggacca t gt ct gat gg aaat ccgacc 2700  
ccaact ggt g caact at t at ct cagccact ggagcct ct g cact gccact ct gccct ca 2760  
cct gct cct c acagt cat ct gccaggag gt ct caccca agaagcacag cact gt ccac 2820  
at cgt gcct c agcgt cgccc t gcct ccct g agaagccgt cagat ct gcc t caggct t ac 2880  
ccacagacag cagt gt ct ca gct ggcacag act gcct gt g t agt gggc cg cccaggacca 2940  
cat cccacc aat t cct cgc t gccaggag agaact aaat cccat gt gcc t t cat t act g 3000  
gat gct gacg t ggaaggt ca gagcaggac t aact gt gc c 3041

<210> 61  
<211> 3041  
<212> DNA  
<213> homo sapi ens

<400> 61  
at ggct t t gc ggggcgccgc gggagcgacc gacaccccgg t gt cct cggc cgggggagcc 60  
cccggcggct cagcgt cct c gt cgt ccacc t cct cgggcg gct cggcct c ggcgggcgcg 120  
gggct gt ggg ccgcgct ct a t gact acgag gct cgcggcg aggacgagct gacgct gcgg 180  
cgcggccagc t ggt ggaggt gct gt cgcag gacgccgccg t gt cgggcga cgagggt gg 240  
t gggcaggcc aggt gcagcg gcgcct cggc at ct t ccccg ccaact acgt ggct ccct gc 300  
cgcccggccg ccagccccgc gccgccgcc t cgcggccca gct ccccgt acacgt cgcc 360  
t t cgagcggc t ggagct gaa ggagct cat c ggcgct gggg gct t cgggca ggt gt accgc 420  
gccacct ggc agggccagga ggt ggccgt g aaggcggcgc gccaggacc ggagcaggac 480  
gcggcggcgg ct gccgagag cgt gcggcgc gaggct cggc t ct t cgccat gct gcggcac 540  
ccaacat ca t cgagct gcg cggcgt gt gc ct gcagcagc cgcacct ct g cct ggt gct g 600  
gagt t cgccc gcggcggagc gct caaccga gcgct ggccg ct gccaacgc cgccccggac 660  
ccgcgcgcgc ccggcccccg ccgcgcgcgc cgcacat ccct c cgcacgt gct ggt caact gg 720  
gccgt gcaga t agcgcgggg cat gct ct ac ct gcat gagg aggcct t cgt gcccat cct g 780  
caccgggacc t caagt ccag caacat t t t g ct act t gaga agat agaaca t gat gacat c 840  
t gcaat aaaa ct t t gaagat t acagat t t t gggc t ggcga gggat ggca caggaccacc 900  
aaaat gagca cagcaggcac ct at gcct gg at ggcccccg aagt gat caa gt ct t cct t g 960  
t t t t ct aagg gaagcgacat ct ggagct at ggagc gct gc t gt gggact gct caccgga 1020  
gaagt cccct at cggggcat t gat ggcct c gccgt ggct t at ggggt agc agt caat aaa 1080  
ct cact t t gc ccat t ccat c cacct gccct gagccgt t t g ccaagct cat gaaagaat gc 1140

t ggcaacaag	accct cat at	t cgt ccat cg	t t t gcct t aa	t t ct cgaaca	gt t gact gct	1200
at t gaagggg	cagt gat gac	t gagat gcct	caagaat ct t	t t cat t ccat	gcaagat gac	1260
t ggaaact ag	aaat t caaca	aat gt t t gat	gagt t gagaa	caaagggaaa	ggagct gcga	1320
t cccgggaag	aggagct gac	t cgggcccgt	ct gcagcaga	agt ct cagga	ggagct gct a	1380
aagcggcgt g	agcagcagct	ggcagagcgc	gagat cgacg	t gct ggagcg	ggaact t aac	1440
at t ct gat at	t ccagct aaa	ccaggagaag	cccaaggt aa	agaagaggaa	gggcaagt t t	1500
aagagaagt c	gt t t aaagct	caaagat gga	cat cgaat ca	gt t t acct t c	agat t t ccag	1560
cacaagat aa	ccgt gcaggc	ct ct cccaac	t t ggacaaac	ggcggagcct	gaacagcagc	1620
agt t ccagt c	ccccgagcag	ccccacaat g	at gccccgac	t ccgagccat	acagt t gact	1680
t cagat gaaa	gcaat aaaac	t t ggggaagg	aacacagt ct	t t cgacaaga	agaat t t gag	1740
gat gt aaaaa	ggaat t t t aa	gaaaaaaggt	t gt acct ggg	gaccaaat t c	cat t caaat g	1800
aaagat agaa	cagat t gcaa	agaaaggat a	agacct ct ct	ccgat ggcaa	cagt cct t gg	1860
t caact at ct	t aat aaaaaa	t cagaaaacc	at gccct t gg	ct t cat t gt t	t gt ggaccag	1920
ccagggt cct	gt gaagagcc	aaaact t t cc	cct gat ggat	t agaacacag	aaaaccaaaa	1980
caaat aaaat	t gcct agt ca	ggcct acat t	gat ct acct c	t t gggaaaga	t gct cagaga	2040
gagaat cct g	cagaagct ga	aagct gggag	gaggcagcct	ct gcgaat gc	t gccacagt c	2100
t ccat t gaga	t gact cct ac	gaat agt ct g	agt agat ccc	cccagagaaa	gaaaacggag	2160
t cagct ct gt	at gggg gcac	cgt cct t ct g	gcat cggg gg	ct ct gggact	ggacct caga	2220
gagct t cat a	aagcacaggc	t gct gaagaa	ccgt t gccca	aggaagagaa	gaagaaacga	2280
gagggaat ct	t ccagcgggc	t t ccaagt cc	cgcagaagcg	ccagt cct cc	cacaagcct g	2340
ccat ccacct	gt ggggaggc	cagcagccca	ccct ccct gc	cact gt caag	t gccct gggc	2400
at cct ct cca	cacct t ct t t	ct ccacaaag	t gcct gct gc	agat ggacag	t gaagat cca	2460
ct ggt ggaca	gt gcacct gt	cact t gt gac	t ct gagat gc	t cact ccgga	t t t t t gt ccc	2520
act gccccag	gaagt ggt cg	t gagccagcc	ct cat gccaa	gact t gacac	t gat t gt agt	2580
gt at caagaa	act t gccgt c	t t cct t cct a	cagcagacat	gt gggaat gt	acct t act gt	2640
gct t ct t caa	aacat agacc	gt cacat cac	agacggacca	cgt ct gat gg	aaat ccgacc	2700
ccaact ggt g	caact at t at	ct cagccact	ggagcct ct g	cact gccact	ct gccct ca	2760
cct gct cct c	acagt cat ct	gccaagggag	gt ct caccca	agaagcacag	cact gt ccac	2820
at cgt gcct c	agcgt cgccc	t gcct ccct g	agaagccgct	cagat ct gcc	t caggct t ac	2880
ccacagacag	cagt gt ct ca	gct ggcacag	act gcct gt g	t agt gggg cg	cccaggacca	2940
cat cccacc	aat t cct cgc	t gccaaaggag	agaact aaat	cccat gt gcc	t t cat t act g	3000
gat gct gacg	t ggaaggt ca	gagcagggac	t aact gt gc	c		3041